



Tax smoothing in a business cycle model with capital-skill complementarity



Konstantinos Angelopoulos^a, Stylianos Asimakopoulos^{b,*}, James Malley^{a,c}

^a University of Glasgow, United Kingdom

^b University of Nottingham, School of Economics, Sir Clive Granger Building, Room B15, University Park, Nottingham, NG7 2RD, United Kingdom

^c CESifo, Germany

ARTICLE INFO

Article history:

Received 11 July 2014

Received in revised form

25 September 2014

Accepted 4 November 2014

Available online 12 November 2014

JEL classification:

E62

E32

J31

Keywords:

Skill premium

Tax smoothing

Optimal fiscal policy

ABSTRACT

This paper undertakes a normative investigation of the quantitative properties of optimal tax smoothing in a business cycle model with state contingent debt, capital-skill complementarity and endogenous skill acquisition under technology and public expenditure shocks. We find that skilled and unskilled labour tax smoothing maintain quantitatively under externalities and exogenous shocks in skill acquisition, as well as when the relative skill supply is exogenously determined. We further find that the government finds it optimal to reduce both the size of the wedge between the marginal rates of substitution and transformation in skill attainment in the long-run and the standard deviation of this wedge over the business cycle. This is achieved by subsidising skill creation and taxing both types of labour income.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

1. Introduction

The celebrated tax smoothing result of Barro (1979) in a partial equilibrium setting has led to a number of important studies on optimal fiscal policy over the business cycle in representative agent general equilibrium models. For example, Lucas and Stokey (1983) formalised labour tax smoothing within a complete markets neoclassical setup without capital when the government had access to state-contingent debt. Chari et al. (1994) generalised this result in a model with capital and labour taxation and showed that Ramsey policy dictated that the labour income tax fluctuated very little in response to aggregate shocks and the *ex ante* capital income tax fluctuated around zero.

The literature has also examined the implications of policy frictions and incomplete asset markets for optimal tax and debt policy through a variety of restrictions to the policy instrument set, government debt and capital income taxation (see e.g. Stockman, 2001; Aiyagari et al., 2002; Angeletos, 2002; Buera and Nicolini, 2004; Farhi, 2010). In contrast, assuming complete asset markets and a complete instrument set, Arseneau and Chugh (2012) considered labour market frictions associated with a division of the labour force into employed and unemployed workers. Their model, with state-contingent debt but no capital, suggested that optimal labour tax volatility depended on whether wages were set efficiently.

* Corresponding author. Tel.: +44 7904085471.

E-mail address: stylianos.asimakopoulos@nottingham.ac.uk (S. Asimakopoulos).

Another important division of the labour force is with respect to the type of labour services workers provide and, in particular, how these complement capital in the production process. This is especially pertinent given the empirical relevance of the wage premium accruing to skilled labour and the roles attributed to capital-skill complementarity, the relative supply of skilled labour and capital augmenting technical progress (see e.g. Katz and Murphy, 1992; Krusell et al., 2000; Hornstein et al., 2005). In an important contribution, which also considers non-homogenous labour, Werning (2007) established the conditions under which optimal labour tax smoothing held in a model with redistribution under complete asset markets when workers differed with respect to their productivity. However, since the distribution of productivity differentials is taken as given, this approach does not account for the endogenous determination of employment type (see e.g. Matsuyama, 2006, who also reviews the literature on upward professional mobility).

In this paper we aim to contribute to the tax smoothing literature by focusing on the differences in the complementarity between capital and skilled/unskilled labour as well as the endogenous determination of the relative skill supply for Ramsey tax policy. Compared to Werning (2007), we focus on aggregate outcomes and abstract from redistribution incentives, by following the literature that examines a division of the labour force into two types of workers. To this end, we work with a representative household which guarantees its members' the same level of consumption (see e.g. Arseneau and Chugh, 2012). We thus stay as close as possible to the representative agent Ramsey analysis of Chari et al. (1994) and extend their model to allow for capital-skill complementarity and endogenous skill formation.

Our goal is thus to undertake a normative investigation of the quantitative properties of optimal taxation of capital and labour income, as well as skill-acquisition expenditure, in the presence of aggregate shocks to total factor productivity (TFP), investment-specific technological change and government spending. We assume complete asset markets, but consider the implications of different assumptions regarding the mechanism driving skill acquisition which can lead to labour market imperfections. Motivated by the analysis and discussions in Goldin and Katz (2008), we focus on three main alternative cases. First, we evaluate the effects of externalities in skill creation given the relevance of peer effects related to neighbourhood and social class. Second, we consider the importance of shocks to exogenous non-fiscal policy interventions and socio-political factors affecting skill formation. Finally, we examine the implications of restricting the ratio of skilled to total workers to remain constant.

In our setup, the government can borrow, tax (or subsidise) skill acquisition expenditure, capital, skilled and unskilled labour income separately, to finance exogenous public spending. All policy instruments are allowed to be state-contingent. In this environment, the optimal taxes on labour income and skill acquisition expenditure are uniquely determined. However, as is well known, when the government has access to both state contingent debt and state contingent capital taxation, the second-best Ramsey allocations do not uniquely pin down optimal debt and capital taxes (see Chari et al., 1994). Hence, following the literature, we discuss the properties of the *ex ante* capital tax rate. Moreover, we also examine the case where debt is restricted to be state uncorrelated, which allows us to calculate the *ex post* capital tax or, if we also allow for state-contingent taxation of income from bonds, the private assets tax.¹

We first find that labour income taxes remain smooth under capital-skill complementarity and the different assumptions on skill acquisition considered. The tax on skilled labour income is more volatile when relative skill supply is endogenous, whereas the unskilled labour income tax becomes more volatile when relative skill supply is fixed. However, given the small magnitudes of the standard deviations, these changes are not quantitatively significant for the volatility of labour taxes.

Although optimal labour taxes do not optimally change much over the business cycle, we further find that capital-skill complementarity leads to different correlations of the taxes on skilled and unskilled labour income with the exogenous productivity shocks. In particular, to smooth labour hours for both types of skill, the government has to make the policy wedges in the skilled and unskilled labour markets move in the same direction as the shock for skilled labour. In contrast, the wedge needs to move in the opposite direction after the very short-run for unskilled labour. The difference arises because technology shocks increase the productivity of skilled labour more than that of unskilled, which creates an incentive to substitute skilled labour hours for unskilled. In turn, this implies that the tax on skilled labour income is positively correlated with exogenous technological shocks, while the tax on unskilled labour income is effectively uncorrelated.

We next find that the skill-acquisition subsidy is the least smooth of the policy instruments that apply to static margins of choice and that its correlation properties follow those of the tax on skilled labour income. This subsidy is used to affect the wedge between the marginal rates of substitution and transformation in relative skill supply. The government finds it optimal to reduce this wedge in the steady-state and also to reduce its standard deviation over the business cycle.

Our results further show that the skill heterogeneity considered, irrespective of the assumptions regarding relative skilled supply that we examine, does not affect the results obtained in the literature regarding the cyclical behaviour of asset taxes. In particular, the *ex ante* tax rate on capital fluctuates around zero and the state contingent private assets and *ex post* capital taxes are near zero and are the most volatile of the tax instruments. Finally, we find that irrespective of the model variant examined, all of the policy instruments except for the *ex post* capital tax and the private assets tax inherit the persistence properties of the shocks.

¹ As shown by Zhu (1992) and Chari et al. (1994), state-contingent capital income taxes allow the government to implement the complete asset markets outcome, despite the lack of access to state-contingent debt.

The remainder of the paper is organised as follows. Sections 2 and 3 present the theoretical model and the Ramsey problem respectively. Section 4 describes the quantitative implementation and Section 5 analyses the results. Section 6 draws the conclusions.

2. Model

We develop a model that extends the complete markets neoclassical setup in Zhu (1992) and Chari et al. (1994) by allowing for a division of the labour force into skilled and unskilled workers, an endogenous skill supply and externalities in skill-acquisition on the household side, and capital-skill complementarity on the goods production side. This setup implies a wage premium for skilled labour, the relative supply of which can be increased by a cost to the household in the form of earmarked training expenditure.² As in Chari et al. (1994) households save in the form of physical capital and state-contingent government bonds.

The household is modelled as an infinitely-lived representative dynasty. The head of the household makes all choices on behalf of its members by maximising the aggregate welfare of the family, ensuring that each household member experiences the same level of consumption irrespective of individual labour market status. This is a commonly employed assumption since Merz (1995), given that it allows for tractability when studying aggregate fluctuations under heterogeneities in the labour market (see e.g. Arseneau and Chugh, 2012 for an example with optimal tax policy).

Firms use capital, skilled and unskilled labour to produce a homogeneous product. Following Katz and Murphy (1992), Krusell et al. (2000) and Hornstein et al. (2005), skilled labour is assumed to be more complementary to capital than unskilled labour. Hence, capital accumulation as well as technological developments and government policies that are capital augmenting, tend to increase the skilled wage premium. In contrast, increases in the relative supply of skilled labour act to reduce the skill premium. Finally, the government can borrow and tax capital, skilled and unskilled labour income separately, to finance subsidies on skill-acquisition expenditure and exogenous public spending.

2.1. Notation

The notation employed throughout follows Ljungqvist and Sargent (2012). In particular, we assume that in every period $t \geq 0$, there is a realization of shocks (stochastic events) $s_t \in S$. Therefore, at each period t there is a history of events $s^t = [s_0, s_1, s_2, \dots, s_t]$ which is known. The unconditional probability of observing a specific history of events s^t is defined as $\pi_t(s^t)$. For $t > \tau$, the conditional probability of having s^t sequence of events given the realization of s^τ is defined as: $\pi_t(s^t|s^\tau)$.

2.2. Households

A representative household is composed of two types of members who provide skilled and unskilled labour services.³ The utility function of the household is given by $U\{c_t(s^t), l_t^{*,s}(s^t), l_t^{*,u}(s^t)\}$ where $U(\cdot)$ is increasing, strictly concave and three times continuously differentiable with respect to its arguments; $c_t(s^t)$ is average consumption of all household members at time t given the history of events s^t ; $l_t^{*,s}(s^t) = \psi_t(s^t)l_t^s(s^t)$ and $l_t^{*,u}(s^t) = [1 - \psi_t(s^t)]l_t^u(s^t)$, denote, respectively, average household leisure time from skilled and unskilled members; $l_t^s(s^t)$ and $l_t^u(s^t)$ are leisure time per skilled and unskilled member respectively; and $\psi_t(s^t)$ is the share of skilled to total household members or the relative skill supply. The time constraints facing each type of member are given by

$$h_t^s(s^t) + l_t^s(s^t) = 1 \quad (1)$$

$$h_t^u(s^t) + l_t^u(s^t) = 1 \quad (2)$$

where, $h_t^s(s^t)$ and $h_t^u(s^t)$ denote, respectively, skilled and unskilled labour hours per member. The household can determine its relative skill supply by incurring an average (over all its members) skill-acquisition expenditure, $e_t(s^t)$, according to the following relation:

$$\psi_t(s^t) = q[e_t(s^t), e_t^e(s^t); \xi] \quad (3)$$

where $q(\cdot)$ is increasing, strictly concave and three times continuously differentiable with respect to $e_t(s^t)$ and $e_t^e(s^t)$; $e_t^e(s^t)$ denotes aggregate, economy-wide skill-acquisition expenditure that the representative household takes as given; and $0 < \xi \leq 1$ is a parameter that captures the extent of externalities in skill creation, with higher values of ξ denoting less externalities. Externalities associated with skill acquisition expenditure capture for example peer effects related to neighbourhood and social class. The case where all expenditure on skill-acquisition is internalised, i.e. when $e_t(s^t) = e_t^e(s^t)$ can be obtained by setting $\xi = 1$.

² This is consistent with the literature on upward professional mobility, where there is a cost associated with achieving a higher professional status (see e.g. Matsuyama, 2006 for a review of several models).

³ Note that the unit mass of household members is equal to the sum of its skilled and unskilled members.

⁴ Since consumption is the same for all members of the household, average and per member consumption are equal.

We can re-arrange (3) for $e_t(s^t)$ to obtain skill-acquisition expenditure as a function of relative skill supply, i.e. $e_t(s^t) \equiv g(\cdot) = g[\psi_t(s^t), e_t^e(s^t); \xi]$, where $\partial g(\cdot)/\partial \psi_t(s^t) \equiv g_\psi(s^t) > 0$, and $\partial g(\cdot)/\partial e_t^e(s^t) \equiv g_{e^e}(s^t) < 0$.⁵ We will denote the solution for $e_t(s^t)$ in (3) in the case where $e_t^e(s^t) = e_t(s^t)$ by $\bar{g}(\cdot) = \bar{g}[\psi_t(s^t)]$. Finally note the function $g(\cdot)$ has the property that $\bar{g}(\cdot) = g[\psi_t(s^t), e_t^e(s^t)]$ when $\xi = 1$.

The household can invest in capital and in state-contingent sequentially traded government bonds that mature fully within a period. The evolution of capital is given by

$$k_{t+1}(s^t) = (1 - \delta)k_t(s^{t-1}) + A_t^{*,k}(s_t)i_t(s^t) \tag{4}$$

where $k_t(s^{t-1})$ is the per member stock of capital at time t given the history of events s^{t-1} ; $i_t(s^t)$ is investment in capital at time t given the history of events s^t ; and $0 < \delta < 1$ is the capital depreciation rate. The capital evolution equation allows for an exogenous process, $A_t^{*,k}(s_t)$, capturing an investment-specific technological change, which has been shown to contribute to output fluctuations (see e.g. Greenwood et al., 2000), as well as the changes in the skill premium (see e.g. Krusell et al., 2000; Lindquist, 2004; Pourpourides, 2011). By using Eq. (4) and defining as $A_t^k(s_t) \equiv 1/A_t^{*,k}(s_t)$, we get the household's sequence of budget constraints:

$$c_t(s^t) + A_t^k(s_t)k_{t+1}(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}|s^t)b_{t+1}(s_{t+1}|s^t) + [1 - \tau_t^s(s^t)]w_t^s(s^t)\psi_t(s^t)h_t^s(s^t) + [1 - \tau_t^u(s^t)]w_t^u(s^t)[1 - \psi_t(s^t)]h_t^u(s^t) + (1 - \delta)A_t^k(s_t) \times k_t(s^{t-1}) + [1 - \tau_t^k(s^t)]r_t(s^t)k_t(s^{t-1}) + b_t(s_t|s^{t-1}) \quad \forall t \tag{5}$$

where $p_t(s_{t+1}|s^t)$ is the pricing kernel for government bonds in terms of t goods and $b_{t+1}(s_{t+1}|s^t)$ is the state s_{t+1} contingent payout value of bonds bought per member at period t ;⁶ $\tau_t^s(s^t)$, $\tau_t^k(s^t)$, $\tau_t^u(s^t)$ are the tax rates on skilled and unskilled labour, capital income respectively; and $\tau_t^a(s^t)$ is a subsidy on skill-acquisition expenditure; $w_t^s(s^t)$ and $w_t^u(s^t)$ are the wage rates of skilled and unskilled labour respectively; and $r_t(s^t)$ is the return to capital.

2.3. First order conditions for households

The objective of the representative household is to maximise:

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) u\{c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t)\} \tag{6}$$

where $0 < \beta < 1$ is the household's subjective discount factor, subject to the sequence of constraints in (5), by choosing $\{c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_{t+1}(s^t)\}_{t=0}^{\infty}$ and $\{b_{t+1}(s_{t+1}, s^t); \forall s_{t+1}^t\}_{t=0}^{\infty}$, given initial values for b_0, k_0 . In each time period t and given history s^t , $\{b_{t+1}(s_{t+1}, s^t)\}_{t=0}^{\infty}$ is a vector of government bonds with one element of the vector for each possible realisation of s_{t+1} . In (6), $u(\cdot)$ is obtained by substituting the time accounting identities for $l_t^{*,s}(s^t)$ and $l_t^{*,u}(s^t)$ and the time constraints (1)– (2) into $U(\cdot)$. This maximisation problem yields six first-order conditions which are reported in Appendix A.

Combining the first-order conditions for consumption, skilled and unskilled labour supply gives the following *atemporal* equilibrium conditions, which equate the marginal rates of substitution between consumption and each type of labour with the average returns to skilled and unskilled labour net of taxes:

$$-\frac{u_{h^s}(s^t)}{u_c(s^t)} = \psi_t(s^t)w_t^s(s^t)[1 - \tau_t^s(s^t)] \tag{7}$$

$$-\frac{u_{h^u}(s^t)}{u_c(s^t)} = [1 - \psi_t(s^t)]w_t^u(s^t)[1 - \tau_t^u(s^t)] \tag{8}$$

Substituting the first-order condition for consumption into the first-order condition for $\psi_t(s^t)$ gives the following *atemporal* condition for relative skill supply:

$$-\frac{u_\psi(s^t)}{u_c(s^t)} = h_t^s(s^t)[1 - \tau_t^s(s^t)]w_t^s(s^t) - h_t^u(s^t)[1 - \tau_t^u(s^t)]w_t^u(s^t) - [1 - \tau_t^a(s^t)]g_\psi(s^t) \tag{9}$$

Condition (9) states that the marginal rate of substitution between consumption and the relative skill supply is equal to the net marginal benefit of increasing the household's share of skilled workers. The latter includes the net increase in labour income, given by the difference between the post-tax labour income from an additional skilled member, $h_t^s(s^t)[1 - \tau_t^s(s^t)]w_t^s(s^t)$, and the post-tax labour income from one less unskilled member, $h_t^u(s^t)[1 - \tau_t^u(s^t)] \times w_t^u(s^t)$. From this, the household needs to deduct the post-subsidy cost for an additional skilled member, $[1 - \tau_t^a(s^t)]g_\psi(s^t)$.

Substituting the first-order condition for consumption and its one-period lead into the first-order conditions for the two assets gives the following *intertemporal* conditions equating the current cost of investing in bonds and capital to the future state-contingent and expected benefits respectively:

$$u_c(s^t)p_t(s_{t+1}|s^t) = \beta\pi_{t+1}(s^{t+1}|s^t)u_c(s^{t+1}) \tag{10}$$

⁵ Note that we follow Ljungqvist and Sargent (2012) in using the notation $\partial X(\cdot)/\partial x_t(s^t) \equiv X_x(s^t)$ for function X and variable x in time t for history s^t .

⁶ Given the period t state $s_t|s^{t-1}$ (or else the history s^t), the income side of the household budget includes revenue from bonds dated $b_t(s_t|s^{t-1})$.

$$E_t \left\{ \frac{\beta u_c(s^{t+1})}{u_c(s^t) A_t^k(s_t)} \left[A_{t+1}^k(s_{t+1})(1-\delta) + (1-\tau_{t+1}^k(s^{t+1}))r_{t+1}(s^{t+1}) \right] \right\} = 1 \tag{11}$$

where $\pi_{t+1}(s^{t+1})/\pi_t(s^t) = \pi_{t+1}(s^{t+1}|s^t)$; E_t is the expectation conditional on information at time t (i.e. history s^t), $E_t x_{t+1}(s^{t+1}) = \sum_{s^{t+1}|s^t} \pi_{t+1}(s^{t+1})/\pi_t(s^t) x_{t+1}(s^{t+1})$; and the summation over $s^{t+1}|s^t$ denotes the sum over all possible histories s^{t+1} such that $\tilde{s}^t = s^t$.

By combining the *intertemporal* conditions (10) and (11) we obtain

$$A_t^k(s_t) = \sum_{s_{t+1}} p_t(s_{t+1}|s^t) \{ [1-\tau_{t+1}^k(s^{t+1})]r_{t+1}(s^{t+1}) + (1-\delta)A_{t+1}^k(s_{t+1}) \} \tag{12}$$

which ensures no-arbitrage between the investment opportunities in bonds and capital.

2.4. Firms

Firms rent capital as well as skilled and unskilled labour from households to maximize their profits using a production technology, $F(\cdot)$, that exhibits constant returns to scale in its three inputs:

$$\Pi_t = F \left[(h_t^{sf}(s^t), h_t^{uf}(s^t), k_t^f(s^t), s_t) - w_t^s(s^t)h_t^{sf}(s^t) - w_t^u(s^t)h_t^{uf}(s^t) - r_t(s^t)k_t^f(s^t) \right] \tag{13}$$

where $F(\cdot)$ incorporates a stochastic shock, s_t , and an f superscript denotes firm quantities. This yields the standard first-order conditions:

$$w_t^s(s^t) = F_{h^{sf}}(s^t) \tag{14}$$

$$w_t^u(s^t) = F_{h^{uf}}(s^t) \tag{15}$$

$$r_t(s^t) = F_{k^f}(s^t). \tag{16}$$

2.5. Government budget and equilibrium conditions

Given a history s^t , the government finances an exogenous stream of expenses, $g_t^e(s_t)$, subsidies to skill-acquisition expenditure, τ_t^a , and its debt obligation, $b_t(s_t|s^{t-1})$, by taxing capital and labour income and by issuing state-contingent debt. Hence, the within-period government budget constraint is given by

$$g_t^e(s_t) = \tau_t^s(s^t)w_t^s(s^t)\psi_t(s^t)h_t^s(s^t) + \tau_t^u(s^t)w_t^u(s^t)[1-\psi_t(s^t)] \times h_t^u(s^t) + \tau_t^k(s^t)r_t(s^t)k_t(s^{t-1}) - \tau_t^a(s^t)\bar{g}(\cdot) + \sum_{s_{t+1}} p_t(s_{t+1}|s^t)b_{t+1}(s_{t+1}|s^t) - b_t(s_t|s^{t-1}). \tag{17}$$

The aggregate consistency condition and market clearing conditions for skilled labour, unskilled labour and capital are given respectively by

$$F(\cdot) = c_t(s^t) + g_t^e(s_t) + \bar{g}(\cdot) + A_t^k(s_t)[k_{t+1}(s^t) - (1-\delta)k_t(s^{t-1})] \tag{18}$$

$$\psi_t(s^t)h_t^s(s^t) = h_t^{sf}(s^t) \tag{19}$$

$$[1-\psi_t(s^t)]h_t^u(s^t) = h_t^{uf}(s^t) \tag{20}$$

$$k_t(s^{t-1}) = k_t^f(s^t). \tag{21}$$

Note that in equilibrium at the aggregate level $e_t^e(s^t) \equiv e_t(s^t)$. The equilibrium condition for $e_t^e(s^t)$ can be obtained by setting $e_t^e(s^t) = e_t(s^t)$ in (3) and re-arranging for $e_t(s^t)$, which is equivalent to

$$e_t^e(s^t) \equiv \bar{g}[\psi_t(s^t)]. \tag{22}$$

Hence, to obtain the equilibrium conditions for the decentralised economy given economic policy, we substitute $e_t^e(s^t)$ using (22) into the first-order conditions of the private economy. This implies that in (5) we replace $g(\cdot)$ by $\bar{g}(\cdot)$ and (9) becomes

$$-\frac{u_{\psi}(s^t)}{u_c(s^t)} = h_t^s(s^t)[1-\tau_t^s(s^t)]w_t^s(s^t) - h_t^u(s^t)[1-\tau_t^u(s^t)]w_t^u(s^t) - [1-\tau_t^a(s^t)]\tilde{g}_{\psi}(s^t) \tag{23}$$

where $\tilde{g}_{\psi}(s^t)$ is obtained by substituting (22) into $g_{\psi}(s^t)$.⁷

⁷ Note that the aggregate conditions (17) and (18) have already been written in terms of $\bar{g}(\cdot)$.

Finally, to obtain the equilibrium conditions in terms of household quantities we substitute the market clearing conditions (19)–(21) into (14)–(16) and (18). In particular, we substitute $h_t^{sf}(s^t)$, $h_t^{uf}(s^t)$, $k_t^f(s^t)$ into $F(\cdot)$, $F_{h^{sf}}(s^t)$, $F_{h^{uf}}(s^t)$ and $F_{k^f}(s^t)$ and denote the resulting quantities respectively by $Y(\cdot)$, $\tilde{F}_{h^s}(s^t)$, $\tilde{F}_{h^u}(s^t)$, $\tilde{F}_k(s^t)$. Therefore, (14)–(16) and (18) become

$$w_t^s(s^t) = \tilde{F}_{h^s}(s^t) \tag{24}$$

$$w_t^u(s^t) = \tilde{F}_{h^u}(s^t) \tag{25}$$

$$r_t(s^t) = \tilde{F}_k(s^t) \tag{26}$$

$$Y(\cdot) = c_t(s^t) + g_t^e(s_t) + \bar{g}(\cdot) + A_t^k(s_t)(k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1})) \tag{27}$$

2.6. Model wedges

In Appendix B we present the social planner's solution for this model economy and express economic efficiency in terms of marginal rates of substitution and marginal rates of transformation that explicitly treat skill acquisition as a primitive. This further allows us, in Appendix C, to obtain the relevant wedges that prevent the private economy from achieving the efficient outcomes, by comparing the first-order conditions that characterise equilibrium outcomes in the private economy to the respective efficient conditions. This analysis demonstrates that, for the equilibrium margins characterising skilled and unskilled labour supply and capital accumulation, the wedges are identical with the relevant labour and capital income taxes. However, this is not the case for the choice of the relative skill supply. In this instance, the wedge, ψ^w , is composed of market, ψ^m , and policy, ψ^p , components

$$\psi^w = \psi^m + \psi^p \tag{28}$$

where

$$\psi^m = \tilde{g}_\psi(s^t) - \bar{g}_\psi(s^t) \tag{29}$$

$$\psi^p = \tau_t^s(s^t)F_{h^{sf}}(s^t)h_t^s(s^t) - \tau_t^u(s^t)F_{h^{uf}}(s^t)h_t^u(s^t) - \tau_t^a(s^t)\tilde{g}_\psi(s^t). \tag{30}$$

Hence, a sufficient condition for eliminating ψ^w is $\tau_t^s(s^t) = \tau_t^u(s^t) = \tau_t^a(s^t) = 0$ and $\xi = 1$. In particular, note that given the Pigouvian nature of taxation in the model with externalities, if $\xi \neq 1$, setting the taxes to zero will not eliminate the wedge between the market and efficient outcomes, since setting taxes to zero will eliminate ψ^p but not ψ^m . However, if $\xi \neq 1$, the government can manipulate the taxes to affect the total wedge. Assuming that the government had access to a lump-sum instrument to finance public spending, all wedges could be eliminated from the market economy by setting $\tau_t^s(s^t) = \tau_t^u(s^t) = \tau_t^k(s^t) = 0$ and

$$\bar{\tau}_t^a(s^t) = \frac{-\bar{g}_\psi(s^t) + \tilde{g}_\psi(s^t)}{\tilde{g}_\psi(s^t)}. \tag{31}$$

The analysis in Appendix C also demonstrates that the tax system with which the government is endowed is complete, since for each margin of adjustment in the market economy there is a unique policy wedge. In other words, the available policy instruments determine each wedge uniquely.

3. The Ramsey problem

The government chooses labour and capital income taxes, the subsidy on skill-acquisition expenditure and next period state-contingent debt to maximise the household's welfare subject to the equilibrium first-order conditions that summarise the reactions of the private economy.⁸ Note that the Ramsey planner internalises the externalities associated with skill-acquisition expenditure.

To solve the Ramsey problem we first derive the present discounted value (PDV) of the household's lifetime budget constraint using the Arrow–Debreu price of the bond and the transversality conditions for bonds and capital. Second, we derive the implementability constraint by substituting out prices and tax rates from the household's present value budget constraint using the first-order conditions for the household and firm. Finally, we derive the optimal Ramsey allocations by maximising the planner's objective function subject to the implementability constraint and the sequence of aggregate resource constraints.

⁸ In Angelopoulos et al. (2013), we develop a model where frictions in both labour and capital markets lead to income inequality between different types of households. We then study the problem of a government which chooses taxes on total income for different households and non-contingent debt. However, following most of the literature on optimal taxation, that setup does not allow for mobility between the different agents.

3.1. Present value of budget constraint

We start with the household's flow budget constraint in equilibrium. Beginning from period 0 and by repeatedly substituting forward one-period budget constraints for the household, we obtain the PDV of the household's lifetime budget constraint:

$$\sum_{t=0}^{\infty} \sum_{s^t} \left(\prod_{i=0}^{t-1} p_i(s_{i+1}|s^i) \right) c_t(s^t) = \sum_{t=0}^{\infty} \sum_{s^t} \left(\prod_{i=0}^{t-1} p_i(s_{i+1}|s^i) \right) \times \{ [(1 - \tau_t^s(s^t)) w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + [(1 - \tau_t^u(s^t)) w_t^u(s^t) \\ \times [1 - \psi_t(s^t)] h_t^u(s^t) - [1 - \tau_t^a(s^t)] \bar{g}(\cdot)] + b_0 + \{ [(1 - \tau_0^k(s_0)) r_0(s_0) + (1 - \delta) A_0^k(s_0)] k_0 \} \} \quad (32)$$

where we have imposed the series of no-arbitrage conditions from (12) $\forall t$ and the following transversality conditions for any s^∞ :

$$\lim_{t \rightarrow \infty} \left(\prod_{i=0}^{t-1} p_i(s_{i+1}|s^i) \right) k_{t+1}(s^t) = 0 \quad (33)$$

$$\lim_{t \rightarrow \infty} \sum_{s_{t+1}} \left(\prod_{i=0}^{t-1} p_i(s_{i+1}|s^i) \right) p_t(s_{t+1}|s^t) b_{t+1}(s_{t+1}|s^t) = 0 \quad (34)$$

which specify that for any possible future history the household does not hold positive or negative valued wealth at infinity. Defining $(\prod_{i=0}^{t-1} p_i(s_{i+1}|s^i)) \equiv q_t^0(s^t)$, $\forall t \geq 1$, with $q_0^0(s^0) \equiv 1$, where $q_t^0(s^t)$ is the Arrow–Debreu price, we can re-write (32) as

$$\sum_{t=0}^{\infty} \sum_{s^t} q_t^0(s^t) c_t(s^t) = \sum_{t=0}^{\infty} \sum_{s^t} q_t^0(s^t) \{ [(1 - \tau_t^s(s^t)) w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + [(1 - \tau_t^u(s^t)) w_t^u(s^t) [1 - \psi_t(s^t)] h_t^u(s^t) - [1 - \tau_t^a(s^t)] \bar{g}(\cdot)] \\ + b_0 + \{ [(1 - \tau_0^k(s_0)) r_0(s_0) + (1 - \delta) A_0^k(s_0)] k_0 \} \} \quad (35)$$

Note that the Arrow–Debreu price satisfies the recursion:

$$q_{t+1}^0(s^{t+1}) = p_t(s_{t+1}|s^t) q_t^0(s^t). \quad (36)$$

Using the first-order condition from the sequential equilibrium for pricing contingent claims (10) and noting that $\pi_0(s^0) = 1$, since, at period 0 the state s^0 is known, the above recursion can be written as

$$q_{t+1}^0(s^{t+1}) = \beta^{t+1} \pi_{t+1}(s^{t+1}) \frac{u_c(s^{t+1})}{u_c(s^0)}. \quad (37)$$

3.2. Implementability constraint

First, note that (37) implies

$$q_t^0(s^t) = \beta^t \pi_t(s^t) \frac{u_c(s^t)}{u_c(s^0)}. \quad (38)$$

Substituting (38) for $q_t^0(s^t)$; the first-order conditions of the firm, (24), (25) and (26) for $w_t^s(s^t)$, $w_t^u(s^t)$ and r_0 , respectively; and the first-order conditions of the household, (7), (8), (23) for $\tau_t^s(s^t)$, $\tau_t^u(s^t)$ and $\tau_t^a(s^t)$ respectively into the present value budget constraint (35), we obtain the implementability constraint⁹:

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) [u_c(s^t) c_t(s^t) + u_{h^s}(s^t) h_t^s(s^t) + u_{h^u}(s^t) h_t^u(s^t) + u_c(s^t) \bar{g}(\cdot) \Omega(s^t)] - A = 0 \quad (39)$$

where $A \equiv A(c_0(s^0), h_0^s(s^0), h_0^u(s^0), \psi_0(s^0); b_0, k_0, \tau_0^k) = u_c(s^0) \{ b_0 + [(1 - \tau_0^k) \tilde{F}_k(s^0) + (1 - \delta) A_0^k(s_0)] k_0 \}$ and

$$\Omega(s^t) = \left\{ - \frac{u_{h^s}(s^t)}{u_c(s^t) \psi_t(s^t)} h_t^s(s^t) + \frac{u_{h^u}(s^t)}{u_c(s^t) (1 - \psi_t(s^t))} h_t^u(s^t) + \frac{u_{\psi}(s^t)}{u_c(s^t)} \right\} / \tilde{g}_{\psi}(s^t).$$

3.3. Pseudo value function

The government maximises (6) subject to the implementability constraint (39) and the sequence of aggregate resource constraints in (27) $\forall t$ by choosing $\{c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_{t+1}(s^t) \forall s^t\}_{t=0}^{\infty}$, given $\{b_0, k_0, \tau_0^k\}$.¹⁰ To achieve this, we

⁹ Note that the *intertemporal* first-order condition (12) has been used already in deriving (35), while the government budget constraint is redundant, since it is a linear combination of the household's budget constraint and the aggregate resource constraint. Therefore, (27) and (39) summarise all the constraints that the government needs to respect.

¹⁰ Note that following the literature we do not examine the problem of initial capital taxation and thus do not allow the government to choose τ_0^k .

follow [Ljungqvist and Sargent \(2012\)](#) and first specify the following within-period pseudo value function:

$$V[c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t); \Phi] = u[c_t(s^t), 1 - h_t^s(s^t), 1 - h_t^u(s^t), \psi_t(s^t)] + \Phi[u_c(s^t)c_t(s^t) + u_{h^s}(s^t)h_t^s(s^t) + u_{h^u}(s^t)h_t^u(s^t) + u_c(s^t)\bar{g}(\cdot)\Omega(s^t)] \quad (40)$$

where Φ is the Lagrange multiplier with respect to the implementability constraint.¹¹ The Lagrangian of the Ramsey planner is defined as

$$J = \sum_{t=0}^{\infty} \beta^t \pi_t(s^t) \{V(c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t); \Phi) + \theta_t(s^t)[Y(\cdot) - c_t(s^t) - g_t^e(s_t) - \bar{g}(\cdot) - (k_{t+1}(s^t) + (1 - \delta)k_t(s^{t-1}))A_t^k(s_t)]\} - \Phi A \quad (41)$$

where $\{\theta_t(s^t); \forall s^t\}_{t=0}^{\infty}$ is a sequence of Lagrange multipliers attached to the aggregate resource constraint.

For a given level of $\{b_0, k_0, \tau_k^k\}$, J is maximized with respect to $\{c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_{t+1}(s^t); \forall s^t\}_{t=1}^{\infty}$ and $c_0(s^0), h_0^s(s^0), h_0^u(s^0), \psi_0(s^0), k_1(s^0)$ yielding the following first-order conditions respectively:

$$V_c(s^t) = \theta_t(s^t), \quad t \geq 1 \quad (42)$$

$$V_{h^s}(s^t) = -\theta_t(s^t)Y_{h^s}(s^t), \quad t \geq 1 \quad (43)$$

$$V_{h^u}(s^t) = -\theta_t(s^t)Y_{h^u}(s^t), \quad t \geq 1 \quad (44)$$

$$V_{\psi}(s^t) = \theta_t(s^t) [\bar{g}_{\psi}(s^t) - Y_{\psi}(s^t)], \quad t \geq 1 \quad (45)$$

$$\theta_t(s^t)A_t^k(s_t) = \beta E_t \theta_{t+1}(s^{t+1}) [Y_k(s^{t+1}) + (1 - \delta)A_{t+1}^k(s_{t+1})], t \geq 0 \quad (46)$$

$$V_c(s^0) = \theta_0(s^0) + \Phi A_c \quad (47)$$

$$V_{h^s}(s^0) = -\theta_0(s^0)Y_{h^s}(s^0) + \Phi A_{h^s} \quad (48)$$

$$V_{h^u}(s^0) = -\theta_0(s^0)Y_{h^u}(s^0) + \Phi A_{h^u} \quad (49)$$

$$V_{\psi}(s^t) = \theta_0(s^0) [\bar{g}_{\psi}(s^0) - Y_{\psi}(s^0)] + \Phi A_{\psi}. \quad (50)$$

The first-order conditions derived in (42)–(50) imply that the system of equations to be solved will be different for $t=0$ and for $t > 0$. These conditions in a non-stochastic environment are presented in [Appendix D](#).

3.4. Capital and asset taxes

As is well known (see e.g. [Zhu, 1992](#); [Chari et al., 1994](#); [Ljungqvist and Sargent, 2012](#)), the Ramsey problem with state-contingent debt cannot uniquely pin down the capital tax rate. Hence, we follow the literature and calculate the optimal *ex ante* capital income tax rate (see [Appendix E](#) for details):

$$\tau_{t+1}^{k,a}(s^t) = \frac{\beta E_t u_c(s^{t+1}) [\bar{F}_k(s^{t+1}) + (1 - \delta)A_{t+1}^k(s_{t+1})] - u_c(s^t)A_t^k(s_t)}{\beta E_t u_c(s^{t+1}) \bar{F}_k(s^{t+1})}. \quad (51)$$

Alternatively, by assuming that government debt is not state-contingent, we can calculate the *ex post* state contingent capital tax (see [Appendix F](#) for the derivation):

$$\tau_t^{k,p}(s^t) = \left(\frac{1}{R_t(s^t)k_t(s^{t-1})} \right) \{g_t^e(s_t) + \tau_t^a(s^t)\bar{g}(\cdot)(t) - \frac{b_{t+1}(s^t)}{R_t(s^t)} + b_t(s^{t-1}) - \tau_t^s(s^t)w_t^s(s^t)\psi_t(s^t)h_t^s(s^t) - \tau_t^u(s^t)w_t^u(s^t)[1 - \psi_t(s^t)]h_t^u(s^t)\} \quad (52)$$

where $R_t(s_t)$ is the state uncontingent or the risk free return to holding government debt. Alternatively, assuming that the government employs a state-contingent tax on income from government bonds, we can calculate the private assets tax, $\xi_t(s^{t+1}|s^t)$ that applies to taxing jointly the income from assets as (see [Appendix F](#) for the derivation):

$$\xi_t(s^{t+1}|s^t) = \left(\frac{1}{F_k(s^{t+1})k_{t+1}(s^t) + b_{t+1}(s^t)} \right) \times \{g_{t+1}^e(s_{t+1}) + b_{t+1}(s^t) - \frac{b_{t+2}(s^{t+1})}{R_{t+1}(s^{t+1})} - \tau_{t+1}^s(s^{t+1})w_{t+1}^s(s^{t+1})\psi_{t+1}(s^{t+1})h_{t+1}^s(s^{t+1}) - \tau_{t+1}^u(s^{t+1})w_{t+1}^u(s^{t+1}) \times [1 - \psi_{t+1}(s^{t+1})]h_{t+1}^u(s^{t+1}) + \tau_{t+1}^a(s^{t+1})\bar{g}(\cdot)(t+1)\}. \quad (53)$$

¹¹ Note that the multiplier Φ is non-negative and measures the disutility of future tax distortions.

4. Quantitative implementation

In this section we quantitatively solve both the non-stochastic and stochastic optimal policy models. Our solution approach follows [Arseneau and Chugh \(2012\)](#). In particular, we first calibrate the non-stochastic model with exogenous policy. Next, we solve the deterministic Ramsey problem, starting from the exogenous policy steady-state, using non-linear methods. Since we are interested in tax smoothing over the business cycle, we then approximate around the steady-state of the deterministic Ramsey problem to solve the stochastic problem and obtain near steady-state dynamics.

4.1. Functional forms

Following [Chari et al. \(1994\)](#) and [Stockman \(2001\)](#), we use a CRRA utility function:

$$u(\cdot) = \frac{\left\{ [c_t(s^t)]^{1-\sigma_1-\sigma_2} [\psi_t(s^t)l_t^f(s^t)]^{\sigma_1} [[1-\psi_t(s^t)]l_t^u(s^t)]^{\sigma_2} \right\}^{\sigma_3}}{\sigma_3} \quad (54)$$

where σ_1 and σ_2 are the weights to leisure in the utility function and σ_3 is the relative risk aversion parameter.

The production side is given by a CES production function that allows for capital-skill complementarity, since the latter has been shown to match the dynamics of the skill premium in the data (see e.g. [Krusell et al., 2000](#); [Lindquist, 2004](#); [Pourpourides, 2011](#)):

$$F(\cdot) = A_t(s_t) \left\{ \mu (h_t^{sf}(s^t))^\alpha + (1-\mu) \left[\rho (k_t^f(s^t))^\nu + (1-\rho) (h_t^{sf}(s^t))^\nu \right]^{\alpha/\nu} \right\}^{1/\alpha} \quad (55)$$

where $A_t(s_t)$ is total factor productivity; $\alpha < 1$, and $\nu < 1$ are the parameters determining the factor elasticities, i.e. $1/(1-\alpha)$ is the elasticity of substitution between capital and unskilled labour and between skilled and unskilled labour, whereas $1/(1-\nu)$ is the elasticity of substitution between equipment capital and skilled labour; and $0 < \mu, \rho < 1$ are the factor share parameters. In this specification, capital-skill complementarity is obtained if $1/(1-\alpha) > 1/(1-\nu)$ which also implies that the skill premium is decreasing in $\psi_t(s^t)$ and increasing in $k_t(s^{t-1})$.

The functional form for the relative skill supply is

$$q[\cdot] = \Psi [e_t(s^t)]^{\gamma\xi} [e_t^e(s^t)]^{\gamma(1-\xi)} \quad (56)$$

where $\Psi > 0$ is the productivity of skill-acquisition; $0 \leq \gamma < 1$ measures the productivity of skill-acquisition expenditure; and $0 < \xi \leq 1$ is the parameter that captures the extent of externalities in creating skill.

4.2. Exogenous policy and calibration

We next present the calibration and steady-state for the exogenous policy model. In particular, we obtain the steady-state of the decentralised competitive equilibrium (DCE). Given initial levels of k_0 , and b_0 , and the five policy instruments $\{\tau_t^s, \tau_t^u, \tau_t^k, \tau_t^g, g_t^e\}$, the non-stochastic DCE system is characterized by a sequence of allocations $\{c_t, h_t^s, h_t^u, \psi_t, k_{t+1}\}_{t=0}^\infty$, prices $\{w_t^s, w_t^u, r_t, R_t\}_{t=0}^\infty$, and the residual policy instrument $\{b_{t+1}\}_{t=0}^\infty$ such that (i) households maximise their welfare and firms maximise their profits, taking policy, prices and aggregate outcomes as given; (ii) the government budget constraint is satisfied in each time period and (iii) all markets clear. Thus, the non-stochastic DCE is composed of the non-stochastic form of the first-order conditions of the household (7), (8), (10), (11) and (23), the government budget constraint (17), the three first-order conditions of the firm (24)–(26), and the aggregate resource constraint (27).

4.2.1. Calibration

The non-stochastic model with exogenous policy is calibrated so that its steady-state is consistent with the annual US data.

Utility: [Table 1](#) below reports the model's quantitative parameters along with an indication of their source. Starting with the share of leisure for each skill type in utility, σ_1 and σ_2 , we calibrate these to 0.35 each so that, in the steady-state, the household devotes about one third of its time to working. The relative risk aversion parameter, $\sigma_3 = -2$ is commonly employed in business cycle models.

Production: The elasticities of substitution between skilled labour and capital and between unskilled labour and capital (or skilled labour) have been estimated by [Krusell et al. \(2000\)](#). Following the literature (see e.g. [Lindquist, 2004](#); [Pourpourides, 2011](#)), we also use these estimates to set $\alpha = 0.401$ and $\nu = -0.495$. The remaining parameters in the production function are calibrated to ensure the steady-state predictions of the model in asset and labour markets are consistent with the data. More specifically, the labour weight in the composite input share $\mu = 0.268$ is calibrated to obtain a labour share of income of approximately equal to 69% and the capital weight in the composite input share, $\rho = 0.52$, is calibrated to obtain a skill premium of 1.64. Both of these targets are consistent with the U.S. data for the period 1970–2011. The target value for the skill premium is from U.S. Census data and the share of labour income in GDP is

Table 1
Model parameters.

Parameter	Value	Definition	Source
$0 < \sigma_1 < 1$	0.350	Weight to skilled leisure in utility	Calibration
$0 < \sigma_2 < 1$	0.350	Weight to unskilled leisure in utility	Calibration
$\sigma_3 < 0$	-2.000	Coefficient of relative risk aversion	Assumption
$\frac{1}{1-\alpha} > 0$	1.669	Cap. equip. to unskilled labour elasticity	Assumption
$0 < \frac{1}{1-\nu} < \frac{1}{1-\alpha}$	0.669	Cap. equip. to skilled labour elasticity	Assumption
$0 < 1 - \mu < 1$	0.732	Share of composite input to output	Calibration
$0 < \rho < 1$	0.520	Share of cap. equip. to composite input	Calibration
$A > 0$	1.000	TFP	Assumption
$A^{*,k} > 0$	1.000	Investment-specific productivity	Assumption
$0 \leq \delta \leq 1$	0.007	Depreciation rate of capital	Calibration
$0 < \beta < 1$	0.960	Time discount factor	Calibration
$0 \leq \gamma < 1$	0.172	Relative skill supply elasticity	Calibration
$0 < \xi \leq 1$	0.700	The share of externality	Calibration
$\Psi > 0$	1.000	Productivity of skill-acquisition	Assumption
τ^k	0.310	Capital income tax rate	Data
τ^u	0.200	Unskilled labour tax rate	Data
τ^s	0.250	Skilled labour tax rate	Data
τ^n	0.220	Effective labour tax rate	Data
τ^a	0.000	Skill-acquisition expenditure tax rate	Assumption
$g^e > 0$	0.046	Government spending	Calibration

from the BEA data on personal income.¹² We also normalize the steady-state values of TFP and investment on equipment to unity (i.e. $A = A^{*,k} = 1$).

Depreciation and time preference: The depreciation rate of capital $\delta = 0.07$ is calibrated to obtain an annual capital to output ratio of about 1.96, which is consistent with the annual data reported by the BEA on capital stocks.¹³ The time discount factor, $\beta = 0.96$, is set to obtain a post-tax post-depreciation annual real rate of return on capital of roughly 4.17%, which coheres with the 4.19% obtained in the data from the World Bank.¹⁴

Relative skill supply: We normalise skill-acquisition productivity, Ψ , to unity. To match the share of skilled workers in total population, ψ , of roughly 44% in the data, we set the productivity of skill-acquisition expenditure, γ , equal to 0.172. This share is consistent with the data from the 2010 U.S. Census which indicates that 43% of the population has a college degree.¹⁵ It also corresponds with a related data set by [Acemoglu and Autor \(2011\)](#) which implies that the average share of the labour force with a college degree is approximately 45%. We set ξ equal to 0.7, which produces a steady-state skill expenditure to output ratio of about 4.3%, which is consistent with the data on tertiary education expenditure for the last four decades from the U.S. National Center for Education Statistics, Digest of Education Statistics.¹⁶

Tax rates and government spending: Finally, we use the ECFIN effective capital and labour tax rates from [Martinez-Mongay \(2000\)](#) to obtain an average tax rate for capital and labour.¹⁷ Therefore, we set the tax rate for capital income $\tau^k = 0.31$ and the two labour income tax rates $\tau^u = 0.20$ and $\tau^s = 0.25$.¹⁸ Given that it is difficult to obtain data which match well with the skill-acquisition expenditure tax (subsidy) rate, τ^a , we set it to zero for the exogenous policy model. We finally set the steady-state value $g^e = 0.046$, to obtain a steady-state debt-to-output ratio, $b/Y = 53\%$, which is equal to the average debt to GDP ratio obtained in the data.¹⁹

Steady-state: The steady-state of the DCE defined and calibrated above is presented in [Table 2](#). The results indicate that the model's predictions for the great ratios match those implied by the data quite well. For example, in the data for 1970–2011: $k/y = 1.895$, $c/y = 0.640$, $i/y = 0.146$, $g^e/y = 0.203$ and $b/y = 0.530$.²⁰ Moreover, the share of skill-acquisition expenditure in GDP, e/y , coheres with US total expenditures for colleges and universities as a share of output of about

¹² The data source is the Current Population Survey, 2011 Annual Social and Economic Supplement from the U.S. Census Bureau.

¹³ Specifically, the BEA Table 1.1 on fixed-assets has been used to obtain the time series for capital stock for 1970–2011.

¹⁴ The data refers to the annual real interest rate from World Bank Indicators database for the period 1970–2011 (i.e. FR.INR.RINR).

¹⁵ This information is obtained from [Table 4](#) of the Census Bureau, Survey of Income and Program Participation.

¹⁶ For instance, in the version of the model without the externality ($\xi = 1$) that will be discussed below, skill expenditure is about 5.3% of output.

¹⁷ In particular, we use the LITR and KITN rates for effective average labour and capital taxes respectively for 1970–2011, as they treat self-employed income as capital income in the calculations.

¹⁸ Note that the calculation of the effective labour income tax rate is equal to 0.22. But since we assume that the skilled and unskilled labour income is taxed differently we decompose the labour income tax into skilled and unskilled tax so as the weighted average of the two tax rates equals 0.22.

¹⁹ The source of that time series is: FRED Economic Data on Gross Federal Debt as a percentage of GDP, 1970–2011.

²⁰ Note that if model prediction for the cost of becoming skilled, $e/y = 0.0433$, is added to the c/y ratio from the model, the sum is very close to the c/y ratio in the data.

4.3% since the mid-seventies.²¹ As pointed out above, the remaining steady-state variables in the exogenous model have been calibrated to match their values in the data.

4.3. Deterministic Ramsey

The deterministic version of the Ramsey problem in (42)–(50) is summarised in Appendix D, (see Eqs. (D.1)–(D.16)) and is solved iteratively, conditional on the calibration described in the previous section. In particular, we first guess a value for Φ and solve Eqs. (D.1)–(D.15) for an allocation $\{c_t, h_t^s, h_t^u, \psi_t, k_{t+1}\}_{t=0}^T$. Then we test whether Eq. (D.16) is binding and increase or decrease the value of Φ if the budget is in deficit or surplus respectively.

The initial conditions for the model's state variables are given by the non-stochastic exogenous steady-state (see Table 2). For the terminal values of the forward looking variables, we assume that after T years the dynamic system has converged to its Ramsey steady-state. This implies that the appropriate terminal conditions are obtained by setting the values for these variables equal to those of the preceding period.

The final system is given by $[(5 \times T) + 1]$ equations, which is solved non-linearly using standard numeric methods (see, e.g. Adjemian et al., 2011). This gives the dynamic transition path from the exogenous to the optimal steady-state. We set $T=250$ to ensure that convergence is achieved. Our results show that this occurs for all endogenous variables within 150 years.²² After we find the optimal allocation for $\{c_t, h_t^s, h_t^u, \psi_t, k_{t+1}\}_{t=0}^T$ we obtain $w_t^s = \tilde{F}_{h^s}(t)$, $w_t^u = \tilde{F}_{h^u}(t)$ and $r_t = \tilde{F}_k(t)$. Additionally, we solve for $\tau_t^s, \tau_t^u, \tau_t^k$ and τ_t^a using the non-stochastic form of (7), (8), (11) and (23) respectively.

The Ramsey steady-state is reported in Table 3. The results are consistent with the messages from the literature initiated by Chamley (1986) on dynamic Ramsey taxation in a deterministic environment (see e.g. Ljungqvist and Sargent, 2012, ch. 16 for a review of this literature). As expected, allowing the government a complete instrument set results in a zero capital tax rate in the long-run.

Compared with the steady-state of exogenous policy, a Ramsey government would increase capital accumulation in the steady-state, by eliminating the *intertemporal* wedge. Moreover, given the externalities in skill creation, the government finds it optimal to subsidise expenditure on skill-acquisition by about 40% in this setup, which leads to an increase in the relative skill supply. Since this increase is encouraged by the skill subsidy, which stimulates overall effective skilled labour hours, the government can tax at a slightly higher rate the higher labour income source, i.e. skilled labour income, which allows for more tax revenue to be generated. Notice that the progressivity of optimal labour income taxation is very small quantitatively. The fall in the skill premium under Ramsey policy suggests that the increase in the relative skill supply has a relatively stronger quantitative impact than the increase in the capital stock. Finally, the government is able to finance part of the required public spending in the long-run from accumulated assets.

We next study the transition dynamics associated with Ramsey policy. Fig. 1 illustrates the dynamic paths implied by optimal policy for the capital tax, the two labour taxes, the skill-acquisition expenditure tax and debt to output as the economy evolves from the exogenous steady-state to the Ramsey steady-state.

The first panel of Fig. 1 shows that in period 1 skilled and unskilled labour are subsidised at rates of 24.1% and 21.3% respectively; and skill-acquisition expenditure is taxed at a rate of 3.6%. In period 2, skilled and unskilled labour taxes are 29.2% and 27.5% respectively and eventually converge to their steady-state values reported in Table 3. Also in period 2, skill-acquisition is subsidised at a rate of 37.7% and converges to about 40% in the steady-state.

The second panel of Fig. 1 shows that in period 1, since capital is already in place, capital income is taxed at a confiscatory rate (approximately 306%). In period 2, the capital income tax is 0.9% and then converges slowly to zero. The high capital taxation in the first period allows the government to create a stock of assets by lending to the household. Government assets increase in future periods and their income is used to subsidise skill-acquisition expenditure and to compensate for the losses from foregone capital income taxation, without the need to resort to high labour income taxes. These transition paths are consistent with previous research.

4.4. Stochastic processes

To move to the analysis of the stochastic Ramsey problem, we need to define the stochastic processes that drive economic fluctuations. In what follows we designate a stochastic state s_t at time t that determines exogenous shocks to total factor productivity, $A_t(s_t)$, investment-specific technological progress, $A_t^{*,k}(s_t)$, and government expenditures, $g_t^e(s_t)$. Following the literature, $A_t(s_t)$, $A_t^{*,k}(s_t)$ and $g_t^e(s_t)$ are assumed to follow stochastic AR(1) processes:

$$\log A_{t+1}(s_{t+1}) = (1 - \rho_A) \log A + \rho_A \log A_t(s_t) + \varepsilon_{t+1}^A \quad (57)$$

$$\log A_{t+1}^{*,k}(s_{t+1}) = (1 - \rho_{A^*k}) \log A^{*,k} + \rho_{A^*k} \log A_t^{*,k}(s_t) + \varepsilon_{t+1}^{A^*k} \quad (58)$$

²¹ Tertiary education expenditure in the USA was higher in the early 1970s, but declined continuously since 1970 to converge to less than 4% over the last two decades.

²² See Fig. 1 below for an illustration of convergence using the policy instruments.

Table 2
Steady-state of exogenous policy.

$\frac{c}{y}$	$\frac{k}{y}$	$\frac{i}{y}$	$\frac{e}{y}$	$\frac{b}{y}$	$\frac{g^e}{y}$	$\frac{W^s}{W^u}$	r^{net}	ψ
0.5823	1.9629	0.1374	0.0433	0.5300	0.2371	1.6400	0.0417	0.4400

Table 3
Steady-state of optimal policy.

$\frac{c}{y}$	$\frac{k}{y}$	$\frac{e}{y}$	$\frac{b}{y}$	$\frac{g^e}{y}$	$\frac{W^s}{W^u}$
0.5565	2.6859	0.0605	-0.8348	0.1950	1.4285
r^s	r^u	r^k	r^d	r^{net}	ψ
0.2654	0.2577	0.0000	0.4005	0.0417	0.4819

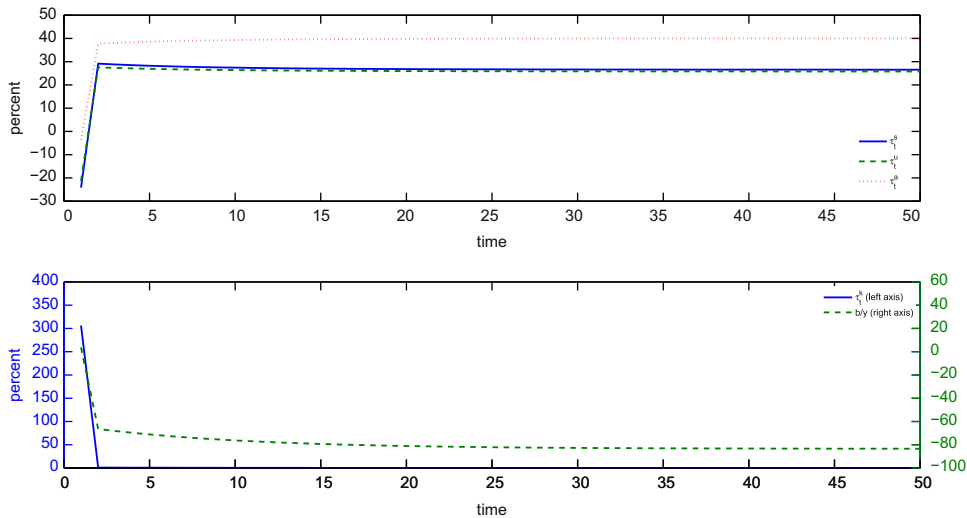


Fig. 1. Transition path of the policy instruments.

$$\log g_{t+1}^e(s_{t+1}) = (1 - \rho_{g^e}) \log g^e + \rho_{g^e} \log g_t^e(s_t) + \varepsilon_{t+1}^{g^e} \tag{59}$$

where ε_t^A , $\varepsilon_t^{A^*k}$ and $\varepsilon_t^{g^e}$ are independently and identically distributed Gaussian random variables with zero means and standard deviations given respectively by σ_A , σ_{A^*k} and σ_{g^e} .

The values for the AR(1) coefficient and the standard deviation for the government expenditures are data-based and are estimated to be: $\rho_{g^e} = 0.65$ and $\sigma_g = 0.017$.²³ The AR(1) coefficient and the standard deviation for the investment specific technological progress are calibrated to $\rho_{A^*k} = 0.70$, $\sigma_{A^*k} = 0.012$ so that the correlation of investment with output and the relative standard deviation of investment to output are approximately equal to 0.85 and 4.3 respectively.²⁴ The autocorrelation parameter of TFP is set equal to 0.80 following Chari et al. (1994) and Stockman (2001), while σ_A is calibrated to match the volatility of output observed in the BEA data.²⁵ More specifically, the standard deviation for TFP is set $\sigma_A = 0.8\%$ to obtain a volatility for output from 1970 to 2011 of about 1%.

²³ The government spending series refers to government consumption expenditures and gross investment from NIPA Table 1.1.5 (1970–2011). To calculate the statistical properties of the cyclical component of the series, we take logs and apply the HP-filter with smoothing parameter equal to 100.

²⁴ The investment series refers to Private Fixed Investment and it is obtained from NIPA Table 5.3.5 (1970–2011). Cyclical investment is calculated using the HP-filter as above.

²⁵ The time series for GDP from 1970 to 2011 is obtained from NIPA Table 1.1.5. Cyclical output is calculated using the HP-filter as above.

4.5. Stochastic Ramsey

We next approximate the dynamic equilibrium paths due to three exogenous shocks using a first-order approximation of the decision rules of the equilibrium conditions under optimal policy in (42)–(46), around the optimal deterministic steady-state of these conditions described above.²⁶ As is common in the literature when characterizing policy dynamics, we also make the auxiliary assumption that the initial state of the economy at $t=0$ is the steady-state under optimal policy. To calculate the business cycle statistics of the relevant quantities of the model under optimal policy, we conduct simulations by shocking all of the exogenous processes, obtain the required moments for each simulation and then calculate their mean value across the simulations. We undertake 1000 simulations, each 400 periods long, to ensure that we have enough periods to approximate lifetime quantities and that the model generated data is stationary.

5. Cyclical properties

We next present the results regarding the key second moments of the stochastic optimal policy problem. We conduct this analysis for the benchmark model developed above and we also examine the robustness of the main results to changes in the skill creation technology.

5.1. Capital and labour taxes

We start with the cyclical properties of Ramsey taxation in the benchmark model. Table 5 presents results on the means, standard deviations relative to output and correlations for the optimal allocations, policy and wedges. The results regarding optimal capital and labour taxes are largely consistent with the literature and thus extend previous findings to a setup with capital-skill complementarity and endogenous skill supply. In particular, the *ex ante* tax rate on capital is effectively zero and has the highest volatility of all tax instruments when debt is state-contingent. Moreover, when debt is not state-contingent, the state contingent private assets and *ex post* capital taxes are near zero, have low correlations with output and are the most volatile of the tax instruments.

Also consistent with the labour tax-smoothing literature, both labour taxes have very low standard deviations relative to output, as the government finds it optimal to minimise the distortions introduced by labour taxes by keeping them relatively smooth over the business cycle and by letting the remaining state-contingent policy instruments respond to exogenous shocks. This finding also coheres with the results regarding the volatility of labour taxes in Werning (2007), albeit in a different setup, since we allow for worker, as opposed to household, heterogeneity.²⁷

Despite the similarity regarding their standard deviations, the labour income taxes exhibit different correlations with output in our setup. In particular, the tax rate on skilled labour income is strongly pro-cyclical, whereas the tax rate on unskilled labour income has a low and negative correlation with output. This difference maintains for both type of technology shocks, since the correlations of τ^s with A and $A^{*,k}$ are positive and high, whereas the correlations of τ^u with A and $A^{*,k}$ are close to zero. Note that the differences in the correlations of the two labour income taxes with the technology shocks are mirrored in the differences in the correlations of skilled and unskilled labour hours with the technology shocks. In particular, skilled labour hours are strongly correlated with both technology shocks, where unskilled labour hours exhibit very low correlations.²⁸ Finally, the correlations for all tax instruments with g^e are very low, as the government uses state-contingent government assets to accommodate temporary pure public finance shocks.

5.2. Skill-acquisition subsidy

We next find that the skill-acquisition subsidy is about four times more volatile than the remaining taxes that affect static margins (i.e. τ^s and τ^u), but not as volatile as the tax instruments that affect the intertemporal margins of adjustment (i.e. the capital and assets taxes). Quantitatively, its standard deviation implies very little fluctuations in τ^a over the business cycle, as, 95% of the time it varies between 39.8% and 40.3%, given that $\sigma_V = 0.0047$. Moreover, we find that τ^a is pro-cyclical and is positively correlated with both technology processes.²⁹

It should be noted here that the volatility of τ^a is due to the existence of the policy distortions in the labour markets and therefore is a characteristic of the second-best nature of Ramsey optimal policy. Recall from the analysis in Section 2.6 that if the government could finance public spending with lump sum instrument, so that $\tau_t^s(s^t) = \tau_t^u(s^t) = \tau_t^k(s^t) = 0$, the optimal skill-acquisition subsidy which would eliminate the wedge introduced by the externality in skill creation, $\bar{\tau}_t^a(s^t)$, is given by (31).

²⁶ We use the perturbation methods in Schmitt-Grohe and Uribe (2003) to solve the dynamic model.

²⁷ Werning (2007) shows that the optimal volatility of labour taxes for households of different ability is zero when the utility function incorporates a constant Frisch-labour supply elasticity. In our model, it can also be shown that the standard deviations of the labour taxes are optimally zero for the constant Frisch-labour supply elasticity case.

²⁸ The dynamic behaviour of the tax rates and their relationships with the remaining endogenous variables are further discussed below when we present responses to temporary shocks.

²⁹ The responses of τ^a to exogenous technology shocks will be discussed in more detail below in the impulse response analysis.

Table 4
Parameters for stochastic processes.

Parameter	Value	Definition
σ_A	0.008	Std. dev. of TFP
ρ_A	0.800	AR (1) coef. of TFP
$\sigma_{A^{*k}}$	0.012	Std. dev. of capital equipment
$\rho_{A^{*k}}$	0.700	AR(1) coef. of capital equipment
σ_{g^e}	0.017	Std. dev. of public spending
ρ_{g^e}	0.650	AR (1) coef. of public spending

Table 5
Stochastic results base model.

x_i	\bar{x}_i	$\frac{\sigma_{x_i}}{\sigma_Y}$	$\rho(x_i, Y)$	$\rho(x_i, A)$	$\rho(x_i, A^{*k})$	$\rho(x_i, g^e)$
Allocations						
c	0.1315	0.5386	0.8655	0.7787	-0.0656	-0.0993
k	0.6347	3.4917	0.6404	0.3900	0.2194	-0.0495
h^s	0.3700	0.3963	0.7114	0.6038	0.7424	0.2452
h^u	0.1715	0.4608	-0.0689	0.0616	0.3771	0.2313
ψ	0.4819	0.2796	0.9252	0.7808	0.1826	-0.0220
w^s	0.6878	2.6241	0.9283	0.8346	0.0444	-0.0571
w^u	0.4815	2.0725	0.9275	0.8471	0.0131	-0.0640
$\overline{w^s}$	1.4285	0.7358	-0.8752	-0.8978	0.2194	0.1126
$\overline{w^u}$						
Policy and wedges						
τ^s	0.2654	0.0776	0.7996	0.6379	0.7352	0.2121
τ^u	0.2577	0.0595	-0.2005	-0.0088	0.1834	0.1817
$\tau^{k,a}$	-6.07e-5	2.7963	0.9370	1.0000	0.0037	0.0020
$\tau^{k,p}$	-0.0047	26.2437	-0.3532	-0.4371	0.0859	0.0983
ξ	7.51e-4	13.7352	0.3526	0.4349	-0.0835	-0.0984
τ^a	0.4005	0.3144	0.8791	0.6621	0.5663	0.1092
ψ^w	0.0215	0.3125	0.9242	0.7629	0.2442	-0.0036
ψ^m	0.0740	0.2069	0.9252	0.7808	0.1826	-0.0220
ψ^p	-0.0525	0.1070	0.9101	0.7185	0.3604	0.0321

It can be shown analytically that for a skill-acquisition technology of the form of (56), this quantity is constant and equal to the size of the externality, i.e. $\bar{\tau}_t^a(s^t) = 1 - \xi$.³⁰ Therefore, since we do not have a lump sum tax instrument in the model, the volatility in the optimal τ^a reported in Table 5 is due to the presence of non-zero labour income taxes.

5.3. Skill acquisition wedge

Our results so far suggest that when the baseline neoclassical model is extended to allow for capital-skill complementarity and endogenous skill supply, the government still finds it optimal to keep the labour market wedges smooth, while the capital tax wedge is effectively eliminated. However, the correlation of the labour wedges with the exogenous technology processes differs for skilled and unskilled labour.

As Eq. (28) makes clear, the wedge in relative skill supply, ψ^w , depends on the subsidy to skill acquisition as well as on the labour taxes which define ψ^p . Moreover, because of the externality in skill creation, ψ^w also depends on ψ^m . To further investigate optimal policy in relation to the wedge in relative skill supply, we present in Table 6 the means and standard deviations of ψ^w and its components for the model with exogenous policy and under the Ramsey planner.

The results in Table 6 show that optimal policy reduces the total wedge in the steady state, by turning the policy component into a subsidy, and also reduces the standard deviation of the total component, by setting the policy instruments over the business cycle to smooth the policy component of the wedge.

5.4. Impulse responses

To further analyse the effects of optimal policy over the business cycle and examine the optimal response of taxation to changes in exogenous productivity, we plot the impulse responses of key endogenous variables, as percent deviations from

³⁰ Note that (56) implies that $\bar{g}_w(s^t) = 1/\gamma\Psi(\psi_t(s^t))^{(1/\gamma)-1}$ and $\bar{g}_\psi(s^t) = (1/\gamma\xi)\Psi(\psi_t(s^t))^{(1/\gamma)-1}$.

Table 6
Skill acquisition wedge.

	Means		Standard deviations (%)	
	Exog. policy	Ramsey	Exog. policy	Ramsey
ψ^w	0.0842	0.0215	0.17	0.15
ψ^m	0.0477	0.0740	0.08	0.10
ψ^p	0.0365	-0.0525	0.10	0.05

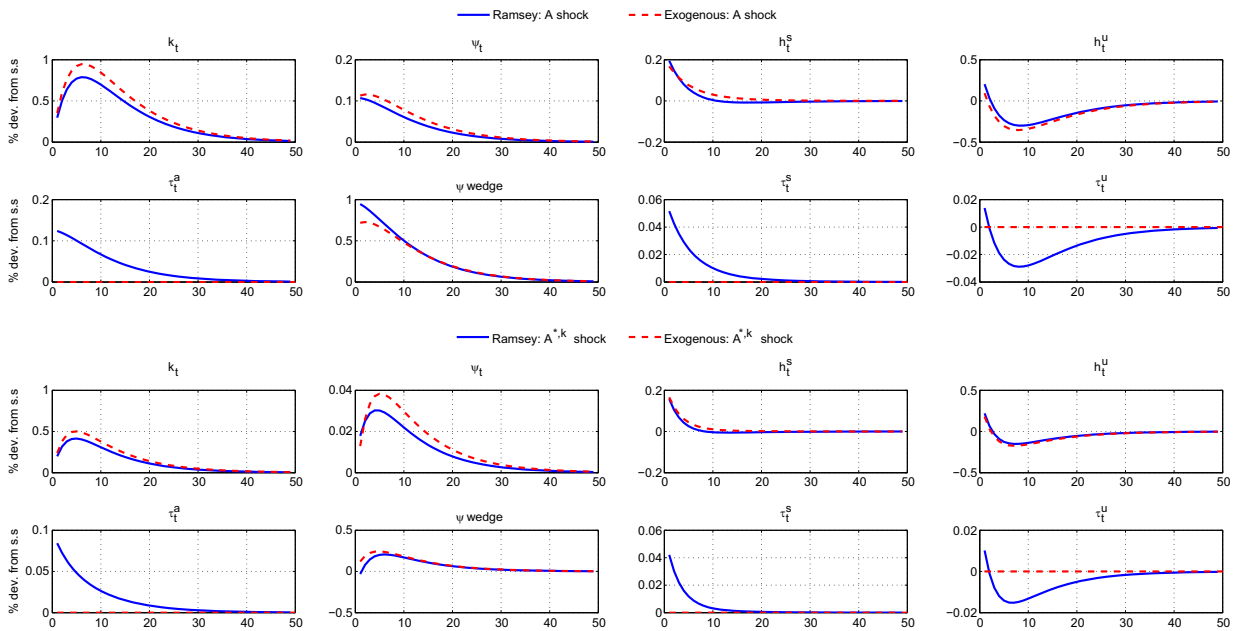


Fig. 2. Impulse responses to 1% temporary productivity shocks.

their respective steady state, after a temporary 1% shock to the exogenous distributions in $\varepsilon_t^A, \varepsilon_t^{A^{*,k}}$. These plots are shown in Fig. 2 above.³¹ To contextualise these effects we also plot the same impulse responses for the model with exogenous policy discussed in Section 4. The continuous lines show the responses under Ramsey policy, while the dashed lines the responses under exogenous policy.

The first point to note in Fig. 2 is that the government optimally smooths the reactions of the inputs in the production process, in response to exogenous technology shocks, relative to the case with exogenous policy. To achieve this increased smoothness in the labour markets, the government needs to make the wedges, which affect the household choices for h^s, h^u and ψ , move in the same direction as these quantities.

In addition to the direct impact of the investment-specific shock, both technology shocks increase the productivity of skilled labour relative to that of unskilled indirectly, via the increase in the capital stock. These productivity effects create more incentives for the household to increase the relative quantity of skilled members. These effects also increase the labour hours of the more productive type of labour and decrease the work time of the least productive (after an initial increase), since both types of workers provide equally valued leisure time.

To mediate the above reactions, the government needs to increase the wedges in skilled labour hours, τ^s , and relative skill supply, ψ^w , while concurrently decreasing the wedge in unskilled labour hours, τ^u . Note that despite the increase in τ^a , the relative skill supply wedge has increased because ψ^w also depends on τ^s and τ^u positively and negatively, respectively, so the changes in the labour taxes tend to raise ψ^w . These movements give rise to the correlations for taxes and wedges summarised in Table 5.

³¹ To save on space, we do not present the impulse responses to ε_t^{β} , since the tax instruments respond very little quantitatively to these shocks, as the government uses state-contingent government assets to accommodate temporary pure public finance shocks.

5.5. Autocorrelations

The results in Table 7 suggest that the labour income taxes and the *ex ante* capital income tax in this model inherit the properties of the exogenous processes. For example, the autocorrelations of these instruments follow the autocorrelations of the exogenous processes. Thus, when the shocks are autocorrelated as in Table 4, so are the tax rates. However, if we assume that the shocks follow *iid* processes, the autocorrelation of the tax rates generally becomes very small.

In contrast, the autocorrelations of the *ex post* capital tax and the private assets tax do not follow the autocorrelations of the exogenous processes. This is again similar to previous findings in the literature. The results in Table 7 finally show that the skill-acquisition subsidy also inherits the properties of the exogenous processes, thus working in a similar fashion to the remaining taxes that affect the static margins.

5.6. Alternative assumptions regarding skill supply

To further evaluate the importance of skill creation and endogenous relative skill supply for optimal taxation in the benchmark model analysed above, we obtain the means and the standard deviations relative to output for taxes and wedges for three alternative assumptions regarding relative skill supply. These are presented in Table 8.³²

5.6.1. No externalities

The first case considered is a model without externalities in skill creation, obtained by setting $\xi = 1$ in $q(\cdot)$. In this case, $\psi^w = \psi^p$, which implies that the skill-acquisition subsidy reflects only policy distortions and should be kept to a minimal level at the steady-state. Moreover, labour income taxation becomes mildly regressive in the steady-state. Despite these long-run adjustments, the relative standard deviations of taxes and subsidies do not change significantly, whereas the wedge in relative skill supply has lower volatility, since this is only driven by the volatility of the policy component of the wedge.

5.6.2. Productivity shock in skill creation

We next consider a case where the skill creation technology includes a productivity series, intended to capture exogenous factors that affect skill creation and social mobility. For instance, it is generally accepted (see e.g. Goldin and Katz, 2008) that social mobility and enrolment in tertiary education, which in turn determine relative skill supply, can depend on policy interventions. These may be the provision of access to skill-related education to members of deprived communities and structural changes in the provision of primary and secondary education (e.g. the curriculum of studies and the definition of school catchment areas).

To broadly capture these potential interventions, we allow for Ψ_t to be determined by an exogenous AR(1) process:

$$\log \Psi_{t+1}(s_{t+1}) = (1 - \rho_\psi) \log \Psi + \rho_\psi \log \Psi_t(s_t) + \varepsilon_{t+1}^\Psi \quad (60)$$

where ε_t^Ψ is an independently and identically distributed Gaussian random variable with zero mean and a standard deviation given by σ_ψ . We set $\rho_\psi = 0.95$ and $\sigma_\psi = 0.009$, so that the model generated series for relative skill supply is consistent with the persistence and standard deviation (relative to output) of the data, for the model with exogenous policy calibrated as in Section 4. We use annual data for the share of college educated to total working population measured in efficiency units from the 1970s from Acemoglu and Autor (2011) and GDP data from the US NIPA accounts to find that the autocorrelation of the cyclical component of relative skill supply is 0.47 and its standard deviation relative to output is 0.27.

The results in Table 8 suggest that optimal labour tax smoothing is not affected in this case and there is only a small increase in the volatility of the skill-acquisition subsidy. The standard deviations of the wedge in relative skill supply and of both of its components have increased significantly. However, given that the increased volatility in skill creation affects efficient as well as market outcomes, the government does not find it optimal to change the cyclical properties of the tax instruments significantly.

5.6.3. Fixed relative skill supply

The final case we consider is when the relative skill supply is fixed, $\psi_t(s^t) = \psi^f = 0.44$ and thus exogenous to the household. This assumption implies that skill creation expenditure also remains fixed at the level that is required to support $\psi_t(s^t) = 0.44$. However, since there is no margin of adjustment for the household regarding the choice of ψ_t in this case, we also normalise $\tau_t^a(s^t) = 0$, since it does not have a role to play in affecting a wedge in the model.³³ In particular, assuming that this restriction is imposed on the market economy, whilst the primitives of the model allow the social planner to choose relative skill supply, we can define the wedge in relative skill supply as the difference between the efficient and the

³² In all cases considered below, we do not re-calibrate the model, except for the specific changes we discuss in the specification for $\psi_t(s^t)$ in (3), to isolate the effects of the changes in skill supply. In each case, we re-calculate the deterministic Ramsey steady-state and obtain the required moments as the economy fluctuates around this steady-state. We do not present results for correlations and autocorrelations as these are similar with those obtained under the benchmark model.

³³ Note that if $\tau_t^a(s^t)$ was left as explicit choice of the government, it would be equivalent to a lump-sum tax, which would violate the second-best nature of the problem that we want to study.

Table 7
Autocorrelations.

	Autocorrelated shocks	iid shocks
τ^s	0.7512	0.0121
τ^u	0.8593	0.0952
$\tau^{k,a}$	0.7897	-0.0023
$\tau^{k,p}$	-0.0362	-0.4203
ξ	-0.0328	-0.4141
τ^a	0.8945	0.2041

Table 8
Stochastic results alternative specifications of q .

X_i	$\xi = 1$		AR process ψ		$\psi_t = 0.44$	
	\bar{X}_i	$\frac{\sigma_{X_i}}{\sigma_Y}$	\bar{X}_i	$\frac{\sigma_{X_i}}{\sigma_Y}$	\bar{X}_i	$\frac{\sigma_{X_i}}{\sigma_Y}$
τ^s	0.2415	0.0766	0.2654	0.0794	0.2574	0.0755
τ^u	0.2512	0.0570	0.2577	0.0605	0.2160	0.1136
$\tau^{k,a}$	-6.07e-5	2.7655	-9.81e-5	2.7168	-3.15e-4	5.3464
$\tau^{k,p}$	-0.0047	25.9045	-0.0080	24.6766	-0.0049	29.9299
ξ	-4.91e-4	21.6957	0.0015	12.4066	5.03e-4	17.9873
τ^a	0.0745	0.3621	0.4005	0.3995	-	-
$\psi^w(\psi^{w,f})^\dagger$	0.0288	0.1372	0.0215	2.5353	0.1439	1.4270
ψ^m	-	-	0.0740	1.1939	-	-
ψ^p	0.0288	0.1372	-0.0525	1.3443	-	-

$^\dagger \psi^{w,f}$ defined in Eq. (61) refers to the results in the last two columns.

market outcome for $\psi_t(s^t)$. This wedge in relative skill supply can be obtained by first combining (B.10) with (C.4) in Appendix C, to write the efficient condition in terms of $\psi_t(s^t)$ and then comparing with the market condition $\psi_t(s^t) = 0.44$ to define the wedge as

$$\psi^{w,f} = \frac{u_{\psi}(s^t)}{u_c(s^t)} + h_t^s(s^t)F_{h^{s,f}}(s^t) - h_t^u(s^t)F_{h^{u,f}}(s^t) - \bar{g}_{\psi}(s^t) + \psi_t(s^t) - \psi^f. \quad (61)$$

As can be seen, $\psi^{w,f}$ is not affected by the policy instruments. The means and the relative standard deviations under $\psi_t(s^t) = 0.44$ are presented in the last column of Table 8. As can be seen, this specification leads to a relatively more progressive labour income taxation and a qualitative change in the relative standard deviations for income taxes, since τ^u is now about 50% more volatile than τ^s (recall that under endogenous relative skill supply in the benchmark model, τ^s is about 30% more volatile than τ^u). However, despite the qualitative difference, the magnitude of the volatilities remains very small, so that labour tax smoothing maintains quantitatively in this specification. Since this specification directly restricts the quantity of relative skill supply in the market economy, the implied wedge derived above is higher and significantly more volatile, compared with the wedge due to externalities and policy distortions.

6. Conclusions

Motivated by the empirical relevance of the wage-skill premium and the roles played by capital-skill complementarity, the relative supply of skilled labour and capital augmenting technical change, this paper contributed to the tax smoothing literature by undertaking a normative investigation of the quantitative properties of optimal taxation of capital and labour income, as well as skill-acquisition expenditure, in the presence of aggregate shocks to total factor productivity (TFP), investment-specific technological change and government spending.

We found that labour tax smoothing maintained in this environment. In particular, labour taxes were the smoothest over the business cycle, followed by a subsidy to skill acquisition, while capital taxes were the most volatile of all instruments. We considered different assumptions and imperfections regarding the mechanism that generated the relative skill supply and further found that although these mattered for the volatility of the wedge between marginal rates of substitution and transformation in skill acquisition, the volatility of the tax instruments was not affected quantitatively.

Acknowledgments

We would like to thank the editor, B. Ravikumar, an anonymous referee, Fabrice Collard, Guido Cozzi, Richard Dennis, Wei Jiang, Charles Nolan, Apostolis Philippopoulos, Kjetil Storesletten, Pedro Telles and seminar participants at the CEF 2014

conference at Oslo for helpful comments and suggestions. The usual disclaimer applies. We are also grateful for financial support from the ESRC, Grant No: ES/I902414/1.

Appendix A. Household's first-order conditions

The household's first-order conditions for consumption, skilled labour supply, unskilled labour supply, debt, capital and the relative skill supply are given respectively by the following relations:

$$u_c(s^t) = \lambda_t(s^t) \quad (\text{A.1})$$

$$u_{h^s}(s^t) = -\lambda_t(s^t)\psi_t(s^t)\{[1 - \tau_t^s(s^t)]w_t^s(s^t)\} \quad (\text{A.2})$$

$$u_{h^u}(s^t) = -\lambda_t(s^t)[1 - \psi_t(s^t)][1 - \tau_t^u(s^t)]w_t^u(s^t) \quad (\text{A.3})$$

$$u_{\psi}(s^t) = -\lambda_t(s^t)\{h_t^s(s^t)[1 - \tau_t^s(s^t)] \times w_t^s(s^t) - h_t^u(s^t)[1 - \tau_t^u(s^t)]w_t^u(s^t)\} + \lambda_t(s^t)[1 - \tau_t^a(s^t)]g_{\psi}(s^t) \quad (\text{A.4})$$

$$\pi_t(s^t)\lambda_t(s^t)p_t(s_{t+1}|s^t) = \beta\pi_{t+1}(s^{t+1})\lambda_{t+1}(s^{t+1}) \quad (\text{A.5})$$

$$\pi_t(s^t)\lambda_t(s^t)A_t^k(s_t) = \beta \sum_{s^{t+1}|s^t} \{\pi_{t+1}(s^{t+1})\lambda_{t+1}(s^{t+1}) \times [r_{t+1}(s^{t+1})[1 - \tau_{t+1}^k(s^{t+1})] + (1 - \delta)A_{t+1}^k(s_{t+1})]\}. \quad (\text{A.6})$$

Appendix B. Efficient allocations

The following Appendix follows the analysis in [Arseneau and Chugh \(2012\)](#), [Appendix C](#). We first obtain the conditions that characterise the efficient allocations for the model economy arising from the social planner's problem. We then describe the static and dynamic efficiency conditions in terms of marginal rates of substitution (MRS) and marginal rates of transformation (MRT), where the MRTs are derived from economic primitives, independent of optimisation.

B.1. Social planner's problem

The social planner's stochastic problem can be written as follows:

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) u[c_t(s^t), \psi_t(s^t), h_t^s(s^t), h_t^u(s^t)] \quad (\text{B.1})$$

$$c_t(s^t) + g_t^e(s_t) + \bar{g}(\cdot) + A_t^k(s_t)(k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1})) - Y(\cdot) = 0 \quad (\text{B.2})$$

where $Y(\cdot)$ denotes the production technology as a function of the economy's resources, i.e. $Y(\cdot) \equiv Y[h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_t(s^{t-1}), s_t]$; and $\bar{g}(\cdot)$ denotes skill acquisition expenditure as a function of relative skill supply, i.e. $e_t(s^t) \equiv \bar{g}(\cdot) = \bar{g}[\psi_t(s^t)]$ and $\bar{g}_{\psi}(s^t) > 0$. Note that in the social planner's problem given by (B1)–(B2), the time constraints (1)–(2) are used to substitute out $l_t^s(s^t)$ and $l_t^u(s^t)$ in the $l_t^{*s}(s^t)$ and $l_t^{*u}(s^t)$ identities respectively which are arguments in the utility function (6); and the capital evolution Eq. (4) has been substituted into the aggregate resource constraint (B.2) for $i_t(s^t)$. The planner maximises (B.1) subject to the sequence of constraints in (B.2) by choosing $\{c_t(s^t), h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_{t+1}(s^t) \forall t=0, \infty\}$, given initial values for k_0 . This yields the following five first-order conditions:

$$u_c(s^t) + \lambda_t^p(s^t) = 0 \quad (\text{B.3})$$

$$u_{h^s}(s^t) - \lambda_t^p(s^t)Y_{h^s}(s^t) = 0 \quad (\text{B.4})$$

$$u_{h^u}(s^t) - \lambda_t^p(s^t)Y_{h^u}(s^t) = 0 \quad (\text{B.5})$$

$$u_{\psi}(s^t) + \lambda_t^p(s^t)\bar{g}_{\psi}(s^t) - \lambda_t^p(s^t)Y_{\psi}(s^t) = 0 \quad (\text{B.6})$$

$$\pi_t(s^t)\lambda_t^p(s^t)A_t^k(s_t) - \beta \sum_{s^{t+1}|s^t} \pi_{t+1}(s^{t+1})\lambda_{t+1}^p(s^{t+1})A_{t+1}^k(s_{t+1}) \times (1 - \delta) - \beta \sum_{s^{t+1}|s^t} \pi_{t+1}(s^{t+1})\lambda_{t+1}^p(s^{t+1})Y_k(s^{t+1}) = 0 \quad (\text{B.7})$$

where λ_t^p is the Lagrange multiplier attached to the aggregate resource constraint.

B.2. Static efficiency

Substituting (B.3) into (B.4)–(B.6) for $\lambda_t(s^t)$ gives the model's static efficiency conditions:

$$-\frac{u_{h^s}(s^t)}{u_c(s^t)} = Y_{h^s}(s^t) \quad (\text{B.8})$$

$$-\frac{u_{h^u}(s^t)}{u_c(s^t)} = Y_{h^u}(s^t) \quad (\text{B.9})$$

$$-\frac{u_{\psi}(s^t)}{u_c(s^t)} = Y_{\psi}(s^t) - \bar{g}_{\psi}(s^t). \quad (\text{B.10})$$

B.3. Intertemporal efficiency

Substituting (B.3) and its one-period lead for $\lambda_t(s^t)$ and $\lambda_{t+1}(s^{t+1})$ into (B.6) gives the model's *intertemporal* efficiency condition:

$$E_t \left\{ \frac{\beta u_c(s^{t+1})}{u_c(s^t) A_t^k} \left[A_{t+1}^k (s_{t+1}) (1 - \delta) + Y_k(s^{t+1}) \right] \right\} = 1. \quad (\text{B.11})$$

B.4. Transformation frontier

The efficiency conditions can be described in terms of appropriately defined marginal rates of substitution and transformation. This permits us to relate efficiency to the general principle that efficient allocations are characterised by the equality of these margins. Following [Arseneau and Chugh \(2012\)](#), we focus on the non-stochastic equivalent of the efficiency conditions, because it allows for the separation of the components of preferences from those of technology. This then allows us to re-express the efficiency conditions (B.8)–(B.11) in terms of MRSs and MRTs. The non-stochastic version of (B.8)–(B.11) is given by

$$MRS_{h_t^s, c_t} \equiv -\frac{u_{h^s}(t)}{u_c(t)} = Y_{h^s}(t) \quad (\text{B.12})$$

$$MRS_{h_t^u, c_t} \equiv -\frac{u_{h^u}(t)}{u_c(t)} = Y_{h^u}(t) \quad (\text{B.13})$$

$$MRS_{\psi_t, c_t} \equiv -\frac{\beta u_{\psi_t}}{u_{c_t}} = Y_{\psi}(s^t) - \bar{g}_{\psi}(s^t) \quad (\text{B.14})$$

$$IMRS_{c_{t+1}, c_t} \equiv \frac{u_c(t)}{\beta u_c(t+1)} = \frac{A_{t+1}^k (1 - \delta) + Y_k(t+1)}{A_t^k}. \quad (\text{B.15})$$

The non-stochastic transformation frontier is defined as³⁴

$$Y(c_t, \psi_t, h_t^s, h_t^u, k_{t+1}; \cdot) \equiv c_t + \bar{g}[\cdot(t)] + A_t^k [k_{t+1} - (1 - \delta)k_t] - Y[\cdot(t)] = 0. \quad (\text{B.16})$$

The function $Y(\cdot)$ jointly describes the two technologies in the model (i.e. goods production and skill creation) and is a function of period t choices. By leading this function one period we obtain

$$Y'(c_t, \psi_{t+1}, h_t^s, h_t^u, k_{t+1}, c_{t+1}, \psi_{t+2}, k_{t+2}, h_{t+1}^s, h_{t+1}^u; \cdot) \equiv c_{t+1} + \bar{g}[\cdot(t+1)] + A_{t+1}^k [k_{t+2} - (1 - \delta)k_{t+1}] - Y[\cdot(t+1)] = 0. \quad (\text{B.17})$$

To capture the dependence between period t and $t+1$ allocations directly, we can view $Y'(\cdot)$ as a generalisation of $Y(\cdot)$ by explicitly allowing for the state variables in period $t+1$ to be determined by choices in period t . We denote as $G(\cdot)$ the representation where period t allocations are incorporated in the period $t+1$ transformation frontier by re-arranging $Y(\cdot)$ for k_{t+1} and substituting into $Y'(\cdot)$:

$$G(c_t, \psi_{t+1}, h_t^s, h_t^u, k_{t+1}, c_{t+1}, \psi_{t+2}, k_{t+2}, h_{t+1}^s, h_{t+1}^u; \cdot) \equiv c_{t+1} + \bar{g}[\cdot(t+1)] + A_{t+1}^k [k_{t+2} - (1 - \delta)k_{t+1}(\cdot)] - Y[\cdot(t+1)] = 0 \quad (\text{B.18})$$

³⁴ As in [Arseneau and Chugh \(2012\)](#) and without loss of generality, in what follows, $g_t^e = 0$ for all t .

where

$$k_{t+1}(\cdot) = \frac{1}{A_t^k} Y[\cdot(t)] + (1 - \delta)k_t - \frac{1}{A_t^k} \bar{g}[\cdot(t)] - \frac{1}{A_t^k} c_t. \tag{B.19}$$

B.5. Static rates of transformation

Differentiating $Y(\cdot)$ with respect to each of its contemporaneous arguments and substituting Y_{c_t} into the remaining condition gives³⁵

$$MRT_{h^s, c_t} \equiv -\frac{Y_{h_t^s}}{Y_{c_t}} = Y_{h^s}(t) \tag{B.20}$$

$$MRT_{h^u, c_t} \equiv -\frac{Y_{h_t^u}}{Y_{c_t}} = Y_{h^u}(t) \tag{B.21}$$

$$MRT_{\psi_t, c_t} \equiv -\frac{Y_{\psi_t}}{Y_{c_t}} = Y_{\psi}(t) - \bar{g}_{\psi}(t). \tag{B.22}$$

Therefore, the conditions for *atemporal* efficiency (B.12)–(B.14) can be expressed in terms of equalisation of MRSs to MRTs.

B.6. Dynamic rates of transformation

To calculate the $IMRT_{c_{t+1}, c_t}$ first note that from (B.19) and $Y(\cdot)$

$$\frac{\partial k_{t+1}}{\partial c_t} = -\frac{1}{A_t^k} = -\frac{Y_{c_t}}{Y_{k_{t+1}}}. \tag{B.23}$$

Differentiating $G(\cdot)$ with respect to c_{t+1} and c_t respectively gives³⁶

$$G_{c_{t+1}} = 1 \tag{B.24}$$

$$G_{c_t} = -A_{t+1}^k (1 - \delta) \frac{\partial k_{t+1}}{\partial c_t} - Y_k(t+1) \frac{\partial k_{t+1}}{\partial c_t} = \frac{A_{t+1}^k (1 - \delta) + Y_k(t+1)}{A_t^k} \tag{B.25}$$

implying

$$IMRT_{c_{t+1}, c_t} \equiv \frac{G_{c_t}}{G_{c_{t+1}}} = \frac{A_{t+1}^k (1 - \delta) + Y_k(t+1)}{A_t^k}. \tag{B.26}$$

Thus we can see that (B.15) implies

$$IMRS_{c_{t+1}, c_t} \equiv \frac{u_c(t)}{\beta u_c(t+1)} = \frac{A_{t+1}^k (1 - \delta) + Y_k(t+1)}{A_t^k} \equiv IMRT_{c_{t+1}, c_t}. \tag{B.27}$$

Therefore, the condition for *intertemporal* efficiency (B.11) can be expressed in terms of MRS and MRT as

$$E_t \left\{ \frac{\beta u_c(s^{t+1}) \left[\frac{A_{t+1}^k (s_{t+1}) (1 - \delta) + Y_k(s^{t+1})}{A_t^k(s_t)} \right]}{u_c(s^t)} \right\} = 1 \Rightarrow E_t \left\{ \frac{IMRT_{c_{t+1}, c_t}}{IMRS_{c_{t+1}, c_t}} \right\} = 1. \tag{B.28}$$

Appendix C. Wedges

This appendix derives the wedges that prevent the first-order conditions in the market economy from satisfying the efficiency conditions obtained in Appendix B. As will be shown, they are due to the presence of distorting taxation as well as the externality in skill creation. To relate the two sets of conditions, note that since we assume the same production technology between the social planner and market economy, we have that

$$Y(h_t^s(s^t), h_t^u(s^t), \psi_t(s^t), k_t(s^{t-1}), s_t) \equiv F(h_t^{s,f} [h_t^s(s^t), \psi_t(s^t)], h_t^{u,f} [h_t^u(s^t), \psi_t(s^t)], k_t^f [k_t(s^{t-1}), s_t]) \tag{C.1}$$

³⁵ Note that since $Y(\cdot)$ is a function of variables dated both t and $t+1$, we explicitly denote the arguments over which the differentiation of $Y(\cdot)$ is carried out.

³⁶ Note that since $G(\cdot)$ is a function of variables dated t , $t+1$ and $t+2$, we explicitly denote the arguments over which the differentiation of $G(\cdot)$ is undertaken.

where $h_t^{sf}(\cdot), h_t^{uf}(\cdot), k_t^f(\cdot)$ are given respectively by the market clearing conditions in (19)–(21). Therefore, we can link $F_{h^{sf}}(s^t), F_{h^{uf}}(s^t)$ and $F_{k^f}(s^t)$ to $Y_{h^s}(s^t), Y_{h^u}(s^t), Y_{\psi}(s^t)$ and $Y_k(s^t)$ as follows:

$$Y_{h^s}(s^t) \equiv \frac{\partial F(\cdot)}{\partial h^s(s^t)} = F_{h^{sf}}(s^t) \frac{\partial h_t^{sf}(s^t)}{\partial h^s(s^t)} = F_{h^{sf}}(s^t) \psi_t(s^t) \tag{C.2}$$

$$Y_{h^u}(s^t) \equiv \frac{\partial F(\cdot)}{\partial h^u(s^t)} = F_{h^{uf}}(s^t) \frac{\partial h_t^{uf}(s^t)}{\partial h^u(s^t)} = F_{h^{uf}}(s^t) [1 - \psi_t(s^t)] \tag{C.3}$$

$$\begin{aligned} Y_{\psi}(s^t) &\equiv \frac{\partial F(\cdot)}{\partial \psi(s^t)} = F_{h^{sf}}(s^t) \frac{\partial h_t^{sf}(s^t)}{\partial \psi(s^t)} + F_{h^{uf}}(s^t) \frac{\partial h_t^{uf}(s^t)}{\partial \psi(s^t)} \\ &= F_{h^{sf}}(s^t) h_t^s(s^t) - F_{h^{uf}}(s^t) h_t^u(s^t) \end{aligned} \tag{C.4}$$

$$Y_k(s^t) \equiv \frac{\partial F(\cdot)}{\partial k(s^{t-1})} = F_{k^f}(s^t) \frac{\partial k_t^f(s^t)}{\partial k(s^{t-1})} = F_{k^f}(s^t). \tag{C.5}$$

C.1. Static wedges

In the decentralised equilibrium with policy distortions, the conditions characterising static trade-offs are given by Eqs. (7), (8) and (23), where the wage rates are given by Eqs. (24) and (25), implying that the static conditions are

$$-\frac{u_{h^s}(s^t)}{u_c(s^t)} = \psi_t(s^t) F_{h^{sf}}(s^t) [1 - \tau_t^s(s^t)] \tag{C.6}$$

$$-\frac{u_{h^u}(s^t)}{u_c(s^t)} = [1 - \psi_t(s^t)] F_{h^{uf}}(s^t) [1 - \tau_t^u(s^t)] \tag{C.7}$$

$$-\frac{u_{\psi}(s^t)}{u_c(s^t)} = h_t^s(s^t) [1 - \tau_t^s(s^t)] F_{h^{sf}}(s^t) - h_t^u(s^t) [1 - \tau_t^u(s^t)] F_{h^{uf}}(s^t) - [1 - \tau_t^a(s^t)] \tilde{g}_{\psi}(s^t). \tag{C.8}$$

Substituting (C.2)–(C.3) into (C.6)–(C.7), we have

$$-\frac{u_{h^s}(s^t)}{u_c(s^t)} = Y_{h^s}(s^t) [1 - \tau_t^s(s^t)] \tag{C.9}$$

$$-\frac{u_{h^u}(s^t)}{u_c(s^t)} = Y_{h^u}(s^t) [1 - \tau_t^u(s^t)]. \tag{C.10}$$

Comparing (C.9)–(C.10) with (B.8)–(B.9) we can see that the skilled and unskilled labour wedges are given by the skilled and unskilled labour taxes respectively. Moreover, note that (C.8) can be re-arranged as

$$\begin{aligned} -\frac{u_{\psi}(s^t)}{u_c(s^t)} &= F_{h^{sf}}(s^t) h_t^s(s^t) - F_{h^{uf}}(s^t) h_t^u(s^t) - \tilde{g}_{\psi}(s^t) + \bar{g}_{\psi}(s^t) - F_{h^{sf}}(s^t) h_t^s(s^t) + F_{h^{uf}}(s^t) h_t^u(s^t) \\ &\quad - \tilde{g}_{\psi}(s^t) + h_t^s(s^t) [1 - \tau_t^s(s^t)] F_{h^{sf}}(s^t) - h_t^u(s^t) [1 - \tau_t^u(s^t)] F_{h^{uf}}(s^t) + \tau_t^a(s^t) \tilde{g}_{\psi}(s^t) \end{aligned} \tag{C.11}$$

which implies using (C.4) that

$$-\frac{u_{\psi}(s^t)}{u_c(s^t)} = Y_{\psi}(s^t) - \bar{g}_{\psi}(s^t) + \tilde{g}_{\psi}(s^t) - \tau_t^a(s^t) \tilde{g}_{\psi}(s^t) - \tau_t^s(s^t) F_{h^{sf}}(s^t) h_t^s(s^t) + \tau_t^u(s^t) F_{h^{uf}}(s^t) h_t^u(s^t)$$

where we can define the wedge in relative skill supply as the difference between the efficient and market margins of adjustment to be

$$\psi^w = \tilde{g}_{\psi}(s^t) - \bar{g}_{\psi}(s^t) - \tau_t^a(s^t) \tilde{g}_{\psi}(s^t) + \tau_t^s(s^t) F_{h^{sf}}(s^t) h_t^s(s^t) - \tau_t^u(s^t) F_{h^{uf}}(s^t) h_t^u(s^t). \tag{C.12}$$

In this case, setting the tax rates to zero will not eliminate the wedge if there are externalities, since $\tilde{g}_{\psi}(s^t) \neq \bar{g}_{\psi}(s^t)$.

C.2. Dynamic wedge

In the decentralised equilibrium with policy distortions, the condition characterising dynamic trade-offs is given by (11), where the interest rate is given by (26), implying that the dynamic condition for capital is

$$E_t \left\{ \frac{\beta u_c(s^{t+1})}{u_c(s^t) A_t^k(s^t)} \left[A_{t+1}^k(s_{t+1})(1 - \delta) + (1 - \tau_{t+1}^k(s^{t+1})) F_{k^f}(s^{t+1}) \right] \right\} = 1. \tag{C.13}$$

We next substitute $F_{k^f}(s^{t+1})$ into (C.13), using (C.5) yielding

$$E_t \left\{ \frac{\beta u_c(s^{t+1})}{u_c(s^t) A_t^k(s^t)} \left[A_{t+1}^k(s_{t+1})(1-\delta) + (1-\tau_{t+1}^k(s^{t+1})) Y_k(s^{t+1}) \right] \right\} = 1. \tag{C.14}$$

Comparing (C.14) with (B.11) we can see that the asset market wedge is equal to the capital tax rate.

Appendix D. Deterministic Ramsey system

In a non-stochastic environment, the first-order conditions derived in (42)–(50) of the main text become

- for $t=0$:

$$V_{h^s}(0) = -[V_c(0) - \Phi A_c] Y_{h^s}(0) + \Phi A_{h^s} \tag{D.1}$$

$$V_{h^u}(0) = -[V_c(0) - \Phi A_c] Y_{h^u}(0) + \Phi A_{h^u} \tag{D.2}$$

$$V_{\psi}(0) = [V_c(0) - \Phi A_c] [\bar{g}_{\psi}(0) - Y_{\psi}(0)] + \Phi A_{\psi} \tag{D.3}$$

$$V_c(0) A_0^k = \beta V_c(1) [Y_k(1) + (1-\delta) A_1^k] + \Phi A_c \tag{D.4}$$

$$Y[\cdot(0)] = c_0 + g_0^e + \bar{g}(\cdot(0)) + (k_1 - (1-\delta)k_0) A_0^k \tag{D.5}$$

- for $t = 1, 2, 3 \dots T-1$:

$$V_{h^s}(t) = -V_c(t) Y_{h^s}(t) \tag{D.6}$$

$$V_{h^u}(t) = -V_c(t) Y_{h^u}(t) \tag{D.7}$$

$$V_{\psi}(t) = V_c(t) [\bar{g}_{\psi}(t) - Y_{\psi}(t)] \tag{D.8}$$

$$V_c(t) A_t^k = \beta V_c(t+1) [Y_k(t+1) + (1-\delta) A_{t+1}^k] \tag{D.9}$$

$$Y[\cdot(t)] = c_t + g_t^e + \bar{g}(\cdot(t)) + (k_{t+1} - (1-\delta)k_t) A_t^k \tag{D.10}$$

- for $t = T$:

$$V_{h^s}(T) = -V_c(T) Y_{h^s}(T) \tag{D.11}$$

$$V_{h^u}(T) = -V_c(T) Y_{h^u}(T) \tag{D.12}$$

$$V_{\psi}(T) = V_c(T) [\bar{g}_{\psi}(T) - Y_{\psi}(T)] \tag{D.13}$$

$$A_T^k = \beta [Y_k(T) + (1-\delta) A_T^k] \tag{D.14}$$

$$Y[\cdot(T)] = c_T + g_T^e + \bar{g}(\cdot(T)) - \delta k_T A_T^k \tag{D.15}$$

- lifetime implementability constraint:

$$\sum_{t=0}^T \beta^t [u_c(t) c_t + u_{h^s}(t) h_t^s + u_{h^u}(t) h_t^u + u_c(t) \Omega(t) \times \bar{g}[\cdot(t)]] - A = 0 \tag{D.16}$$

where $A = u_c(0) \{ b_0 + [(1-\tau_0^k) Y_k(0) + (1-\delta) A_0^k] k_0 \}$ and $\Omega(t) = \{ -(u_{h^s}(t)/u_c(t) \psi_t) h_t^s + (u_{h^u}(t)/u_c(t) (1-\psi_t)) h_t^u + u_{\psi}(t)/u_c(t) \} / \bar{g}_{\psi}(t)$; the Lagrange multiplier θ_t has been replaced with $V_c(t)$ using (42) and the notation $X(t)$ denotes the time period t quantity of X .

Appendix E. ex ante capital tax

First, assume that the government uses a capital tax that is not state-contingent, so that its value for period $t+1$ is decided using the history s^t . Second, define this uncontingent tax as $\tau_{t+1}^{k,a}(s^t)$ and note that it needs to satisfy the Euler-

equation from (11) in the main text, so that the Ramsey allocations are preserved:

$$u_c(s^t)A_t^k(s_t) = \beta E_t \left\{ u_c(s^{t+1}) \left[\tilde{F}_k(s^{t+1}) \left[1 - \tau_{t+1}^k(s^{t+1}) \right] + (1 - \delta)A_{t+1}^k(s_{t+1}) \right] \right\} \quad (\text{E.1})$$

where we have used $\tilde{F}_k(s^{t+1}) = r_{t+1}(s^{t+1})$. Hence, $\tau_{t+1}^{k,a}(s^t)$ needs to satisfy:

$$u_c(s^t)A_t^k(s_t) = \beta E_t \left\{ u_c(s^{t+1}) \left[\tilde{F}_k(s^{t+1}) \left[1 - \tau_{t+1}^{k,a}(s^t) \right] + (1 - \delta)A_{t+1}^k(s_{t+1}) \right] \right\}. \quad (\text{E.2})$$

By comparing (E.2) with (E.1), we see that $\tau_{t+1}^{k,a}(s^t)$ needs to satisfy:

$$E_t \left\{ u_c(s^{t+1}) \left[\tilde{F}_k(s^{t+1}) \left[1 - \tau_{t+1}^{k,a}(s^t) \right] + (1 - \delta)A_{t+1}^k(s_{t+1}) \right] \right\} = E_t \left\{ u_c(s^{t+1}) \left[\tilde{F}_k(s^{t+1}) \left[1 - \tau_{t+1}^k(s^{t+1}) \right] + (1 - \delta)A_{t+1}^k(s_{t+1}) \right] \right\} \quad (\text{E.3})$$

implying that

$$\tau_{t+1}^{k,a}(s^t) = \frac{E_t u_c(s^{t+1}) \left[\tau_{t+1}^k(s^{t+1}) \tilde{F}_k(s^{t+1}) \right]}{E_t u_c(s^{t+1}) \tilde{F}_k(s^{t+1})}. \quad (\text{E.4})$$

This gives $\tau_{t+1}^{k,a}(s^t)$ the *ex ante* capital tax interpretation, since, by multiplying both numerator and denominator in (E.4) by $k_{t+1}(s^t)$, this expression provides the expected tax revenue from capital income as share of the expected capital income, where the expectation is calculated using information at period t .

To obtain the *ex ante* rate stated in Eq. (51) of the main text, we first expand the Euler-equation (E.1):

$$u_c(s^t)A_t^k(s_t) = \beta E_t u_c(s^{t+1}) \tilde{F}_k(s^{t+1}) - \beta E_t u_c(s^{t+1}) \tau_{t+1}^k(s^{t+1}) \tilde{F}_k(s^{t+1}) + \beta E_t u_c(s^{t+1}) (1 - \delta) A_{t+1}^k(s_{t+1}) \quad (\text{E.5})$$

and note that $E_t u_c(s^{t+1}) \tau_{t+1}^k(s^{t+1}) \tilde{F}_k(s^{t+1})$ in (E.5) equals $\tau_{t+1}^{k,a}(s^t) E_t u_c(s^{t+1}) \times \tilde{F}_k(s^{t+1})$, using (E.4). Substituting this expression back into (E.5) we obtain

$$u_c(s^t)A_t^k(s_t) = \beta E_t u_c(s^{t+1}) \tilde{F}_k(s^{t+1}) - \beta \tau_{t+1}^{k,a}(s^t) E_t u_c(s^{t+1}) \tilde{F}_k(s^{t+1}) + \beta E_t u_c(s^{t+1}) (1 - \delta) A_{t+1}^k(s_{t+1}). \quad (\text{E.6})$$

Finally solving (E.6) for $\tau_{t+1}^{k,a}(s^t)$ gives the *ex ante* capital tax rate reported in Eq. (51) of the main text.

Appendix F. Uncontingent debt

F.1. Ex-post capital tax

The treatment of state-uncontingent debt and presentation follows Chari et al. (1994) and Ljungqvist and Sargent (2012, ch. 16). Assume that the government issues uncontingent debt, $b_{t+1}(s^t)$ which has a risk-free return $R_t(s^t)$. The budget constraint of the government in period t is written as

$$g_t^s(s^t) = \tau_t^s(s^t) w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + \tau_t^u(s^t) w_t^u(s^t) [1 - \psi_t(s^t)] \\ \times h_t^u(s^t) - \tau_t^a(s^t) \bar{g}(\cdot)(t) + \tau_t^k(s^t) r_t(s^t) k_t(s^{t-1}) + \frac{b_{t+1}(s^t)}{R_t(s^t)} - b_t(s^{t-1}) \quad (\text{F.1})$$

The budget constraint of the household in period t is given by

$$c_t(s^t) + k_{t+1}(s^t) A_t^k(s_t) + \frac{b_{t+1}(s^t)}{R_t(s^t)} + [1 - \tau_t^a(s^t)] \bar{g}(\cdot)(t) = (1 - \tau_t^s(s^t)) w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + [1 - \tau_t^u(s^t)] w_t^u(s^t) \\ \times [1 - \psi_t(s^t)] h_t^u(s^t) + (1 - \delta) k_t(s^{t-1}) A_t^k(s_t) + [1 - \tau_t^k(s^t)] \times r_t(s^t) k_t(s^{t-1}) + b_t(s^{t-1}) \quad (\text{F.2})$$

which implies that the first-order condition with respect to holding bonds is

$$\frac{1}{R_t(s^t)} = \beta E_t \frac{u_c(s^{t+1})}{u_c(s^t)} \quad (\text{F.3})$$

In turn, this implies that the risk-free (or uncontingent) return needs to satisfy

$$\frac{1}{R_t(s^t)} = \sum_{s^{t+1}|s^t} p_t(s_{t+1}|s^t). \quad (\text{F.4})$$

To obtain an expression for $b_{r+1}(s^r)$ for a given period r , we work as follows. We multiply the budget constraint of the household in (F.2) for periods r and $r+1$ by $\pi_r(s^r)$ and $\pi_{r+1}(s^{r+1})$ respectively, sum the resulting budget constraint in $r+1$ over all possible realisations s_{r+1} and add it to the budget constraint in period r . We then use the first-order conditions of the household to simplify the expression and continue this forward iterative process until time period $T \rightarrow \infty$. By imposing the appropriate transversality conditions we obtain an expression for $b_{r+1}(s^r)$ as a function of identified equilibrium paths

given in

$$b_{r+1}(s^r) = R_r(s^r) \sum_{t=r+1}^{\infty} \sum_{s^t} (\beta^{t-r} \pi_t(s^t) [u_c(s^t) c_t(s^t) + u_{h^s}(s^t) h_t^s(s^t) + u_{h^u}(s^t) h_t^u(s^t) + u_c(s^t) \Omega(s^t) \times \bar{g}(\cdot(t))] / [\pi_r(s^r) u_c(s^r)] - R_r(s^r) k_{r+1}(s^r). \quad (\text{F.5})$$

Hence we can use (F.3) to obtain $R_t(s^t)$, (F.5) to find $b_{t+1}(s^t)$ and finally (F.1) to calculate the *ex-post* capital tax reported in Eq. (52) of the main text.

F.2. Private assets tax

Assume that the government issues uncontingent debt, $b_{t+1}(s^t)$, which has a risk-free return $R_t(s^t)$, satisfying (F.4), but which is taxed using a state-contingent tax $v_{t+1}(s^{t+1})$. The budget constraint of the government is now written as

$$g_t^e(s_t) = \tau_t^s(s^t) w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + \tau_t^u(s^t) w_t^u(s^t) \times [1 - \psi_t(s^t)] h_t^u(s^t) - \tau_t^a(s^t) \bar{g}(\cdot(t)) + \tau_t^k(s^t) \times r_t(s^t) k_t(s^{t-1}) + \frac{b_{t+1}(s^t)}{R_t(s^t)} - [1 - v_t(s^t)] b_t(s^{t-1}) \quad (\text{F.6})$$

while the budget constraint of the household becomes

$$c_t(s^t) + k_{t+1}(s^t) A_t^k(s_t) + \frac{b_{t+1}(s^t)}{R_t(s^t)} + [1 - \tau_t^a(s^t)] \bar{g}(\cdot(t)) = [1 - \tau_t^s(s^t)] w_t^s(s^t) \psi_t(s^t) h_t^s(s^t) + [1 - \tau_t^u(s^t)] w_t^u(s^t) \times [1 - \psi_t(s^t)] h_t^u(s^t) + (1 - \delta) k_t(s^{t-1}) A_t^k(s_t) + [1 - \tau_t^k(s^t)] \times r_t(s^t) k_t(s^{t-1}) + [1 - v_t(s^t)] b_t(s^{t-1}) \quad (\text{F.7})$$

which implies that the first-order condition with respect to holding bonds is

$$\frac{1}{R_t(s^t)} = \sum_{s^{t+1} | s^t} \beta \pi_{t+1}(s^{t+1} | s^t) \frac{u_c(s^{t+1})}{u_c(s^t)} [1 - v_{t+1}(s^{t+1})]. \quad (\text{F.8})$$

This implies that the asset tax must satisfy

$$E_t u_c(s^{t+1}) v_{t+1}(s^{t+1}) = 0 \quad (\text{F.9})$$

which implies that at time period t , the expected value of the asset tax in period $t+1$, valued in terms of utility, has to be equal to zero. Therefore, (F.9) implies that $R_t(s^t)$ in this case is given by (F.3) as well. Moreover, to obtain (F.5), we substitute the household budget constraints in (F.7) forward, using the household first-order conditions, the transversality conditions, the restriction in (F.9) and the restriction that the asset tax in the initial period under consideration is zero. Note that this restriction is equivalent to making the zero capital tax assumption in the initial period.

The private assets tax is defined as the tax revenue from assets over income from assets. In particular:

$$\xi_t(s^{t+1} | s^t) = \frac{\tau_{t+1}^k(s^{t+1}) \bar{F}_k(s^{t+1}) k_{t+1}(s^t) + v_{t+1}(s^{t+1}) b_{t+1}(s^t)}{\bar{F}_k(s^{t+1}) k_{t+1}(s^t) + b_{t+1}(s^t)}. \quad (\text{F.10})$$

Solving (F.6) for $v_t(s^t) b_t(s^{t-1})$ and substituting this into (F.10) we have the expression for $\xi_t(s^{t+1} | s^t)$ reported in Eq. (53) of the main text.

References

- Acemoglu, D., Autor, D., 2011. Skills, tasks and technologies: implications for employment and earnings. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Elsevier, Amsterdam, pp. 1043–1171.
- Adjemian, S., Bastani, H., Karamé, F., Juillard, M., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M., Villemot, S., 2011. Dynare: Reference Manual Version 4. Dynare Working Papers 1, CEPREMAP, revised Jan 2014.
- Aiyagari, S., Marcet, A., Sargent, T., Seppala, J., 2002. Optimal taxation without state-contingent debt. *J. Polit. Econ.* 110 (6), 1220–1254.
- Angeletos, G.M., 2002. Fiscal policy with non-contingent debt and the optimal maturity structure. *Q. J. Econ.* 117 (3), 1105–1131.
- Angelopoulos, K., Asimakopoulou, S., Malley, J., 2013. The Optimal Distribution of the Tax Burden Over the Business Cycle. CESifo Working Paper series 4468.
- Arseneau, D., Chugh, S., 2012. Tax smoothing in frictional labor markets. *J. Polit. Econ.* 120, 926–985.
- Barro, R.J., 1979. On the determination of the public debt. *J. Polit. Econ.* 87 (6), 940–971.
- Buera, F., Nicolini, J.P., 2004. Optimal maturity of government debt without state contingent bonds. *J. Monet. Econ.* 51 (3), 531–554.
- Chamley, C., 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54 (3), 607–622.
- Chari, V., Lawrence, C., Kehoe, P., 1994. Optimal fiscal policy in a business cycle model. *J. Polit. Econ.* 102 (4), 617–652.
- Farhi, E., 2010. Capital taxation and ownership when markets are incomplete. *J. Polit. Econ.* 118 (5), 908–948.
- Goldin, C., Katz, L., 2008. *The Race between Education and Technology*. The Belknap Press of Harvard University Press.
- Greenwood, J., Hercowitz, Z., Krusell, P., 2000. The role of investment-specific technological change in the business cycle. *Eur. Econ. Rev.* 44 (1), 91–115.
- Hornstein, A., Krusell, P., Violante, G., 2005. The effects of technical change on labor market inequalities. In: Aghion, P., Durlauf, S. (Eds.), *Handbook of Economic Growth*, vol. 1., Elsevier, Amsterdam, pp. 1275–1370.
- Katz, L., Murphy, K., 1992. Changes in relative wages, 1963–1987: supply and demand factors. *Q. J. Econ.* 107, 35–78.
- Krusell, P., Ohanian, L., Rios-Rull, J., Violante, G., 2000. Capital-skill complementarity and inequality: a macroeconomic analysis. *Econometrica* 68, 1029–1053.
- Lindquist, M., 2004. Capital-skill complementarity and inequality over the business cycle. *Rev. Econ. Dyn.* 7, 519–540.
- Ljungqvist, L., Sargent, T., 2012. *Recursive Macroeconomic Theory*, 3rd edition The MIT Press, Cambridge, MA.
- Lucas, R., Stokey, N., 1983. Optimal fiscal and monetary policy in an economy without capital. *J. Monet. Econ.* 12 (1), 55–93.

- Martinez-Mongay, C., 2000. ECFIN's Effective Tax Rates. Properties and Comparisons with Other Tax Indicators. Economic Paper No. 146, (October), Brussels, European Commission.
- Matsuyama, K., 2006. The 2005 Lawrence R. Klein Lecture: emergent class structure. *Int. Econ. Rev.* 47, 327–360.
- Merz, M., 1995. Search in the labor market and the real business cycle. *J. Monet. Econ.* 36 (2), 269–300.
- Pourpourides, P., 2011. Implicit contracts and the cyclicity of the skill-premium. *J. Econ. Dyn. Control* 35, 963–979.
- Schmitt-Grohe, S., Uribe, M., 2003. Closing small open economy models. *J. Int. Econ.* 61, 163–185.
- Stockman, D., 2001. Balanced-budget rules: welfare loss and optimal policies. *Rev. Econ. Dyn.* 4, 438–459.
- Werning, I., 2007. Optimal fiscal policy with redistribution. *Q. J. Econ.* 122 (3), 925–967.
- Zhu, X., 1992. Optimal fiscal policy in a stochastic growth model. *J. Econ. Theory* 58 (2), 259–289.