Time-Consistent Consumption Taxation

Sarolta Laczo & Raffaele Rossi

Working Paper No. 857 April 2018 ISSN 1473-0278

School of Economics and Finance

Queen Mary
University of London

Time-Consistent Consumption Taxation*

Sarolta Laczó† Raffaele Rossi‡

April 2018

Abstract

We characterise optimal tax policies when the government has access to consumption taxation and cannot credibly commit to future policies. We consider a neoclassical economy where factor income taxation is distortionary within the period, due to endogenous labour and capital utilisation and non-tax-deductibility of depreciation. Contrary to the case where only labour and capital income are taxed, the optimal time-consistent policies with consumption taxation are remarkably similar to their Ramsey counterparts. 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.

JEL classification: E62, H21.

Keywords: fiscal policy, Markov-perfect policies, consumption taxation, variable capital utilisation

*We thank Charles Brendon, Davide Debortoli, Andrea Lanteri, Campbell Leith, Yang K. Lu, Albert Marcet, Ricardo Nunes, Evi Pappa, Vito Polito, José Víctor Ríos-Rull, and Maurizio Zanardi. We also thank seminar participants at CREI/UPF, IAE/UAB, Bank of England, University of York, University of Manchester, Lancaster University, University of Surrey and University of Shieffield and participants of the Max Weber June Conference in Florence, the Barcelona GSE Summer Forum 'Macro and Micro Perspectives on Taxation,' CEF in Oslo, PET in Seattle, EEA in Toulouse, MMF in Durham, and Barcelona GSE Winter Workshops for useful comments and suggestions. We received funding from the Spanish Ministry of Science and Innovation under grant ECO2008-04785 and the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 612796. Laczó also acknowledges funding from the JAE-Doc programme co-financed by the European Social Fund. We thank the Institut d'Anàlisi Econòmica (IAE-CSIC) for hospitality while advancing this project.

†Queen Mary University of London, IAE-CSIC, and CEPR, School of Economics and Finance, Mile End Road, London, E1 4NS, United Kingdom. Email: s.laczo@qmul.ac.uk.

‡University of Manchester, Department of Economics, Oxford Road, Manchester, M13 9PL, United Kingdom. Email: raffaele.rossi@manchester.ac.uk.

1 Introduction

Most of the literature on optimal fiscal policy rules out consumption taxation, a policy instrument used in most industrialised economies. For example, as of early 2018, the value-added tax (VAT) on standard items ranges from 17 to 27 percent in European Union countries. The literature on optimal consumption taxation includes Coleman (2000), who finds, under the assumption that the fiscal authority can commit to future policies, that replacing income taxes with consumption taxes would lead to large welfare gains in the United States. Correia (2010) extends this result to a heterogeneous-agents framework. Two recent contributions highlight the role of consumption taxation as a tool to relax a constraint of the monetary authority on the nominal interest rate, either as a result of the zero lower bound (Correia et al., 2013) or in a monetary union (Farhi et al., 2014). Our study finds a new benefit of consumption taxation: time-consistent policies and the resulting allocations are almost identical to those under commitment.

Our results are derived in a neoclassical model with endogenous labour supply and variable capital utilisation.[1] The government has to finance spending on public goods and has access to three types of proportional taxes: capital income, labour income, and consumption taxes. A key element of our baseline environment is that all tax instruments are distortionary within the period. Given endogenous labour, both the consumption tax and the labour tax affect the consumption-leisure choice of households. Given endogenous capital utilisation and that depreciation is not tax-deductible, the capital tax distorts the capital utilisation rate. Lump-sum taxes and debt or asset accumulation are not available to the government. These assumptions serve to avoid trivial solutions to the problem of raising fiscal revenues.[2]

Our key assumptions resulting in a distortionary capital tax within the period are motivated by (i) the fact that in reality capital utilisation is not fixed, and (ii) the fact that in practice depreciation is deductible according to accounting formulae, and is not based on the actual loss of value in capital due to usage. Hence, capital taxes introduce a wedge between the return on capital services and their cost in terms of capital depreciation. To

[1] Greenwood et al. (1988) and many others highlight the importance of taking into account variable capital utilisation when modelling business cycles.

[2] Two recent papers (Debortoli and Nunes, 2013; Debortoli et al., 2017) allow for debt in a Markov-perfect policy setting, but they exclude capital. Coleman (2000) and Correia (2010) allow for debt in a Ramsey context with consumption taxation, and impose either upper limits on tax rates or that tax rates cannot vary over time, in order to avoid a large initial capital levy. Zhu (1995) shows that Ramsey policies can be made time-consistent through public debt restructuring, when debt can be denominated in the after-tax wage and after-tax return on capital. Domínguez (2007) establishes a similar result when there are delays in tax policy implementation, as in Klein and Ríos-Rull (2003). Our results reply on a simpler and existing instrument, proportional consumption taxes.

our knowledge, Zhu (1995) was the first to argue that it is important to consider the capital utilisation margin when studying optimal fiscal policy.

The novelty of our paper lies in analysing time-consistent fiscal policies when the policy-maker has access to consumption taxation, in addition to factor income taxation. The existing literature on Markov-perfect policies, starting with the seminal paper Klein et al. (2008), finds that lack of commitment alters greatly the characteristics of optimal policies and the resulting allocations when capital and/or labour income are taxed. The closest contributions to our study are Martin (2010) and Debortoli and Nunes (2010). These papers study Markov-perfect equilibria in the same environment as we do, but they allow only factor income taxation. They find that optimal time-consistent taxation differs greatly from its Ramsey counterpart, and the tax rates are close to those in the United States.[3]

Our results can be summarised as follows. First, 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 commitment. Optimal taxation involves no tax on capital and taxing consumption and subsidising labour at the same rate, as long as private consumption is larger than labour income, as it is the data. The result comes from the fact that any constant consumption tax rate is non-distortionary with respect to the household's consumption-saving decision.[4] Given that Ramsey policies achieve the first best, they are time-consistent. In other words, the steady states under Ramsey and Markov-perfect policy-making coincide. However, these tax policies include an unrealistically large (several hundred percent) labour subsidy.

Second, we study the case where subsidising labour is prohibited, as in Coleman (2000). In this case, the labour income tax is zero. The Ramsey (Markov) policy-maker taxes consumption at 22.3 (22.1) percent at the steady state in our baseline calibration, and sets the capital income tax to zero (0.4 percent). Looking at the transition from the status quo, optimal consumption and capital tax rates vary little over time and with the level of capital, under both Ramsey and Markov policy-making.[5]

[3] Martin (2010) analyses the case where the capital income tax is non-distortionary within the period as well, and shows that a Markov-perfect equilibrium does not exist given standard calibrations. Below we discuss this case further, and we also solve our model under alternative assumptions on the tax-deductibility of depreciation and on capital utilisation.

[4] Note that the government has only three instruments but faces four constraints, in particular, its own balanced-budget constraint and three optimality conditions of the private sector: consumption-saving, consumption-leisure, and capital utilisation. Then, setting the capital tax to zero leaves both the Euler equation and capital utilisation undistorted, and setting the labour income tax equal to minus the consumption tax leaves the consumption-leisure margin undistorted. Finally, the consumption tax can be chosen to raise sufficient revenue to satisfy the government's budget constraint.

[5] Coleman (2000) finds that the capital tax rate stays at its upper bound (50 or 100 percent) for a few

The intuition behind these results is the following. Firstly, at the steady state, taxing consumption causes only intratemporal distortion, while taxing capital income causes intertemporal distortion as well, which yields the standard result that capital should not be taxed in the long run, under commitment. At the Markov equilibrium and in the initial periods under commitment, the policy-maker optimally taxes already installed capital as much as possible, as it is viewed as a non-distortionary source of revenue. In our environment, firstly, our balanced-budget requirement limits the initial capital levy. Secondly and most importantly, the capital income tax distorts capital utilisation. Martin (2010) and Debortoli and Nunes (2010) show that with only factor income taxation, the Markov policy-maker's desire to tax 'initial' capital dominates, and hence there are large differences in policies and allocations between Ramsey and Markov governments.[6]

On the contrary, taxing consumption partly taxes the initial capital stock, akin to the capital tax and as opposed to the labour tax. Further, a flat consumption tax does not distort intertemporal decisions at all, while the capital tax of all periods except the initial one does. An additional trade-off the government faces is in terms of intratemporal distortions: the capital tax impacts the capital utilisation margin, while the consumption tax distorts the consumption-leisure margin. The latter distortion turns out to be the least important quantitatively in determining optimal policy, while the intratemporal distortion caused by the capital tax plays a key role.

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. With consumption taxation the welfare gains over the existing tax system in the United States are 7.745 (7.744) percent under Ramsey (Markov) policies, while without taxing consumption the gains are 7.06 (4.92) percent. Remarkably, we find higher welfare under discretion when the policy-maker has access to consumption taxation than under commitment when the government can tax only

years and the consumption tax is high initially as well. The key differences between our baseline environment and his is that he (i) does not impose a balanced-budget requirement and (ii) assumes that the capital tax in non-distortionary within the period. Below we discuss in detail the role of the different assumptions.

[6] In our baseline calibration, the Ramsey policy-maker sets the labour income tax to 24.0 and the capital income tax to zero at the steady state. It 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 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 tax rates in the Markov equilibrium are highly sensitive to the Frisch elasticity of labour supply, which equals 3 in our baseline calibration. When this elasticity is equal to 1, we recover the approximately equal labour and capital income tax rates of the literature (Martin, 2010; Debortoli and Nunes, 2010).

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 also the cyclical properties of tax rates and allocations under a Ramsey and a time-consistent policy-maker are very similar.

Our main results are robust to modifying utility parameters such as the Frisch elasticity, assuming that government spending is exogenous, etc. Our results are much weakened if the capital income tax rate is non-distortionary within the period. This happens if depreciation is tax-deductible and/or capital is fully utilised. The gains from commitment with consumption taxation increase to 0.388 (0.633) percent with full (variable) capital utilisation, and the capital tax rate is not close to zero at the Markov steady state (10.3 percent with non-tax-deductible depreciation and capital fully utilised, 23-24 percent with tax-deductible depreciation), while consumption is taxed at approximately 16 percent. Martin (2010) has shown that a Markov equilibrium does not exist for standard calibrations with only factor income taxation when the capital tax is non-distortionary. Therefore, taxing consumption ensures that a Markov equilibrium exists in such cases as well.

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

2 The model

The economy is populated by a representative household, a representative firm, and a utilitarian policy-maker. The household decides on consumption, saving, leisure, and the capital utilisation rate. The firm operates in perfectly competitive markets, maximises profits, and uses capital services and labour as production inputs. The policy-maker spends on public consumption which yields utility to households, and raises revenues via proportional taxes on labour income, capital income, and consumption. Depreciation, which depends on the capital utilisation rate, is not tax-deductible. Lump-sum taxes are not available, and the government has to balance its budget in each period. Time is discrete.

A few comments are in order about our main assumptions before describing the economic environment in mathematical terms. First, endogenous capital utilisation (see Greenwood et al., 1988, Greenwood et al., 2000, and many others) and non-tax-deductible depreciation imply that taxing capital is distortionary within the period. Zhu (1995) was the the first to argue that variable capital utilisation should be taken into account when analysing optimal fiscal policy. Martin (2010) shows, allowing only factor income taxation, that a Markov-perfect equilibrium does not exist when the capital tax is non-distortionary within the period, see also Debortoli and Nunes (2010). Variable capital utilisation is in line with the fact that capital is not fully utilised in reality and is a standard assumption in bigger Dynamic Stochastic General Equilibrium (DSGE) models. Further, in reality depreciation which is tax-deductible does not depend on the actual capital utilisation rate, instead it is given by accounting rules.[7,8] To understand the role of these assumptions, we solve our model under alternative assumptions as well. In particular, we assume that depreciation is tax-deductible and/or capital is fully utilised, so that the capital tax is no longer distortionary within the period. See Section 4.3.3 and Appendix C.

Second, we assume that the government operates under a balanced-budget rule. We do this 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 and Ríos-Rull, 2003; Ortigueira, 2006; Klein et al., 2008; Azzimonti et al., 2009; Martin, 2010; Debortoli and Nunes, 2010). Second, while Debortoli and Nunes (2013) and Debortoli et al. (2017)) allow for debt in a Markov-perfect policy setting, they exclude capital. We leave the study of Markov-perfect policies in an environment with both capital and government debt to future work, and focus on the trade-offs between distortions generated by different tax instruments with and without commitment in this paper. The seminal paper of Coleman (2000) on Ramsey policies with consumption taxation does not impose a balanced-budget requirement, and imposes upper bounds on the tax rates instead.[9] Below we compare are results to his, in order to highlight the role our balanced-budget requirement plays.

Finally, we consider government spending to be a choice variable of the fiscal authority,

[7]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.

[8]Instead of endogenous capital utilisation, Klein and Ríos-Rull (2003) and Mateos-Planas (2010) assume that the capital income tax is chosen one or more periods in advance. In this way the current government internalises the distortionary 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.

[9]Without any constraint, the government could set the initial capital tax to hundreds of percent and accumulate assets to then pay for public spending from the return on its assets. In this way it could fully solve the public finance problem, hence there would no policy trade-offs – a setting of limited interest theoretically or practically.

following the literature on Markov-perfect policies (e.g., Klein et al., 2008). At least part of government consumption can be adjusted in response to changes in the economy, and we are interested in studying the optimal policy mix both on the revenue and the spending side. We have repeated our analysis with exogenous government spending as a robustness check, and the results are unaltered.

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

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

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

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

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

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

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

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

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

It is straightforward to derive a standard Euler equation,

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

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

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

The optimal rate of capital utilisation is where the marginal benefit from utilising more capital in terms of after-tax income equals its marginal cost in terms of higher depreciation. Equation (6) implies that capital income taxation is distortionary within the period.[10]

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

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

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

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

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

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

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

The resource constraint in this economy is

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

[10] If actual depreciation were tax-deductible, i.e., the capital tax were levied on $\left[r\left(\mathbf{a}^t\right)v\left(\mathbf{a}^t\right)-\delta\left(v\left(\mathbf{a}^t\right)\right)\right]k\left(\mathbf{a}^{t-1}\right)$, then $\delta_v\left(\mathbf{a}^t\right)=r\left(\mathbf{a}^t\right)$ would hold, that is, the capital tax would not distort capital utilisation. We analyse this case as well below.

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

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

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

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

The characterisation of the first best is presented in Appendix A.1.

We can define competitive equilibria as follows.

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

(i) private sector optimisation taking government policies and prices as given, that is,

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

(ii) market clearing, (10), and

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

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

3 The policy problems

Both with and without commitment, the policy-maker maximises the household's lifetime utility over competitive equilibria. We assume, following most of the literature, that the policy-maker moves first in each period. We use a version of the primal approach, i.e., we write the policy problems in terms of allocations and substitute for prices and tax rates. We also substitute for output to simplify. However, we keep the consumption tax rate as a decision variable along with the allocations. This will be useful when constraining the tax rates.[11]

First, one can eliminate three variables, output $y\left(\mathbf{a}^t\right)$ and prices $w\left(\mathbf{a}^t\right)$ and $r\left(\mathbf{a}^t\right)$, and three equations, (7), (8), and (9), in the definition of competitive equilibria, Definition 2. Second, the government's budget constraint and the resource constraint jointly imply that the household's budget constraint, (3), holds. Then six conditions are left which characterise competitive equilibria. Third, one can use the household's consumption-leisure optimality condition and the government's budget constraint to express the labour and capital income tax rates. The details of these derivations are in Appendix A.2. Then there remain four constraints: (2) and, $\forall a^t$,

$$c\left(\mathbf{a}^t\right)+g\left(\mathbf{a}^t\right)+k\left(\mathbf{a}^t\right)=a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}+\left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right), \tag{12}$$

$$\begin{aligned}\frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)}=\beta\mathbb{E}_t\Bigg(&\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)}\Bigg[1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)+a_{t+1}v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\\&-\frac{g\left(\mathbf{a}^{t+1}\right)-\tau^c\left(\mathbf{a}^{t+1}\right)c\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\Bigg]-u_\ell\left(\mathbf{a}^{t+1}\right)\frac{h\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\Bigg),\end{aligned} \tag{13}$$

$$\begin{aligned}\delta_v\left(\mathbf{a}^t\right)=a_t&\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}-\frac{g\left(\mathbf{a}^t\right)-\tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\\&-\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)}\left(1+\tau^c\left(\mathbf{a}^t\right)\right)\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}.\end{aligned} \tag{14}$$

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

$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)}\leq\frac{1}{1+\tau^c\left(\mathbf{a}^t\right)}\left(1-\gamma\right)a_t\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma}. \tag{15}$$

Below we write the policy problems in a general form including the constraint (15). We will, however, also study the case where (15) is ignored and the case without consumption taxation, i.e., $\tau^c\left(\mathbf{a}^t\right)=0$, $\forall\mathbf{a}^t$, to compare our results with the existing literature.

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

3.1 The Ramsey policy-maker's problem

The Ramsey policy-maker maximises (1) choosing consumption tax rates $\{\tau^c(\mathbf{a}^t)\}_{t=0}^{\infty}$ and allocations $\{g(\mathbf{a}^t), c(\mathbf{a}^t), \ell(\mathbf{a}^t), h(\mathbf{a}^t), k(\mathbf{a}^t), v(\mathbf{a}^t)\}_{t=0}^{\infty}$, subject to (2), (12), (13), (14), and (15), $k(\mathbf{a}^0)$ and the productivity process given. We assign the Lagrange multipliers $\lambda_1(\mathbf{a}^t), ..., \lambda_5(\mathbf{a}^t)$ to the five constraints, respectively.

Note that future decision variables enter into the household's current Euler equation, (13), hence the Ramsey problem is not recursive using only the natural state variables, a and k. Following Marcet and Marimon (1998/2017), the Lagrange multiplier on the Euler equation, $\lambda_3(\mathbf{a}^t)$, with $\lambda_3(\mathbf{a}^0) = 0$, can be introduced as a co-state variable to write a Bellman equation. The Ramsey problem can then be solved numerically by standard policy function iteration. The value function and the policy functions are time-invariant on the extended state space, the current value of which is denoted (a, k, λ_3). Next period's productivity a' is given exogenously. The policy-maker chooses the functions $\mathcal{K}'()$ and $\Lambda_3'()$, as well the policy functions for the control variables, i.e., $\mathcal{T}^c()$, $\mathcal{C}()$, $\mathscr{L}()$, $\mathcal{H}()$, $\mathcal{G}()$, $\mathcal{V}()$, $\Lambda_1'()$, $\Lambda_2'()$, $\Lambda_4'()$, and $\Lambda_5'()$, and the value function $\mathcal{W}()$.

Let unindexed variables denote the values of the policy functions for the control variables at this state, i.e., $c = \mathcal{C}(a, k, \lambda_3)$, and so on; and $k' = \mathcal{K}'(a, k, \lambda_3)$, $\lambda_3' = \Lambda_3'(a, k, \lambda_3)$, and $W = \mathcal{W}(a, k, \lambda_3)$. Then we can write

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

$\lambda_5 \geq 0$, with complementary slackness conditions.[12] Appendix A.3 presents the first-order conditions of the Ramsey policy-maker's 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

[12] Note that we have assumed from the beginning that the time constraint, (2), and the resource constraint, (12), will bind, and we know that $\lambda_1 > 0$ and $\lambda_2 > 0$.

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. It also takes into account the optimising behaviour of the private sector. Note that, as under Ramsey policy-making, the fiscal authority moves first in every period, and commits within the period.[13]

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 et al., 2008). In a stationary Markov-perfect equilibrium, all governments employ the same policy rules. Hence, the rules must satisfy a fixed-point property: if the current policy-maker anticipates that all future governments will follow the policy rules $\{\mathcal{T}^c(a,k), \mathcal{C}(a,k), \mathscr{L}(a,k), \mathcal{H}(a,k), \mathcal{G}(a,k), \mathcal{V}(a,k), \mathcal{K}'(a,k)\}$, and similar rules for the Lagrange multipliers, then it finds it optimal to follow the same rules.

Let $\mathcal{U}_c() = u_c(\mathcal{C}(), \mathscr{L}(), \mathcal{G}())$, and similarly for $\mathcal{U}_\ell()$. Then we have

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

$\lambda_5 \geq 0$, with complementary slackness condition.[14] Appendix A.4 presents the first-order conditions of the time-consistent policy-maker's problem.

[13]See Ortigueira (2006) on the importance of the assumption on the intratemporal timing of actions.

[14]It can be useful to think about the policy problems and the primal approach in the terminology of oligopoly. Both in the Ramsey and the Markov problems, given that the policy-maker moves first in every period, it is a Stackelberg leader and the private sector is the follower. Therefore, one can use the private sector's optimality conditions to simplify the problem, as is done according to the primal approach. In the Markov problem there is strategic interaction between the policy-makers of different time periods as well. Today's policy-maker is a Stackelberg leader when interacting with tomorrow's policy-maker, as the former clearly has a first-mover advantage. Therefore, one can plug in the policy functions of tomorrow's policy-maker into today's policy-maker's problem, as we have done in this section.

3.3 Analytical results

We present analytical results for the steady state (i) in the case where tax rates are unrestricted and (ii) excluding a labour subsidy. In both cases we consider an economy without productivity shocks, i.e., we set $a = a' = 1$. We assume that private consumption is larger than labour income, as in the data. We also consider cases where depreciation is tax-deductible and/or capital is fully utilised, i.e., $v = v' = 1$. Importantly, under these assumption, the capital income tax is no longer distortionary within the period. The details of the model with tax-deductible depreciation are in Appendix C. Our analytical results characterise steady states.[15]

Result 1. *Assume that the government has access to labour income, capital income, and consumption taxation, and all tax rates are unrestricted. Then the Ramsey steady state with* $\tau^c > 0$, $\tau^c = -\tau^h$, *and* $\tau^k = 0$ *corresponds to the first best, and hence it is time-consistent. This holds in our basline model, when capital is fully utilised, and when depreciation is tax-deductible.*

Proof. In Appendix B. □

However, the required tax rates are larger than 100 percent in absolute value given the relatively small difference between the consumption and the labour income tax bases in reality, see our quantitative analysis below. See also Coleman (2000). We now turn to the case where a labour subsidy is excluded.

Result 2. *When* $\tau^h \geq 0$ *is imposed, the Ramsey policy-maker taxes only consumption at the steady state. This holds in our baseline model, when capital is fully utilised, and when depreciation is tax-deductible.*

Proof. In Appendix B. □

Coleman (2000) proves similar results in a framework without a balanced-budget requirement and with capital fully utilised and depreciation tax-deductible. In that case and allowing for a labour subsidy, as in Result 1, constant taxes can be set over the transition period as well, hence the Ramsey planner can implement the first best in all periods.

[15] 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.

4 Quantitative analysis

We now turn to numerical methods and solve a calibrated version of our economy, to assess quantitatively the optimal fiscal policy mix and welfare with and without consumption taxation and with and without commitment.

4.1 Calibration

We consider the model period to be a year and specify the utility function as

$$u(c, \ell, g) = \log(c) - \alpha_\ell \frac{(1-\ell)^{1+1/\varphi}}{1+1/\varphi} + \alpha_g \log(g), \tag{16}$$

where φ is the (constant) Frisch elasticity of labour supply, while α_ℓ and α_g are the weights of leisure and public goods relative to private consumption, respectively. Given an intertemporal elasticity of substitution equal to 1, we set the Frisch elasticity of labour supply, φ, equal to 3, as in Trabandt and Uhlig (2011).[16] We assume that the depreciation rate is an increasing and convex function of capital utilisation, following Greenwood et al. (1988) and many others. That is, $\delta(v) = \eta v^\chi$, with $\eta > 0$ and $\chi > 1$. Finally, we assume that aggregate productivity follows an AR(1) process with persistence parameter ρ and standard deviation of the shock σ_a.

To pin down β, γ, α_ℓ, η, and χ, we use the private sector's first-order conditions and the resource constraint at steady state to match average macroeconomic ratios from United States data for the period 1996-2010.[17] We take average capacity utilisation for all industries from the Federal Reserve Board, and we compute all other macroeconomic ratios using data provided by Trabandt and Uhlig (2012).[18] The private sector takes as given the effective tax rates. 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.[19] We target the average labour income share (60.9 percent), private consumption[20] over GDP (69.6 percent), public

[16]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. We check the robustness of our results to a wide range of values of φ, 0.4 to 5, see below.

[17]We choose this relatively short time period to better proxy current 'steady-state' macro ratios, as some, notably the labour income share, have changed substantially over time.

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

[19]We convert it to a capital tax rate without depreciation allowance, in line with our model, taking revenue from capital income taxation as given. That is, $\tau^k_\delta (rv - \delta(v)) = \tau^k rv = \tau^k \gamma \frac{y}{vk}$. This gives $\tau^k = 0.253$.

[20]In order to properly account for the different tax bases, our definition of consumption includes all products and services that are subject to VAT, i.e., non-durables, durables, and services. This is in line with Coleman (2000). However, we have also considered the alternative assumption of assigning durables to investment as a robustness check. The results, available upon request, are similar for this alternative definition of consumption.

consumption over GDP (15.5 percent), capital over GDP (2.349), and the fraction of time worked for the working age population (24.9 percent[21]). To calibrate α_g, we assume that g found in the data is optimally chosen in the sense that $u_c = u_g$. Finally, we calibrate the AR(1) coefficients of the technological progress in a standard way.[22]

The calibrated parameter values are presented in Table 1.

TABLE 1: Calibrated parameters

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

Note that we have not taken into account the household's and the government's budget constraint. In reality tax revenues are raised not only to finance public consumption, but also in order to redistribute resources from richer to poorer households. At the status-quo steady state, tax revenues are higher by 11.7 percent of GDP than public consumption. In order to satisfy the budget constraints, one can imagine that the government gives a lump-sum transfer of 11.7 percent of GDP to the representative household. Viewed through the lens of a representative-agent model, this is a source of inefficiency and will imply additional welfare gains for all optimal tax reforms.

4.2 Solution method

First, we solve the Ramsey problem using policy function iteration. This consists of the following steps. We discretise the state variables $k \in [\underline{k}, \overline{k}]$ and $\lambda_3 \in [\underline{\lambda}_3, \overline{\lambda}_3]$. In the stochastic case, we approximate the estimated AR(1) process by a 3-state Markov chain

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

[22]We match the unconditional persistence and standard deviation of total output in our economy with fixed tax rates to those of the de-trended US GDP for the period $1996 - 2010$. Note that while the value of the persistence parameter is lower than usual in the RBC literature as a whole, it is in line with calibrated RBC models with variable capital utilisation such as Greenwood et al. (2000).

following Galindev and Lkhagvasuren (2010) and Kopecky and Suen (2010).[23] Once we have found the endogenous collocation nodes, we guess the policy functions at each grid point. At each iteration we solve the system of non-linear equations at each grid point, and we approximate globally the policy functions of next period using cubic splines.[24]

Second, using the solution to the Ramsey problem by policy function iteration as initial guess, we solve it again parameterising the policy functions using cubic splines. Then, we solve the time-consistent policy-maker's problem in the same way.[25] In this solution algorithm we iterate until the parameters of the policy functions converge to high accuracy. The resulting policy functions are well behaved. We use this algorithms to simulate the dynamic paths of tax policies and allocations.[26]

4.3 Results

In this section we present the results for our baseline model, where the capital income tax is distortionary within the period, as well as results under alternative assumptions. We describe the steady state and the transition below, and we discuss our results with productivity shocks in Appendix D.

4.3.1 Steady state

Table 2 shows the allocations and the tax rates at steady state for five policy models.[27] 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.

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

$$\Pi = \begin{bmatrix} 0.655 & 0.308 & 0.036 \\ 0.154 & 0.692 & 0.154 \\ 0.036 & 0.308 & 0.655 \end{bmatrix}.$$

[24]Chebyshev polynomials work as well.

[25]The derivatives of next period's policy functions with respect to the endogenous state k' are computed using the Compecon Matlab package by Fackler and Miranda (2004).

[26]We have solved the time-consistent policy-maker's problem by policy function iteration as well, using the Ramsey solution as initial guess. The resulting tax rates and allocations are identical up to many decimals for the two methods for both the Ramsey and Markov problems.

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

Table 2: Tax rates and allocations at steady state

Variable	unrestricted	$\tau^h \geq 0$ Ramsey	$\tau^h \geq 0$ Markov	$\tau^c = 0$ Ramsey	$\tau^c = 0$ Markov
Consumption tax rate	3.245	0.223	0.221	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.065
Capital income tax rate	0.000	0.000	0.004	0.000	0.198
Capital	1.801	1.548	1.539	1.467	1.106
Hours worked	0.320	0.275	0.277	0.261	0.283
Income	0.572	0.492	0.491	0.466	0.439
Consumption-income ratio	0.654	0.654	0.652	0.654	0.723
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. In this case, the Ramsey policy-maker (Column 2) taxes consumption at 22.3 percent at the steady state and sets the labour and capital income taxes to zero.[28] 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. The policy-maker uses the consumption tax, because it is less distortionary.

Taxing consumption is less distortionary than taxing labour for the following reasons. First, note that, at the steady state, neither tax distorts intertemporal decisions, only intratemporal ones. The consumption-leisure margin is distorted by the tax wedge $\xi \equiv \frac{1-\tau^h}{1+\tau^c}$, and $\xi \leq 1$ as long as tax rates are non-negative. It is easy to see that for any tax rate $\tilde{\tau} > 0$, ξ is closer to 1 when consumption rather than labour income is taxed. That is, the same consumption tax distorts the consumption-leisure margin less than the labour income tax. Note also that the difference in ξ increases with the tax rate $\tilde{\tau}$. Relatedly, a 100 percent labour tax would imply that the economy shuts down, or, the labour tax Laffer curve peaks below 100 percent, while the economy would still function with a consumption tax of 100 percent.[29] In addition, as long as private consumption is larger than labour income as a share of GDP, as in the data and in our model, then raising a given amount of public revenue requires a lower consumption tax rate than labour tax rate, which magnifies the difference between the two instruments when it comes to distorting the household's consumption-leisure choice.

[28] 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.

[29] See Trabandt and Uhlig (2011) for more details on the labour and consumption tax Laffer curves.

Without consumption taxation, the Ramsey policy-maker (Column 4) taxes only labour income at the steady state (the Chamley-Judd result), 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. However, the tax rates strongly depend on the Frisch elasticity of labour supply ($\varphi = 3$ in our baseline calibration) in this case, while not in the other policy scenarios.[30]

Turning to the allocations, firstly, at the Ramsey steady states and at the first best, the consumption-income ratio and the public spending-income ratio are the same. This is due to log utility of consumption.[31] Secondly, comparing Ramsey and Markov steady states without consumption taxation, Markov policies imply significantly lower long-run capital and income, a higher consumption-income ratio, and a lower public spending-income ratio, which are all due to more distortions caused by taxation.[32] 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 almost as high as at the Ramsey steady state. Notably, they are higher without commitment but with consumption taxation than under Ramsey policy but taxing only labour and capital income. The consumption- and public spending-income ratios change very little as a result of the change in commitment. To summarise, the most striking feature of tax rates and allocations at steady state is that the Ramsey and Markov policies and allocations are very similar with consumption taxation.

Table 2 also shows the per-period utilities and the welfare-equivalent consumption losses compared to the first best at the steady state.[33] With consumption taxation and imposing $\tau^h \geq 0$ (without consumption taxation), the steady-state welfare-equivalent consumption losses amount to 6.1 (18.4) percent in the case of time-consistent policy and to 5.9 (8.7)

[30] We have also considered (i) $\varphi = 0.4$, which is in line with recent micro estimates such as Domeij and Floden (2006) (see also Guner et al., 2012), (ii) $\varphi = 1$, which is often chosen in the macro literature (e.g. Christiano et al., 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 et al., 2007). We adjust α_ℓ appropriately in each case as described in Section 4.1. As φ increases from 0.4 to 5, τ^h decreases from 17.3 to 4.2 percent and τ^k increases from 7.8 to 22.1 percent. This is because increasing the elasticity of labour supply increases the distortion cause by τ^h compared to τ^k. With $\varphi = 1$ we recover the result of the existing literature (Martin, 2010; Debortoli and Nunes, 2010) that the two income tax rates are near-equal at the Markov steady state. The results are available upon request.

[31] See Motta and Rossi (2018).

[32] The result that the public spending-income ratio is lower under Markov policy was first noted by Klein et al. (2008) in an environment where only labour income taxes are available.

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

percent under Ramsey. 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.[34]

We have verified that our conclusions are robust to various parameter changes.[35] Our conclusions remain unaltered in all cases. We have also verified whether our results still hold when public consumption, g, is exogenous rather than endogenous. Remarkably, under all parameterisations we considered and irrespective of whether government spending is endogenous or exogenous, taxing consumption is more important than being able to commit. That is, welfare is always higher at the Markov steady state with consumption taxation than under commitment without taxing consumption. The results are available upon request.

4.3.2 Dynamics

In this section, we study whether our main results on the usefulness of consumption taxation in terms of mitigating the commitment problem of the policy-maker and improving welfare hold once we take into account transitional dynamics. Further, we aim to shed light on the key trade-offs faced by the policy-makers during the transition. In addition, we quantify for different taxation and/or commitment scenarios (i) the welfare gains from the status quo, (ii) the gains from commitment, and (iii) the gains from taxing consumption.

In order to do this, we perform the following policy exercise. Initially the economy is at the status-quo steady state. In period 1 a new policy-maker enters into office. It can either be a Markov policy-maker or have access to a commitment technology. Figures 1 and 2 show

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

[35] In addition to different values for the Frisch elasticity of labour supply φ, see above, we have considered (i) a coefficient of relative risk aversion for private consumption equal to 2 (in this case we have recalibrated 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), (ii) $\beta = 0.96$, the most-commonly used value in the macro literature for yearly models, (iii) $\chi = 1.8$ 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), and (iv) $\alpha_g = 0.3$, which reduces the consumption tax base compared to the baseline.

the dynamics of the tax rates and the allocations without and with access to consumption taxation, respectively, for both Ramsey and Markov policy-makers.[36]

The key feature of the results presented in Figures 1 and 2 is that with access to consumption taxation the whole dynamic paths of taxes and allocations hardly differ with and without commitment, while they differ substantially when consumption is not taxed, as in Martin (2010) and Debortoli and Nunes (2010).

The intuition behind these results is the following. First, consider the case without consumption taxation. A high capital tax rate in the first few periods is partly levied on the initial capital stock, which is a non-distortionary way to raise revenue. Our balanced-budget assumption and the intratemporal distortion caused by the capital tax limit the initial capital levy in our environment, but it is still the key driver of policies in the first few years. The Ramsey planner takes into account that the capital income tax in all periods except period 1 encourages people to consume more in earlier periods, reducing investment and growth. The desire of the policy-maker not to distort saving is what drives long-run tax policies under commitment, as is well known. In our environment with endogenous labour and capital utilisation, the government balances the above considerations with the distortion the capital income tax creates for the capital utilisation rate and the labour income tax's impact on the consumption-leisure margin. The contemporaneous distortion of the capital tax modifies the calculation of the Ramsey government only marginally compared to a standard setting with capital fully utilised. The policy-maker raises the capital income tax less in the initial period and makes a greater use of the labour income tax.

The time-consistent policy-maker differs in not internalising the impact of the current capital income tax on past consumption-saving decisions. Therefore, the desire to tax 'initial' capital plays a key role in determining policy. The Markov policy-maker balances this consideration with only the intratemporal distortions caused by the two tax instruments. This leads to dramatic differences in policies and allocations between Ramsey and Markov governments in this setting, see also Martin (2010) and Debortoli and Nunes (2010).

On the contrary, taxing consumption partly taxes the initial capital stock as well. Further, a time-constant consumption tax does not distort intertemporal decisions, while the capital tax of all periods except period 1 does. Further, in our framework with endogenous capital utilisation, the initial capital tax distorts capital utilisation, while the consumption tax distorts the consumption-leisure margin. This last distortion turns out to be the least im-

[36] Figures representing the dynamics of tax policies and allocations starting from the Markov steady states with and without consumption taxation, and from steady states without consumption taxation are available upon request.

Figure 1: Ramsey and Markov policies without consumption taxation starting from the status quo

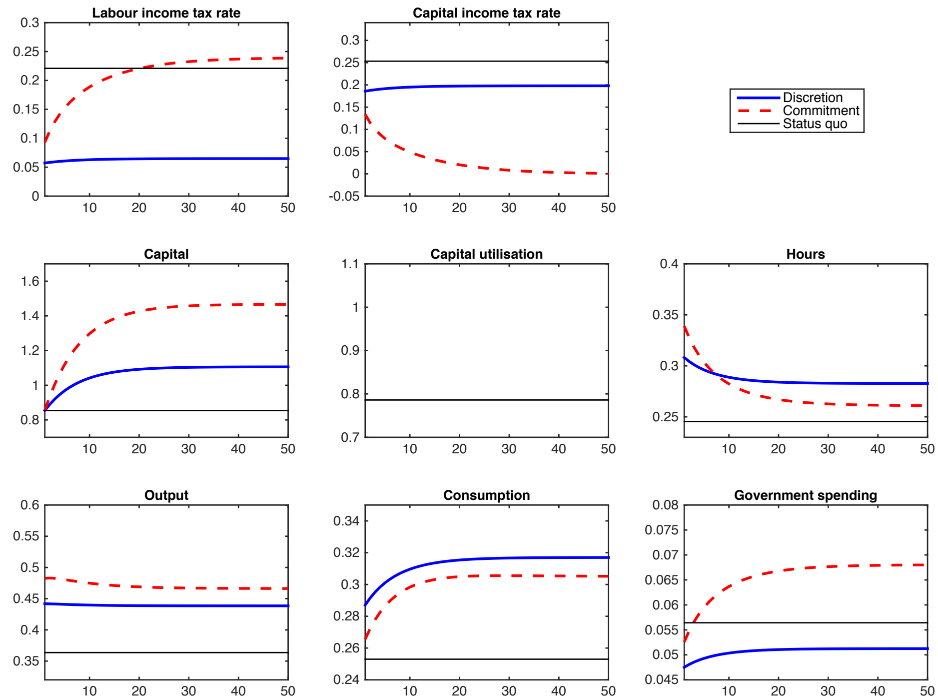
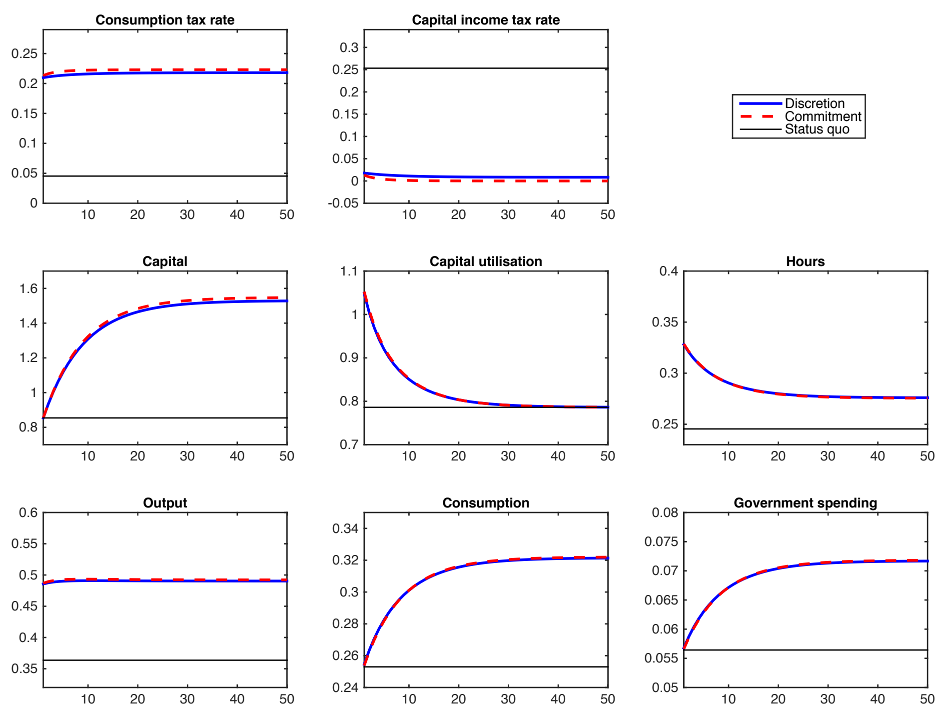


Figure 2: Ramsey and Markov policies with consumption taxation and no labour subsidy starting from the status quo



portant quantitatively in determining optimal policy. The distortion the capital income tax causes within the period combined with fact that the consumption tax is partially levied on initial capital imply that it is optimal to raise most public revenue from taxing consumption, including in period 1. Then, the time-inconsistency features of policies under commitment are negligible, and the intuition for the Markov equilibrium with consumption taxation follows from the Ramsey case. At the Markov equilibrium in every period the trade-offs are as in the initial period of the Ramsey equilibrium, hence fiscal revenue is raised mainly from taxing consumption by the Markov government as well in all periods.

Using our simulation results above, one can compute the welfare gains in terms of welfare-equivalent consumption from the different taxation and commitment scenarios compared to the existing tax system. We assume that initially the economy is at the status-quo steady state, described in Section 4.1. At time 1, a new policy-maker enters into office. It 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.

We also quantify the welfare gains from commitment both with and without consumption taxation, starting from the corresponding Markov steady state. The welfare gains from commitment are 0.0003 percent with consumption taxation and 2.01 percent taxing labour income instead. Hence, the gains from commitment are negligible with access to consumption taxation, while they are substantial without. Finally, the gains from taxing consumption without (with) commitment are 2.77 (1.21) percent, taking into account the transition from the Markov (Ramsey) steady state without consumption taxation. Hence, the welfare gains from taxing consumption are much larger under discretion than under commitment.

Table 3 summarises our welfare results including transitions.

TABLE 3: Welfare gains in welfare-equivalent consumption (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

4.3.3 On the tax-deductibility of depreciation, variable capital utilisation, and budget balance

We now study the role of three key assumptions adopted in our model, namely, (i) non-tax-deductibility of depreciation, (ii) variable capital utilisation, and (iii) the balanced-budget requirement imposed on the government. First of all, we analyse the case where depreciation is not tax-deductible and capital is fully utilised. To do this, we set $v = 1$ and $\lambda_4 = 0$ in all periods. Second, we consider the case where depreciation is tax-deductible and capital utilisation is endogenous. Third, we assume that depreciation is tax-deductible and capital is fully utilised, as in Klein et al. (2008).[37] In all three cases the capital income tax is no longer distortionary within the period. Further, note that the key difference between the first and third alternative scenarios is that the capital income tax base is a lot smaller when depreciation is tax-deductible. Appendix C contains the model setup and the Ramsey and Markov policy problems and their first-order conditions with tax-deductible depreciation.

The key remaining difference between our third alternative environment and that of Coleman (2000) is that he does not impose a balanced-budget requirement. Therefore, we are able to highlight the role of this assumption. Solving our policy problems with both capital and government debt is beyond the scope of this paper, as well as of the literature on Markov-perfect policies, to our knowledge.

First, we discuss the case where the policy-maker can tax labour and capital income but not consumption. If the labour income tax rate has to be non-negative, then only capital is taxed at the Markov solution in all periods (Klein et al., 2008). Martin (2010) shows that no Markov-perfect equilibrium exists for standard calibrations when the capital income tax is non-distortionary within the period and labour can be subsidised. See also Debortoli and Nunes (2010). More precisely, for any given rate of time preference, a Markov-perfect equilibrium exists only if the preference for the public good is low enough, much lower than in standard calibrations. This is because the Markov government wants to tax capital and subsidise labour, and ends up confiscating the whole capital stock, hence the economy shuts down (Martin, 2010).

We now turn our attention to the case where the policy-maker can tax consumption, and solve for both Ramsey and Markov equilibria. Table 4 presents steady-state results with consumption taxation under three sets of alternative assumptions.

Figures 3 and 4 present the dynamic paths of policies with tax-deductible depreciation, with endogenous capital utilisation and full capital utilisation, respectively. We omit the

[37] Martin (2010) analyses this case as well, alongside variable capital utilisation, without consumption taxation.

Table 4: Tax rates and allocations at steady state, alternative assumptions

Variable	non-tax-ded. δ, $v = 1$		tax-ded. δ, variable v		tax-ded. δ, $v = 1$	
	Ramsey	Markov	Ramsey	Markov	Ramsey	Markov
Consumption tax rate	0.223	0.160	0.223	0.156	0.223	0.160
Capital income tax rate	0.000	0.103	0.000	0.231	0.000	0.236
Capital	1.215	1.031	1.548	1.144	1.215	1.031
Hours worked	0.288	0.292	0.275	0.287	0.288	0.292
Income	0.505	0.478	0.492	0.473	0.505	0.478
Consumption-income ratio	0.618	0.638	0.654	0.654	0.618	0.638
Public spending-income ratio	0.138	0.142	0.146	0.146	0.138	0.142

dynamic paths for the case with tax-deductible depreciation and capital fully utilised, as only the values of the capital tax rate differ from Figure 4, as at steady state, see Table 4. Further, the results are similar in all three scenarios, therefore we discuss them jointly. That is, the crucial assumption is whether the capital income tax is distortionary within the period.

We first discuss the Ramsey equilibria. The key difference from our baseline model is a higher capital tax in the first few years. This is because no intratemporal distortion caused by capital taxation is present, which creates an incentive for the Ramsey planner to tax capital at a higher rate initially. The government uses consumption taxes as well from the beginning, due to the fact that this tax instrument is partly levied on the initial capital stock as well, and an increasing path of consumption taxes discourages saving. Capital income taxes converge to zero, while consumption taxes increase during the transition to satisfy the government's budget constraint. The paths of allocations are similar to those in our baseline model.

The key difference between the scenario presented in Figure 4 and the setting of Coleman (2000) is that we assume that the government has to balance its budget. Coleman (2000) imposes upper bounds on the capital income tax rate (0.5 or 1) instead, to avoid trivial solutions to the public finance problem. Without a balanced-budget requirement and given an upper bound on the capital tax, it is optimal to impose a higher consumption tax initially to tax initial capital as much as possible.[38]

Turning to the Markov solutions, first of all, Markov equilibria exist, as opposed to without consumption taxation and allowing for a labour subsidy. This follows from what happens at the Ramsey equilibrium in the initial period. Consumption is taxed rather than subsidised optimally, and the capital income tax is not confiscatory, for our standard calibration. The same is true in each period when policies are time-consistent.

[38] With an upper bound of 100 percent on the capital income tax, Coleman (2000) finds an initial consumption tax of 117 percent.

FIGURE 3: Ramsey and Markov policies with consumption taxation and no labour subsidy starting from the status quo, with depreciation allowance

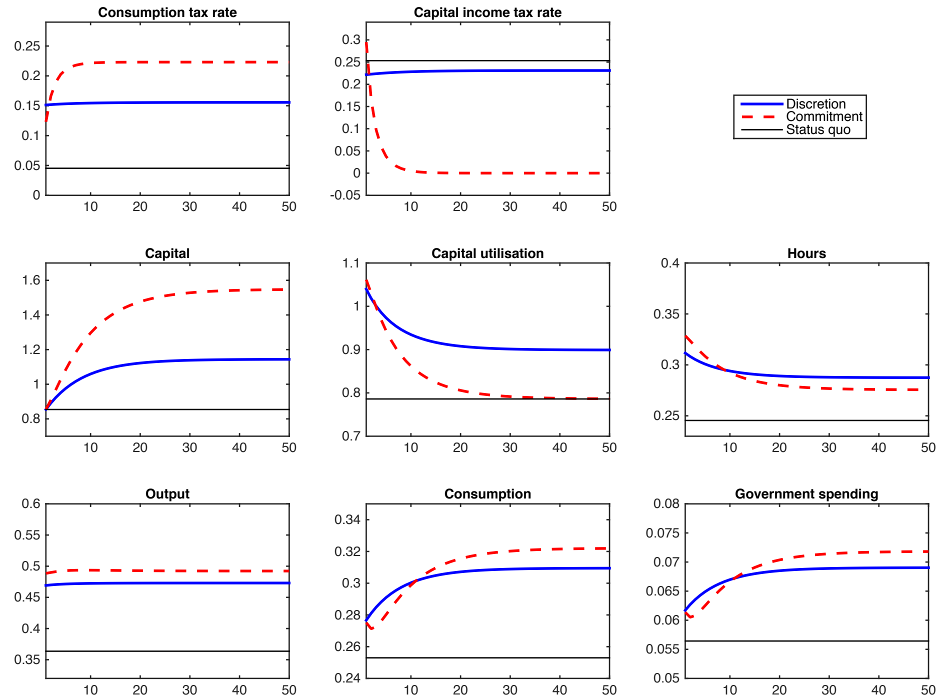
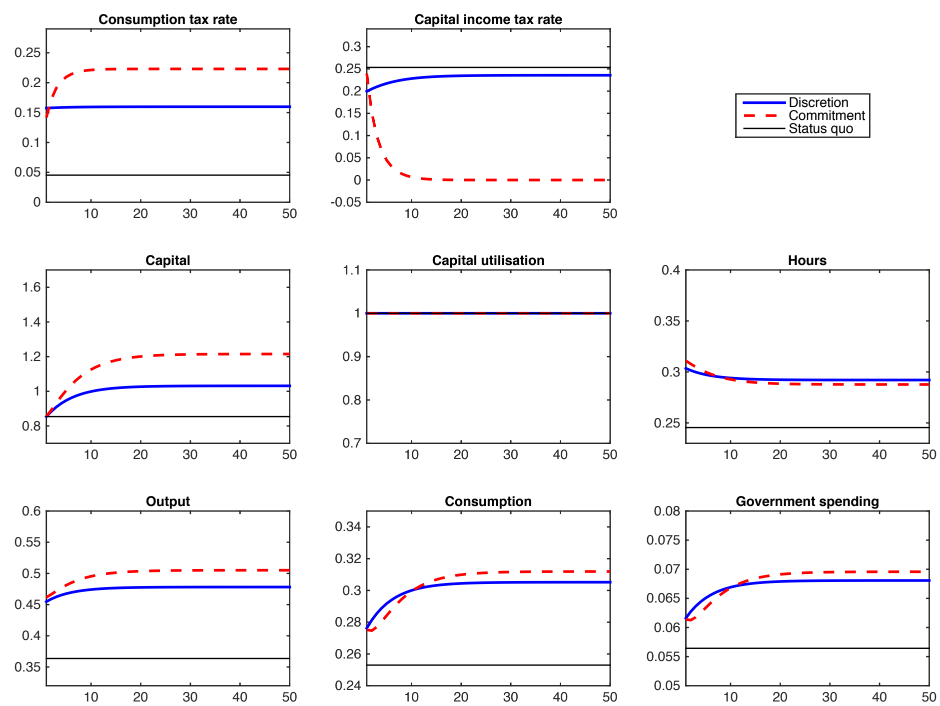


FIGURE 4: Ramsey and Markov policies with consumption taxation and no labour subsidy starting from the status quo, with depreciation allowance and full capital utilisation



Contrary to our baseline model, the capital income tax rate is not close to zero, instead it is approximately 10 percent with non-tax-deductible depreciation and capital fully utilised, while it is between 22 and 23 percent with tax-deductible depreciation. The difference in the capital tax rate depending on the tax-deductibility of depreciation is due to the difference in the tax base. Consumption is taxed at approximately 16 percent in all periods in all three scenarios. As such, this tax remains an important source of revenues. These policies significantly inhibit capital accumulation. When capital utilisation is endogenous, its rate is significantly higher at the Markov equilibrium than the Ramsey.

Finally, we highlight the welfare gains from commitment and the gains compared to the existing tax system for all three alternative environments, see Table 5. Our baseline results are much weakened when depreciation is tax-deductible and/or capital is fully utilised. The value of commitment is 0.633 (0.338) percent with endogenous (full) capital utilisation, compared to 0.0003 at baseline. These numbers are still significantly lower than without consumption taxation at baseline (2.01 percent).[39]

Table 5: Welfare gains in welfare-equivalent consumption (percent), alternative assumptions

Welfare gains...	non-tax-ded. δ, $v = 1$	tax-ded. δ, variable v	tax-ded. δ, $v = 1$
...from commitment	0.338	0.633	0.338
	...compared to the existing tax system		
Ramsey	6.816	7.793	6.816
Markov	6.447	7.709	6.447

These results highlight that our assumption that the capital income tax is distortionary within the period is key to the result that Ramsey and Markov-perfect tax policies and allocations are almost identical.

5 Concluding remarks

This paper has studied the properties of time-consistent optimal fiscal policies when the policy-maker has access to consumption taxation. 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. A crucial assumption for this result is that the capital income tax is distortionary within the period, due to variable capital utilisation and non-tax-deductibility of

[39] Note that we cannot find similar numbers without intratemporal distortion caused by capital taxation and without consumption taxation, because no Markov-perfect equilibrium exists for standard calibrations such as ours, as Martin (2010) shows, see above.

(actual) depreciation. Our results reinforce the conclusions of Zhu (1995) on the importance of the capital utilisation margin when studying optimal taxation.

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 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 a time-consistent policy maker can only tax labour and capital income, the welfare gains are reduced to 4.92 percent. With consumption taxation the welfare gains of commitment are negligible, while they are substantial with only factor income taxation. The gains from taxing consumption are much larger without than with commitment.

In this paper we have considered a representative-agent framework, hence we studied the optimal tax mix from an efficiency perspective only. An important task for future research is to analyse the distributional impact of different tax instruments with and without commitment in a model with heterogeneous 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.

Debortoli, D. and R. Nunes (2013). Lack of Commitment and the Level of Debt. *Journal of the European Economic Association 11*(5), 1053–1078.

Debortoli, D., R. Nunes, and P. Yared (2017). Optimal Time-Consistent Government Debt Maturity. *The Quarterly Journal of Economics 132*(1), 55–102.

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.

Galindev, R. and D. Lkhagvasuren (2010). Discretization of Highly Persistent Correlated AR(1) Shocks. *Journal of Economic Dynamics and Control 34*(7), 1260–1276.

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.

Kopecky, K. A. and R. M. Suen (2010). Finite State Markov-Chain Approximations to Highly Persistent Processes. *Review of Economic Dynamics 13*(3), 701–714.

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 (1998/2017). Recursive Contracts. Mimeo.

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

Mateos-Planas, X. (2010). Demographics and the Politics of Capital Taxation in a Life-Cycle Economy. *American Economic Review 100*(1), 337–363.

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 (2018). Optimal Fiscal Policy with Consumption Taxation. *Journal of Money, Credit and Banking* (forthcoming).

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.

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 Analytical characterisations

A.1 First-best allocation

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

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

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

$$u_c\left(\mathbf{a}^t\right) = \lambda_2\left(\mathbf{a}^t\right), \tag{17}$$
$$u_\ell\left(\mathbf{a}^t\right) = \lambda_1\left(\mathbf{a}^t\right), \tag{18}$$
$$a_t(1-\gamma)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma} = \lambda_1\left(\mathbf{a}^t\right), \tag{19}$$
$$u_g\left(\mathbf{a}^t\right) = \lambda_2\left(\mathbf{a}^t\right), \tag{20}$$
$$\lambda_2\left(\mathbf{a}^t\right) = \beta\mathbb{E}_t\left(\lambda_1\left(\mathbf{a}^{t+1}\right)\left[a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)^{\gamma}\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\right)^{1-\gamma} + 1 - \delta\left(v\left(\mathbf{a}^{t+1}\right)\right)\right]\right), \tag{21}$$
$$\delta_u\left(\mathbf{a}^t\right) = a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \tag{22}$$
$$\ell\left(\mathbf{a}^t\right) + h\left(\mathbf{a}^t\right) = 1, \tag{23}$$
$$c\left(\mathbf{a}^t\right) + g\left(\mathbf{a}^t\right) + k\left(\mathbf{a}^t\right) = a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma} h\left(\mathbf{a}^t\right)^{1-\gamma} + \left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right). \tag{24}$$

Straightforward combinations of (17)-(24) lead to the following equations which characterise the first-best allocation:

$$u_g\left(\mathbf{a}^t\right) = u_c\left(\mathbf{a}^t\right), \tag{25}$$
$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} = a_t(1-\gamma)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma}, \tag{26}$$

$$h\left(\mathbf{a}^t\right)+\ell\left(\mathbf{a}^t\right)=1, \tag{27}$$

$$u_c\left(\mathbf{a}^t\right)=\beta\mathbb{E}_t\left(u_c\left(\mathbf{a}^{t+1}\right)\left[1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)+a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\right]\right), \tag{28}$$

$$c\left(\mathbf{a}^t\right)+g\left(\mathbf{a}^t\right)+k\left(\mathbf{a}^t\right)=a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}+\left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right), \tag{29}$$

$$\delta_v\left(\mathbf{a}^t\right)=a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \tag{30}$$

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

A.2 Constraints of the policy problems

The following six equations characterise competitive equilibria once three variables (output $y\left(\mathbf{a}^t\right)$ and prices $w\left(\mathbf{a}^t\right)$ and $r\left(\mathbf{a}^t\right)$) and four equations ((7), (8),(9), and (3)) are eliminated in Definition 2: $\forall\mathbf{a}^t$,

$$h\left(\mathbf{a}^t\right)+\ell\left(\mathbf{a}^t\right)=1,\ \forall\mathbf{a}^t, \tag{31}$$

$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)}=\frac{1-\tau^h\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)}a_t\left(1-\gamma\right)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma}, \tag{32}$$

$$\begin{aligned}\frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)}=\beta\mathbb{E}_t\Bigg(&\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)}\Big[1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)\\&+\left(1-\tau^k\left(\mathbf{a}^{t+1}\right)\right)a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\Big]\Bigg),\end{aligned} \tag{33}$$

$$\delta_v\left(\mathbf{a}^t\right)=\left(1-\tau^k\left(\mathbf{a}^t\right)\right)a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \tag{34}$$

$$g\left(\mathbf{a}^t\right)=a_t\left[\tau^k\left(\mathbf{a}^t\right)\gamma+\tau^h\left(\mathbf{a}^t\right)\left(1-\gamma\right)\right]\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}+\tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right), \tag{35}$$

$$c\left(\mathbf{a}^t\right)+g\left(\mathbf{a}^t\right)+k\left(\mathbf{a}^t\right)=a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}+\left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right), \tag{36}$$

which is (12) in the main text.

Then, we use the household's intratemporal optimality condition (32), and the government's budget constraint, (35), to express the labour and capital income tax rates when history $\mathbf{a}^t$ has occurred, respectively, as

$$\tau^h\left(\mathbf{a}^t\right)=1-\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)}\frac{1+\tau^c\left(\mathbf{a}^t\right)}{\left(1-\gamma\right)a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{-\gamma}}, \tag{37}$$

$$\tau^k\left(\mathbf{a}^t\right)=\frac{g\left(\mathbf{a}^t\right)-\tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right)}{\gamma a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}}-\frac{1-\gamma}{\gamma}\tau^h\left(\mathbf{a}^t\right). \tag{38}$$

Replacing for $\left(1-\tau^k\left(\mathbf{a}^{t+1}\right)\right)$ in (33) using (38) and in turn for $\tau^h\left(\mathbf{a}^{t+1}\right)$ using (37), we can write the Euler equation as

$$\frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} = \beta\mathbb{E}_t\left(\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)}\left[1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)+a_{t+1}v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\right.\right.$$
$$\left.\left.-\frac{g\left(\mathbf{a}^{t+1}\right)-\tau^c\left(\mathbf{a}^{t+1}\right)c\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\right]-u_\ell\left(\mathbf{a}^{t+1}\right)\frac{h\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\right), \tag{39}$$

which is (13) in the main text. Similarly, we can eliminate $\tau^k\left(\mathbf{a}^t\right)$ from (34) and rewrite it as

$$\delta_v\left(\mathbf{a}^t\right) = a_t\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma} - \frac{g\left(\mathbf{a}^t\right)-\tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}$$
$$-\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)}\left(1+\tau^c\left(\mathbf{a}^t\right)\right)\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}, \tag{40}$$

which is (14) in the main text.

A.3 First-order conditions of the Ramsey policy-maker's problem

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

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

$$0 = \frac{1}{\left(1+\tau^c\right)^2}\left\{-\lambda_3' u_c + \lambda_3 u_c\left[1-\delta(v)+av\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{k}\right]\right.$$
$$\left.-\lambda_5\left(1-\gamma\right)a\left(\frac{vk}{h}\right)^{\gamma}\right\}-\lambda_4\left(\frac{c}{vk}-\frac{u_\ell}{u_c}\frac{h}{vk}\right)-\lambda_3\frac{u_c}{1+\tau^c}\frac{c}{k}, \tag{41}$$

$$0 = u_c-\lambda_2+\lambda_3'\frac{u_{cc}}{1+\tau^c}-\lambda_3\frac{u_{cc}}{1+\tau^c}\left[1-\delta(v)+av\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{k}\right]$$
$$-\lambda_3\frac{u_c}{1+\tau^c}\frac{\tau^c}{k}-\lambda_4\left[\frac{\tau^c}{vk}+\frac{u_\ell}{u_c^2}u_{cc}\left(1+\tau^c\right)\frac{h}{vk}\right]+\lambda_5\frac{u_\ell}{u_c^2}u_{cc}, \tag{42}$$

$$0 = u_\ell-\lambda_1+\lambda_3 u_{\ell\ell}\frac{h}{k}+\lambda_4\frac{u_{\ell\ell}}{u_c}\left(1+\tau^c\right)\frac{h}{vk}-\lambda_5\frac{u_{\ell\ell}}{u_c}, \tag{43}$$

$$0 = -\lambda_1+\lambda_2 a\left(1-\gamma\right)\left(\frac{vk}{h}\right)^{\gamma}-\lambda_3\left[\frac{u_c}{1+\tau^c}a\left(1-\gamma\right)\frac{v^\gamma}{h^\gamma k^{1-\gamma}}-\frac{u_\ell}{k}\right] \tag{44}$$
$$-\lambda_4\left[a\left(1-\gamma\right)h^{-\gamma}\left(vk\right)^{\gamma-1}-\frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{1}{vk}\right]-\lambda_5\frac{1}{1+\tau^c}a\gamma\left(1-\gamma\right)\left(vk\right)^{\gamma}h^{-\gamma-1},$$

$$0 = u_g-\lambda_2+\lambda_3\frac{u_c}{1+\tau^c}\frac{1}{k}+\lambda_4\frac{1}{vk}, \tag{45}$$

$$
\begin{aligned}
0 = & -\lambda_2 + \beta \mathbb{E} \left(\lambda_2' \left[a' v'^{\gamma} \gamma \left(\frac{h'}{k'} \right)^{1-\gamma} + 1 - \delta(v') \right] \right. \\
& + \lambda_3' \left[\frac{u_c'}{1+\tau^{c\prime}} \left(a' v'^{\gamma} (1-\gamma) k'^{\gamma-2} h'^{1-\gamma} - \frac{g' - \tau^{c\prime} c'}{k'^2} \right) - u_\ell' \frac{h'}{k'^2} \right] \\
& + \lambda_4' \left[a' (1-\gamma) k'^{\gamma-2} \left(\frac{h'}{v'} \right)^{1-\gamma} - \frac{g' - \tau^{c\prime} c'}{v' k'^2} - \frac{u_\ell'}{u_c'} (1+\tau^{c\prime}) \frac{h'}{v' k'^2} \right] \\
& \left. + \lambda_5' \frac{1}{1+\tau^{c\prime}} a' (1-\gamma) \gamma \frac{v'^{\gamma}}{k'^{1-\gamma} h'^{\gamma}} \right),
\end{aligned} \tag{46}
$$

$$
\begin{aligned}
0 = & \lambda_2 \left(a\gamma v^{\gamma-1} k^{\gamma} h^{1-\gamma} - \delta_v k \right) + \lambda_3 \frac{u_c}{1+\tau^c} \left[\delta_v - a\gamma v^{\gamma-1} \left(\frac{h}{k} \right)^{1-\gamma} \right] \\
& + \lambda_4 \left[a (1-\gamma) v^{\gamma-2} \left(\frac{h}{k} \right)^{1-\gamma} - \frac{g - \tau^c c}{v^2 k} - \frac{u_\ell}{u_c} (1+\tau^c) \frac{h}{v^2 k} + \delta_{vv} \right] \\
& + \lambda_5 \frac{1}{1+\tau^c} a (1-\gamma) \gamma v^{\gamma-1} \left(\frac{k}{h} \right)^{\gamma},
\end{aligned} \tag{47}
$$

$$
0 = \ell + h - 1, \tag{48}
$$

$$
0 = c + g + k_{t+1} - a (vk)^{\gamma} h^{1-\gamma} - (1 - \delta(v)) k, \tag{49}
$$

$$
0 = -\frac{u_c}{1+\tau^c} + \beta \mathbb{E} \left(\frac{u_c'}{1+\tau^{c\prime}} \left[1 - \delta(v') + a' v' \left(\frac{h'}{v' k'} \right)^{1-\gamma} - \frac{g' - \tau^{c\prime} c'}{k'} \right] - u_\ell' \frac{h'}{k'} \right), \tag{50}
$$

$$
0 = a \left(\frac{h}{vk} \right)^{1-\gamma} - \frac{g - \tau^c c}{vk} - \frac{u_\ell}{u_c} (1+\tau^c) \frac{h}{vk} - \delta_v, \tag{51}
$$

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

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

A.4 First-order conditions of the time-consistent policy-maker's problem

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

$$
0 = -\lambda_3 \frac{1}{(1+\tau^c)^2} u_c - \lambda_4 \left(\frac{c}{vk} - \frac{u_\ell}{u_c} \frac{h}{vk} \right) - \lambda_5 \frac{1}{(1+\tau^c)^2} (1-\gamma) a \left(\frac{vk}{h} \right)^{\gamma}, \tag{53}
$$

$$
0 = u_c - \lambda_2 + \lambda_3 \frac{u_{cc}}{1+\tau^c} - \lambda_4 \left[\frac{\tau^c}{vk} + \frac{u_\ell}{u_c^2} u_{cc} (1+\tau^c) \frac{h}{vk} \right] + \lambda_5 \frac{u_\ell}{u_c^2} u_{cc}, \tag{54}
$$

$$
0 = u_\ell - \lambda_1 + \lambda_4 \frac{u_{\ell\ell}}{u_c} (1+\tau^c) \frac{h}{vk} - \lambda_5 \frac{u_{\ell\ell}}{u_c}, \tag{55}
$$

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

$$- \lambda_5 \frac{1}{1+\tau^c}(1-\gamma)\gamma a (vk)^{\gamma} h^{-\gamma-1}, \tag{56}$$

$$0 = u_g - \lambda_2 + \lambda_4 \frac{1}{vk}, \tag{57}$$

$$0 = \beta \mathbb{E}\frac{\partial \mathcal{W}(a',k')}{\partial k'} - \lambda_2 - \beta\lambda_3 \mathbb{E}\left(\frac{\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}}{\partial k'}\left[1-\delta\left(\mathcal{V}(a',k')\right)\right.\right. \tag{58}$$

$$\left. + a'\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'} - \frac{\mathcal{G}(a',k')}{k'} + \frac{\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}\right] - \frac{\partial \mathcal{U}_\ell(a',k')\frac{\mathcal{H}(a',k')}{k'}}{\partial k'}$$

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

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

$$+ \lambda_5 \frac{1}{1+\tau^c} a(1-\gamma)\gamma v^{\gamma-1}\left(\frac{k}{h}\right)^{\gamma}, \tag{59}$$

where

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

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

Applying the envelope theorem gives

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

hence,

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

Plugging this condition into (58), we obtain

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

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

$$0=\ell+h-1, \tag{60}$$

$$0=c+g+k'-a(vk)^{\gamma}h^{1-\gamma}-(1-\delta(v))k, \tag{61}$$

$$\begin{aligned}0 =& -\frac{u_c}{1+\tau^c}+\beta\mathbb{E}\left(\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',kx')}\left[1-\delta(\mathcal{V}(a',k'))+a'\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'}\right.\right.\\
&\left.\left.-\frac{\mathcal{G}(a',k')-\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}\right]-\mathcal{U}_\ell(a',k')\frac{\mathcal{H}(a',k')}{k'}\right),\end{aligned} \tag{62}$$

$$0=a\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{vk}-\frac{u_\ell}{u_c}(1+\tau^c)\frac{h}{vk}-\delta_v, \tag{63}$$

$$0\geq\frac{u_\ell}{u_c}-\frac{1}{1+\tau^c}(1-\gamma)a\left(\frac{vk}{h}\right)^{\gamma}, \tag{64}$$

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

B Proofs

Proof of Result 1. Consider first the case where capital is fully utilised, i.e., $v = 1$. Beside the technological constraints, i.e., the time constraint (2) and the resource constraint (12), which constrain the first best as well, the government faces only three constraints: its budget constraint plus the Euler and consumption-leisure optimality condition of households. Given a constant τ^c at the steady state, comparing (28) and (50) gives $\tau^k = 0$. Also, comparing (26) and (32) gives $\frac{1-\tau^h}{1+\tau^c} = 1$, hence $\tau^c = -\tau^h$. Finally, as long as consumption is larger than labour income, the only way to raise revenue to finance g is by setting $\tau^c > 0$. In the case where households choose the capital utilisation rate, the policy-maker has to satisfy an additional incentive constraint, (34), while it has no more instruments. However, given that $\tau^k = 0$, the capital utilisation margin is not distorted, hence the first-best steady state can still be implemented. In addition, it is well known that when a Ramsey equilibrium attains the first best, it is time-consistent. A similar argument can be made with tax-deductible depreciation.

Proof of Result 2. Consider our baseline framework first, i.e., endogenous capital utilisation and non-tax-deductible depreciation. By the usual Kuhn-Tucker argument, if (52) 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 (50) and (52) 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 (52) as equality again, we can rewrite (46) 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 (45) $\lambda_2 - \lambda_3 \frac{u_c}{1+\tau^c}\frac{1}{k} - \lambda_4 \frac{1}{vk} = u_g > 0$, hence $g = \tau^c c$ if $\lambda_5 = 0$. Finally, from (44), this holds if $\lambda_1 = \lambda_2 (1-\gamma)\left(\frac{vk}{h}\right)^{\gamma} = \lambda_2 w$, which says that the marginal value of leisure relative to consumption is the real wage, which obviously holds. Now, with $v = 1$ and $\lambda_4 = 0$ and/or with tax-deductible depreciation it easy to see that the above argument still holds.

C With tax-deductible depreciation

We focus on the deterministic case (and drop aggregate productivity a from the equations to simplify), and use the notation of the recursive version of our model. We assume that the capital tax is levied on $(rv - \delta(v))k$, i.e., (actual) depreciation is tax-deductible.

C.1 Environment and policy problems

The description of the economy is modified as follows. The government's and the household's budget constraints are, respectively,

$$g = \tau^k \left(rv - \delta\left(v\right)\right) k + \tau^h wh + \tau^c c,$$

$$\left(1+\tau^c\right) c + k' = \left(1-\tau^k\right) rvk + \tau^k \delta\left(v\right) k + \left(1-\tau^h\right) wh + \left(1-\delta\left(v\right)\right) k.$$

The private sector's optimality conditions change as follows. The labour income tax can be expressed from the household's consumption-leisure FOC as in (37), as before, and we can express the capital income tax from the government's budget constraint as

$$\tau^k = \frac{g - \left(1-\gamma\right) h \left(\frac{vk}{h}\right)^{\gamma} + \frac{u_\ell}{u_c} h \left(1+\tau^c\right) - \tau^c c}{\left[\gamma v \left(\frac{h}{vk}\right)^{1-\gamma} - \delta\left(v\right)\right] k},$$

where we have replaced for τ^h. The household's Euler now is

$$\frac{u_c}{1+\tau^c} = \beta \mathbb{E}_t \left(\frac{u_c'}{1+\tau^{c\prime}} \left[1 - \delta\left(v'\right) + \left(1-\tau^{k\prime}\right) \gamma v' \left(\frac{h'}{v'k'}\right)^{1-\gamma} + \tau^{k\prime} \delta\left(v'\right) \right] \right).$$

Replacing for $\tau^{k\prime}$ and rearranging give

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

that is, the Euler is unchanged. Finally, the first-order condition with respect to v now is

$$\delta_v = r = \gamma \left(\frac{h}{vk}\right)^{1-\gamma},$$

hence capital income taxation is not distortionary within the period.

The recursive Lagrangian of the Ramsey policy-maker's problem is

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

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

The Markov policy-maker solves

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

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

C.2 First-order conditions of the Ramsey policy-maker's problem

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

$$0 = \frac{1}{(1+\tau^c)^2} \left\{ -\lambda_3' u_c + \lambda_3 u_c \left[1 - \delta(v) + v \left(\frac{h}{vk} \right)^{1-\gamma} - \frac{g - \tau^c c}{k} \right] \right.$$
$$\left. -\lambda_5 (1-\gamma) \left(\frac{vk}{h} \right)^\gamma \right\} - \lambda_3 \frac{u_c}{1+\tau^c} \frac{c}{k}, \tag{65}$$

$$0 = u_c - \lambda_2 + \lambda_3' \frac{u_{cc}}{1+\tau^c} - \lambda_3 \frac{u_{cc}}{1+\tau^c} \left[1 - \delta(v) + v \left(\frac{h}{vk} \right)^{1-\gamma} - \frac{g - \tau^c c}{k} \right]$$
$$- \lambda_3 \frac{u_c}{1+\tau^c} \frac{\tau^c}{k} + \lambda_5 \frac{u_\ell}{u_c^2} u_{cc}, \tag{66}$$

$$0 = u_\ell - \lambda_1 + \lambda_3 u_{\ell\ell} \frac{h}{k} - \lambda_5 \frac{u_{\ell\ell}}{u_c}, \tag{67}$$

$$0 = -\lambda_1 + \lambda_2 (1-\gamma) \left(\frac{vk}{h} \right)^\gamma - \lambda_3 \left[\frac{u_c}{1+\tau^c} (1-\gamma) \frac{v^\gamma}{h^\gamma k^{1-\gamma}} - \frac{u_\ell}{k} \right]$$
$$- \lambda_4 \gamma (1-\gamma) h^{-\gamma} (vk)^{\gamma-1} - \lambda_5 \frac{1}{1+\tau^c} \gamma (1-\gamma) (vk)^\gamma h^{-\gamma-1}, \tag{68}$$

$$0 = u_g - \lambda_2 + \lambda_3 \frac{u_c}{1+\tau^c} \frac{1}{k}, \tag{69}$$

$$0 = -\lambda_2 + \beta \mathbb{E} \left(\lambda_2' \left[v'^\gamma \gamma \left(\frac{h'}{k'} \right)^{1-\gamma} + 1 - \delta(v') \right] \right.$$
$$+ \lambda_3' \left\{ \frac{u_c'}{1+\tau^{c\prime}} \left[v'^\gamma (1-\gamma) k'^{\gamma-2} h'^{1-\gamma} - \frac{g' - \tau^{c\prime} c'}{k'^2} \right] - u_\ell' \frac{h'}{k'^2} \right\}$$
$$\left. - \lambda_4' \gamma (\gamma - 1) k'^{\gamma-2} \left(\frac{h'}{v'} \right)^{1-\gamma} + \lambda_5' \frac{1}{1+\tau^{c\prime}} (1-\gamma) \gamma \frac{v'^\gamma}{k'^{1-\gamma} h'^\gamma} \right), \tag{70}$$

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

$$- \lambda_4 \left[\gamma\left(\gamma-1\right) v^{\gamma-2} \left(\frac{h}{k}\right)^{1-\gamma} - \delta_{vv}\right] + \lambda_5 \frac{1}{1+\tau^c} \left(1-\gamma\right) \gamma v^{\gamma-1} \left(\frac{k}{h}\right)^{\gamma}, \tag{71}$$

$$0 = \ell + h - 1, \tag{72}$$

$$0 = c + g + k_{t+1} - (vk)^{\gamma} h^{1-\gamma} - \left(1 - \delta\left(v\right)\right) k, \tag{73}$$

$$0 = -\frac{u_c}{1+\tau^c} + \beta \mathbb{E} \left(\frac{u_c'}{1+\tau^{c\prime}} \left[1 - \delta\left(v'\right) + v' \left(\frac{h'}{v'k'}\right)^{1-\gamma} - \frac{g' - \tau^{c\prime} c'}{k'}\right] - u_\ell' \frac{h'}{k'}\right), \tag{74}$$

$$0 = \gamma \left(\frac{h}{vk}\right)^{1-\gamma} - \delta_v, \tag{75}$$

$$0 \geq \frac{u_\ell}{u_c} - \frac{1}{1+\tau^c} \left(1-\gamma\right) \left(\frac{vk}{h}\right)^{\gamma}, \tag{76}$$

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

When capital is fully utilised, we set $v = 1$ and $\lambda_4 = 0$, and ignore (71) and (75).

C.3 First-order conditions of the time-consistent policy-maker's problem

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

$$0 = -\lambda_3 \frac{1}{\left(1+\tau^c\right)^2} u_c - \lambda_5 \frac{1}{\left(1+\tau^c\right)^2} \left(1-\gamma\right) \left(\frac{vk}{h}\right)^{\gamma}, \tag{77}$$

$$0 = u_c - \lambda_2 + \lambda_3 \frac{u_{cc}}{1+\tau^c} + \lambda_5 \frac{u_\ell}{u_c^2} u_{cc}, \tag{78}$$

$$0 = u_\ell - \lambda_1 - \lambda_5 \frac{u_{\ell\ell}}{u_c}, \tag{79}$$

$$0 = -\lambda_1 + \lambda_2 \left(1-\gamma\right) \left(\frac{vk}{h}\right)^{\gamma} - \lambda_4 \gamma (1-\gamma) h^{-\gamma} (vk)^{\gamma-1} - \lambda_5 \frac{1}{1+\tau^c} \left(1-\gamma\right) \gamma \left(vk\right)^{\gamma} h^{-\gamma-1}, \tag{80}$$

$$0 = u_g - \lambda_2, \tag{81}$$

$$0 = -\lambda_2 + \beta \mathbb{E} \left(\Lambda_2\left(k'\right) \left[\gamma \mathcal{V}\left(k'\right)^{\gamma} \left(\frac{\mathcal{H}\left(k'\right)}{k'}\right)^{1-\gamma} + 1 - \delta\left(\mathcal{V}\left(k'\right)\right)\right]\right.$$

$$\left. -\Lambda_4 \gamma (\gamma-1) \left(\frac{h'}{v'}\right)^{1-\gamma} k'^{\gamma-2} + \Lambda_5\left(k'\right) \frac{1}{1+\mathcal{T}^c\left(k'\right)} \left(1-\gamma\right) \gamma \frac{\mathcal{V}\left(k'\right)^{\gamma}}{k'^{1-\gamma} \mathcal{H}\left(k'\right)^{\gamma}}\right)$$

$$-\beta \lambda_3 \mathbb{E} \left(\frac{\frac{\mathcal{U}_c(k')}{1+\mathcal{T}^c(k')}}{\partial k'} \left[1 - \delta\left(\mathcal{V}\left(k'\right)\right) \quad + \frac{\mathcal{H}\left(k'\right)^{1-\gamma} \mathcal{V}\left(k'\right)^{\gamma} k'^{\gamma}}{k'} - \frac{\mathcal{G}\left(k'\right)}{k'} + \frac{\mathcal{T}^c\left(k'\right) \mathcal{C}\left(k'\right)}{k'}\right]\right.$$

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

where we have applied the envelop theorem,

$$0 = \lambda_2 \left(\gamma v^{\gamma-1} k^{\gamma} h^{1-\gamma} - \delta_v k\right) - \lambda_4 \left[\gamma(\gamma-1)\left(\frac{h}{k}\right)^{1-\gamma} v^{\gamma-2} - \delta_{vv}\right] + \lambda_5 \frac{1}{1+\tau^c}(1-\gamma)\gamma v^{\gamma-1}\left(\frac{k}{h}\right)^{\gamma}. \tag{82}$$

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

$$0 = \ell + h - 1, \tag{83}$$

$$0 = c + g + k' - (vk)^{\gamma} h^{1-\gamma} - (1 - \delta(v)) k, \tag{84}$$

$$0 = -\frac{u_c}{1+\tau^c} + \beta \mathbb{E}\left(\frac{\mathcal{U}_c(k')}{1+\mathcal{T}^c(k')}\left[1 - \delta(\mathcal{V}(k')) + \frac{\mathcal{H}(k')^{1-\gamma}\mathcal{V}(k')^{\gamma} k'^{\gamma}}{k'} - \frac{\mathcal{G}(k') - \mathcal{T}^c(k')\mathcal{C}(k')}{k'}\right] - \mathcal{U}_{\ell}(k')\frac{\mathcal{H}(k')}{k'}\right), \tag{85}$$

$$0 = \gamma\left(\frac{h}{vk}\right)^{1-\gamma} - \delta_v, \tag{86}$$

$$0 \geq \frac{u_\ell}{u_c} - \frac{1}{1+\tau^c}(1-\gamma)\left(\frac{vk}{h}\right)^{\gamma}, \tag{87}$$

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

When capital is fully utilised, we set $v = 1$ and $\lambda_4 = 0$, and ignore (82) and (86).

D Response to aggregate productivity shocks

We now study the cyclical properties of policy instruments and allocations. We are interested in whether the close similarity between Ramsey and Markov policies when consumption is taxed optimally still holds when the economy faces aggregate productivity shocks. For each policy scenario, we simulate the model and calculate sample statistics from the simulated data.[40] The results of this exercise for our baseline calibration are reported in Table 6.[41]

When consumption taxes are not available and the policy-maker can credibly commit, we recover well-known labour tax smoothing result of the Ramsey literature (Chari et al., 1994).

[40]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.

[41]We have also solved the stochastic model with $\varphi = 1$ to check the robustness of the cyclical properties of tax rates and allocations. The main features remain unchanged. The results are available upon request.

Table 6: Cyclical properties of taxes and allocations

	$\tau^h \geq 0$		$\tau^c = 0$	
	Ramsey	Markov	Ramsey	Markov
	Consumption tax		Labour income tax	
Mean	0.223	0.221	0.240	0.065
Standard deviation	0.005	0.004	0.002	0.002
Coefficient of variation	0.021	0.019	0.009	0.026
Autocorrelation	0.609	0.640	0.922	0.507
Correlation with output	0.9998	0.996	-0.700	-0.853
Capital income tax				
Mean	0.000	0.004	0.000	0.198
Standard deviation	0.008	0.007	0.009	0.003
Coefficient of variation	10.807	1.699	30.693	0.017
Autocorrelation	0.609	0.638	0.463	0.504
Correlation with output	-0.9995	-0.996	-0.801	-0.909
Public spending				
Mean	0.072	0.072	0.068	0.051
Standard deviation	0.001	0.001	0.001	0.001
Coefficient of variation	0.013	0.013	0.019	0.020
Autocorrelation	0.976	0.976	0.776	0.809
Correlation with output	0.400	0.431	0.880	0.886
Public spending-income ratio				
Mean	0.146	0.146	0.146	0.117
Standard deviation	0.004	0.004	0.004	0.002
Coefficient of variation	0.027	0.026	0.026	0.020
Autocorrelation	0.521	0.524	0.514	0.504
Correlation with output	-0.894	-0.893	-0.936	-0.885
Consumption				
Mean	0.322	0.322	0.305	0.317
Standard deviation	0.004	0.004	0.005	0.005
Coefficient of variation	0.013	0.013	0.017	0.016
Autocorrelation	0.976	0.976	0.939	0.935
Correlation with output	0.400	0.431	0.621	0.675
Hours				
Mean	0.275	0.276	0.261	0.283
Standard deviation	0.005	0.005	0.007	0.006
Coefficient of variation	0.018	0.018	0.027	0.022
Autocorrelation	0.524	0.527	0.513	0.503
Correlation with output	0.859	0.856	0.936	0.895
Output				
Mean	0.492	0.492	0.467	0.439
Standard deviation	0.015	0.015	0.020	0.016
Coefficient of variation	0.030	0.030	0.042	0.036
Autocorrelation	0.603	0.607	0.570	0.597
Welfare-eq. consumption loss	0.054	0.058	0.080	0.168

Under discretion, the policy-maker uses both capital and labour income taxes in response to unexpected productivity changes. This is due to the fact that the Markov policy-maker is less able to smooth the effects of random productivity events intertemporally, as it is less able to use the capital income tax as shock absorber, given that it relies heavily on this instrument to raise fiscal revenue. 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.[42]

The differences between Ramsey and Markov policies are greatly reduced when the policy-maker can tax consumption. As in the case without consumption taxation, under commitment the coefficient of variation of capital income taxes is larger than that of the alternative tax instrument, in this case, the consumption tax. However, under Markov policies capital taxes still play the main role in absorbing shocks, as opposed to without consumption taxation. A new feature of tax policies with consumption taxation is that the consumption tax rate is highly procyclical. The capital income tax rate remains countercyclical, but its correlation with output increases when consumption is taxed. All allocation variables, namely, consumption, public spending, hours, and output vary less with consumption taxation than without.

Finally, we compute long-run expected welfare as a percentage increase in welfare-equivalent consumption 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.[43] The values 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.

[42] 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.

[43] 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.

School of Economics and Finance

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/