CFCM

CENTRE FOR FINANCE, CREDIT AND MACROECONOMICS

Working Paper 14/17

The optimal distribution of the tax burden over the business cycle

Konstantinos Angelopoulos, Stylianos Asimakopoulos and James Malley

Produced By:

Centre for Finance, Credit and Macroeconomics
School of Economics
Sir Clive Granger Building
University of Nottingham
University Park
Nottingham
NG7 2RD

Tel: +44(0) 115 951 4763
Fax: +44(0) 115 951 4159
suzanne.robey@nottingham.ac.uk

The optimal distribution of the tax burden over the business cycle*

Konstantinos Angelopoulos
University of Glasgow

Stylianos Asimakopoulos
University of Nottingham

James Malley
University of Glasgow and CESifo

December 30, 2014

Abstract

This paper analyses optimal income taxation over the business cycle for households differentiated by labour skill, income and wealth. A model incorporating capital-skill complementarity in production and differential access to labour and capital markets is developed to capture the cyclical characteristics of the US economy, as well as the empirical observations on wage (skill premium) and wealth inequality. We first find that, under a fully flexible budget, the income taxation burden over the business cycle is spread roughly equally across the high-, middle- and low-income households. However, under a balanced budget restriction, the burden is distributed least favourably to the middle-income and most favourably to the high income households.

Keywords: optimal taxation, business cycle, skill premium, income distribution
JEL Classification: E24, E32, E62
Corresponding author: jim.malley@glasgow.ac.uk

*We would like to thank Marco Bassetto, Andrew Clausen, Fabrice Collard, Richard Dennis, Michael Hatcher, Wei Jiang, Matthew Lindquist, Ioana Moldovan, Charles Nolan, Apostolis Philippopoulos, Fabien Postel-Vinay, participants at the Royal Economic Society 2014 Annual Conference, the University of Kent Workshop on Labor Market, Public Policy and the Business Cycle, and the Universities of Aberdeen and Nottingham for helpful comments and suggestions. We are also grateful for financial support from the ESRC, Grant Nos. RES-062-23-2292 and ES/I902414/1 but the views expressed here are entirely our own.

1 Introduction

There is a considerable literature that aims to characterize the properties of optimal income tax policy over the business cycle in models with and without market imperfections as well as with and without restrictions to the set of policy instruments available to the government (see, e.g. the work reviewed in Mankiw *et al.* (2009)). For example, in a representative-agent model with a frictionless labour market, Chari *et al.* (1994) find that the labour income tax should vary little over the business cycle and remain a-cyclical. Werning (2007) extends this result and shows that the optimal volatility of labour taxes for households of different ability should also be very low. A further extension by Angelopoulos *et al.* (2014), in a representative household business cycle model with state contingent debt, capital-skill complementarity and endogenous skill acquisition find that optimal skilled and unskilled labour tax smoothing holds. In contrast, Stockman (2001) shows that a balanced-budget restriction leads to an increase in the optimal volatility of labour relative to capital taxes in a representative-agent model. Additionally, Arseneau and Chugh (2012) find that under search frictions in the labour market, the optimal labour income tax becomes very volatile and counter-cyclical.

However, the literature has not examined optimal income taxation over the business cycle under limited participation of households in asset and skilled labour markets. This is despite the empirical evidence on increased wage inequality associated with capital-skill complementarities in production[1] and the importance of "hand-to-mouth" consumers for macroeconomic stabilisation policy.[2] Moreover, the implications for the business cycle properties of optimal income taxes under a balanced budget have not yet been examined in this environment. This budgetary restriction is particularly relevant in the post financial crisis political environment which favours limiting the use of debt to respond to fluctuations in most advanced economies. In such a setting, the revenue requirements for governments that are faced with exogenous aggregate shocks need to be financed by unpleasant taxes. Thus, a pertinent question for policymaking becomes how to distribute the tax burden over the business cycle.

In light of the above, we aim to analyse optimal income taxation over the business cycle in a model that captures the long-run and cyclical characteristics of the U.S. economy and the empirical observations on wage (skill premium) and wealth inequality. To this end, we develop a model economy

[1] See, e.g. Hornstein *et al.* (2005) and Acemoglu and Autor (2011) for a review of the literature on wage inequality and the skill premium.

[2] See, e.g. the papers by Campbell and Mankiw (1989), Mankiw (2000) and Galí *et al.* (2007).

characterised by capital-skill complementarity in the production process, a labour market which is fragmented with respect to skill, and capital market frictions that lead to the exclusion of a subset of the population from holding assets.

Our model thus consists of three types of households, representing high, middle and low income groups, as well as two labour markets, i.e. skilled and unskilled labour. We assume that the first type of household provides skilled labour services, capturing the supply of college-educated workers in the labour market. The other two types of households, in contrast, are excluded from the skilled labour market, due to lack of college-level education, and can thus only provide unskilled labour.[3]

Capital market frictions in the form of transaction costs associated with financial intermediation imply that households in our model differ with respect to their participation in the asset markets. Following the contributions of Campbell and Mankiw (1989), Mankiw (2000) and Galí *et al.* (2007), we assume that a subset of the households do not have any savings and thus only earn labour income which is totally consumed. We further assume that these households offer unskilled labour services, so that the three types of households in the economy are defined as, high income skilled agents who own assets and face the lowest transactions costs, middle income unskilled agents who also own assets and low income unskilled agents who do not have access to the asset markets.

In contrast to the representative agent optimal taxation literature of e.g. Chari *et al.* (1994), Stockman (2001) and Arseneau and Chugh (2012), our modeling emphasises the importance of imperfections that differentiate households with respect to wage and asset income. In particular, we assume socio-economic barriers, in the form of family background and neighborhood and other peer effects, that exclude certain socio-economic groups from the skilled labour market. We focus on the skill premium following the literature which highlights the importance of a low relative skill supply, in the form of a slow-down in University enrolment over recent decades, as a main driver of wage inequalities (see e.g. Goldin and Katz (2008)). Compared to the heterogeneous agent literature of optimal taxation assuming unobserved innate productivity differences (see e.g. the work reviewed and analysed in Kocherlakota (2010)), our approach focuses on a quantitative evaluation of the business cycle properties of income taxes for three key socio-economic groups, as opposed to the long-run implications of Ramsey income taxation.

We calibrate a version of the model with exogenous tax policy to the

[3]See, for example, Goldin and Katz (2008) for the importance of insufficient growth in college education in explaining wage inequality.

U.S. quarterly data and find that the model fits the key long-run stylized facts as well as the cyclical properties of the data very well, including the empirical findings that the skill premium is effectively a-cyclical and not volatile.[4] Having established the empirical relevance of the model, we then characterize optimal policy, by allowing a Utilitarian government to choose the income tax rates and debt optimally over the business cycle to maximise aggregate welfare given its revenue requirements.

To decide how to set the three income taxes over the business cycle, the government must evaluate a key trade-off. On one hand, unfavourable income taxes, in the form of higher volatility and counter-cyclicality, imply higher marginal social welfare losses when applied to the lower income/consumption groups. This negative effect is particularly strong for hand-to-mouth households, since asset market exclusions imply that they lack the means to smooth consumption over time. At the same time, the tax revenue gains from counter-cyclical and more volatile taxes also increase with the wealth of the tax base to which they apply. Therefore, the government has an incentive to keep the taxes that apply to households with less wealth smoother.

On the other hand, unfavourable taxes directed at the skilled households propagate more in the economy given that the factor supply choices of these households have a greater effect on capital accumulation. This is because the high income skilled households supply more capital than either of the unskilled households, and because of the complementarity between capital and skill. Thus, the skilled income tax implies higher distortions at the aggregate level than the other income taxes. This creates a strong incentive for the government to smooth the tax for the skilled household. By extension, income taxes to the middle income group of unskilled workers are less distorting than the skilled but more distorting than taxes to the hand-to-mouth workers whose factor supply choices affect the propagation mechanism the least.

The quantitative evaluation of this trade-off in the benchmark case, where the government chooses taxes and debt over the business cycle, leads to optimal income taxes which have similar business cycle properties. More specifically, income taxes that apply to the middle-income households are more volatile than the remaining taxes, but the differences are not significant. All taxes are less volatile than output but exhibit sizeable variation over the business cycle compared to models without wage and asset inequality. Similarly, the output correlations of the taxes are all negative and similar across the household types. Finally debt responds countercyclically and is

[4]See e.g. Lindquist (2004) and Pourpourides (2011) for similar exercises in model evaluation with wage inequality and the skill premium.

more volatile than output. Thus, debt performs its expected shock absorbing role over the cycle.

A balanced budget restriction implies that income taxes must be used more actively over the business cycle to meet the revenue requirements of the government. In this case, the quantitative evaluation of the above trade-off leads to a large increase in the volatility of the middle income group and no change in the volatility of the tax to the skilled. In contrast to the other taxes rates, we further find that the volatility of the tax to hand-to-mouth households depends on the assumed degree of risk aversion. The balanced-budget restriction also leads to substantial increases in the counter-cyclicality of the income taxes applying to the unskilled (low and middle income) households, whereas the cyclicality of the skilled households is not much affected.

Overall, the consequences of a balanced budget restriction in terms of higher volatility and counter-cyclicality of the optimal income taxes are distributed least favourably to the middle-income group and most favourably to the high income group. As a result of the exclusions in asset and labour markets under capital-skill complementarity, the importance that skilled workers enjoy in production and the vulnerability of hand-to-mouth consumers to income fluctuations imply that the middle income group is, on balance, the least distorting tax base over the business cycle.

The rest of the paper is organised as follows. Section 2 sets out the model structure. Sections 3 and 4 describe the cyclical properties of the model under exogenous and optimal fiscal policy respectively. Finally, the conclusions are presented in Section 5.

2 Model

We next develop a business cycle model to capture key features of an economy characterised by limited participation in labour and asset markets. We first consider a fragmented labour market, so that there are separate markets for "skilled" and "unskilled" labour, defined as workers with and without college education. We assume that there are socio-economic barriers that do not allow mobility between the two types of labour.[5] This is motivated by

[5] When looking at longer horizons, it is natural to allow for mobility from unskilled to skilled labour due to human capital investment and university education (see e.g. He (2012) and Angelopoulos *et al.* (2013) for models incorporating the joint determination of the relative skill supply and the skill premium). In such contexts, the microfoundations that lead to socio-economic exclusion and/or social mobility are important for long-run outcomes and transitional dynamics (see e.g. Matsuyama (2006) and Aghion and Howitt

empirical evidence which suggests that in business cycle frequencies the share of college educated population in the data has low volatility and is effectively uncorrelated with output. More specifically, using the data in Acemoglu and Autor (2011), we find that the standard deviation of the cyclical component of the skilled population share, relative to that of output, is 0.27, while its correlation with output is -0.18.[6] These findings suggest that the socio-economic barriers limiting access to education and in turn participation in labour markets are indeed more restrictive in shorter, business cycle horizons.

This environment naturally leads to wage inequality. Following the skill premium literature (see e.g. Hornstein *et al.* (2005), Goldin and Katz (2008) and Acemoglu and Autor (2011) for reviews), we assume that the production process involves skilled and unskilled labour inputs which have different degrees of complementarity with capital. The production technology we employ involves structures and equipment capital and it is assumed that skilled labour complement equipment capital relatively more than unskilled labour.

We also allow for asset holding costs when participating in capital and public debt markets (see e.g. Schmitt-Grohé and Uribe (2003) and Benigno (2009)). These capture the outlays associated with financial intermediation.[7] Given inequalities in asset ownership, and specifically, evidence that suggests higher wealth for skilled relative to unskilled workers, we distinguish these costs between skilled and unskilled households.[8] This is motivated, for instance, by assuming that higher education and professional class imply higher financial literacy, which in turn reduces the reliance on financial intermediation for investing in asset markets (see, e.g., Bernheim and Garrett (2003) for evidence that financial literacy increases asset market participation). This leads to different asset holdings across workers, and in particular, we assume that a subset of the population is excluded from the asset markets (see, e.g. Aghion and Howitt (2009)) for capital market imperfections that may lead to limited market participation and agent heterogeneity). Excluded or hand-to-mouth households cannot smooth consumption and thus consume all their (labour) income (see e.g. Campbell and Mankiw (1989), Mankiw (2000) and

(2009, ch. 6)). Here, focusing on business cycle frequencies, we take the barriers that lead to the split in the labour force to skilled and unskilled workers as given.

[6]This is obtained using annual data for the share of college educated population measured in efficiency units, 1963-2008, from Acemoglu and Autor (2011) and GDP data from the U.S. National Income and Product Accounts (NIPA). The cyclical component of the series is obtained using the HP-filter with a smoothing parameter of 100.

[7]See Philippon (2014) on the importance of financial intermediation over the 20th century.

[8]Data from the 2010 U.S. Census, which will be discussed below in more detail, indicate that the wealth of the population with at least a bachelor degree is two and half times more than those without a bachelor degree.

Galí *et al.* (2007)).

The above assumptions lead to an economy with three types of households: (i) skilled, s, who save and provide skilled labour; (ii) unskilled, u, who save and provide unskilled labour; and (iii) hand-to-mouth, h, who do not save and provide unskilled labour.[9] Given the previous discussion, the composition of the population is assumed to be constant and exogenous. For simplicity, we also assume that the total size of the population, N, is constant. The above implies that $N = N_s + N_u + N_h$, where we define $n_s = N_s/N$, $n_u = N_u/N$, and $n_h = 1 - n_s - n_u$. There are also N identical firms and a government.

2.1 Production and firms

Each firm maximises its profits in perfectly competitive markets, by using labour and capital inputs to produce output, Y_t. The production function follows the specification in Krusell *et al.* (2000) which has been shown to match the behavior of the skill premium in the data.[10] In particular, there are two types of capital used in production, capital in structures and equipment, denoted respectively as $K_t^{f,q}$ and $K_t^{f,e}$ and two types of labour, skilled and unskilled, denoted respectively as $h_{s,t}^f$ and $h_{u,t}^f$. The constant returns to scale (CRS) production function is given by a constant elasticity of substitution (CES) specification:

$$\begin{aligned} Y_t = A_t \left(K_t^{f,q}\right)^{\alpha} \{\lambda \left[\nu \left(K_t^{f,e}\right)^{\rho} + (1-\nu)\left(h_{s,t}^f\right)^{\rho}\right]^{\varphi/\rho} + \\ + (1-\lambda)\left(h_{u,t}^f\right)^{\varphi}\}^{\frac{1-a}{\varphi}} \end{aligned} \tag{1}$$

where $0 < a, \lambda, \nu < 1$; $-\infty < \varphi, \rho < 1$; A_t is total factor productivity; φ, and ρ are the parameters determining the factor elasticities, i.e. $1/(1-\varphi)$ is the elasticity of substitution between equipment capital and unskilled labour and between skilled and unskilled labour, whereas $1/(1-\rho)$ is the elasticity of substitution between equipment capital and skilled labour; and a, λ, ν are the factor share parameters. In this specification, capital-skill complementarity is obtained if $1/(1-\rho) < 1/(1-\varphi)$.

Following the literature, A_t is assumed to follow a stochastic exogenous AR(1) process:

$$\log(A_{t+1}) = (1-\rho_A)\log(A) + \rho_A \log(A_t) + \varepsilon_t^A \tag{2}$$

[9] A similar population decomposition is considered in the analysis of UK policy reforms in Angelopoulos *et al.* (2014).

[10] Recent studies in the dynamic general equilibrium (DGE) literature which employ this specification include, e.g. Lindquist (2004), Pourpourides (2011) and He (2012).

where ε_t^A is independently and identically distributed Gaussian random variable with zero mean and standard deviation given respectively by σ_A.

Taking prices and policy variables as given, firms maximise profits:

$$\Pi_t = Y_t - w_{s,t}h_{s,t}^f - w_{u,t}h_{u,t}^f - r_t^e K_t^{f,e} - r_t^q K_t^{f,q} \tag{3}$$

subject to the technology constraint in (1). In equilibrium, profits are zero. The optimality conditions for the firm are given in Appendix A.

2.2 Households

Households, denoted with the subscript $j = s, u, h$, maximize expected lifetime utility:

$$U_j = E_t \sum_{t=0}^{\infty} \beta^t u(C_{j,t}, l_{j,t}) \tag{4}$$

where E_t is the conditional expectations operator at period t; $0 < \beta < 1$ is a constant discount factor; $C_{j,t}$ and $l_{j,t}$ are private consumption and leisure respectively at period t; and $u(\cdot)$ is the utility function which is increasing, strictly concave and three times continuously differentiable with respect to its arguments:

$$u(C_{j,t}, l_{j,t}) = \frac{\left[\left(\mu C_{j,t}^{\gamma} + (1-\mu)\, l_{j,t}^{\gamma}\right)^{\frac{1}{\gamma}}\right]^{(1-\sigma)}}{1-\sigma} \tag{5}$$

where $\sigma > 1$ is the coefficient of relative risk aversion; $0 < \gamma < 1$, with $\frac{1}{1-\gamma}$ representing the elasticity of substitution of leisure for consumption; and $0 < \mu < 1$ measures the share of consumption in utility.

A household of type j faces the following time constraint:

$$1 = l_{j,t} + h_{j,t} \tag{6}$$

where $h_{j,t}$ is hours worked in period t. Additionally, skilled and unskilled households, $m = s, u$, face the following budget constraint:

$$\begin{aligned} &(1-\tau_{m,t})(w_{m,t}h_{m,t} + r_t^q K_{m,t}^q + r_t^e K_{m,t}^e) + R_t^b B_{m,t} = C_{m,t} + \\ &+ I_{m,t}^q + I_{m,t}^e + B_{m,t+1} + \psi_m \left[\left(K_{m,t}^q\right)^2 + \left(K_{m,t}^e\right)^2 + (B_{m,t})^2\right] \end{aligned} \tag{7}$$

while hand-to-mouth households face the constraint:

$$(1-\tau_{h,t})w_{u,t}h_{h,t} = C_{h,t} \tag{8}$$

where $I^q_{m,t}$ and $I^e_{m,t}$ are investment in structures and equipment capital, $K^q_{m,t}$ and $K^e_{m,t}$ respectively; $B_{m,t}$ is the stock of government bonds; $w_{m,t}$ is the wage rate; r^q_t, r^e_t, and r^b_t are the net returns to holding structures and equipment capital and bonds respectively; $R^b_t \equiv \left(1 + r^b_t\right)$ is the gross return to bonds; and $\psi_m > 0$ measure the holding costs for capital and bonds.

The above budget constraints capture several key differences between the households populating) the model. First, households differ in their labour income, as there are different wage rates for skilled and unskilled households. Second, the households also differ in their capital and bond income, since they face different holding costs. In particular skilled and unskilled households which save face finite holding costs, modelled here as quadratic functions of the capital and debt stocks, following e.g. Persson and Tabellini (1992) and Benigno (2009). These differ between households so that they can be distinguished with respect to their steady-state holdings of wealth. Moreover, the hand-to-mouth households implicitly face holding costs that are infinite, so that they are excluded from the asset markets. Third, for each level of income, as reflected by the household type, there is a different income tax rate.

Finally, for $m = s, u$, the laws of motion for capital are:

$$K^q_{m,t+1} = (1 - \delta^q)K^q_{m,t} + I^q_{m,t} \tag{9}$$

$$K^e_{m,t+1} = (1 - \delta^e)K^e_{m,t} + A^e_t I^e_{m,t} \tag{10}$$

where, $0 \leq \delta^i \leq 1$. The capital equipment evolution equation allows for an exogenous process, A^e_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), and Pourpourides (2011)). The investment-specific technological change, A^e_t, is assumed to follow a stochastic exogenous AR(1) process:

$$\log\left(A^e_{t+1}\right) = (1 - \rho_{A^e})\log\left(A^e\right) + \rho_{A^e}\log\left(A^e_t\right) + \varepsilon^{A_e}_t \tag{11}$$

where $\varepsilon^{A_e}_t$ is independently and identically distributed Gaussian random variable with zero mean and standard deviation given by σ_{A^e}.

An increase in the efficiency level of investment in capital equipment, A^e_t, favours the productivity of skilled workers more than the productivity of unskilled workers. Hence, the model is consistent with the empirical evidence that points to skill-biased technical change and a rising skill premium over the recent decades (see e.g. Katz and Murphy (1992) and Krusell *et al.* (2000); also see Hornstein *et al.* (2005) and Acemoglu and Autor (2011) for reviews).

Each household $m = s, u$ chooses $\{C_{m,t},\ h_{m,t},\ I^q_{m,t},\ I^e_{m,t},\ K^q_{m,t+1},\ K^e_{m,t+1},\ B_{m,t+1}\}^{\infty}_{t=0}$, to maximise (4) subject to (6)-(10), by taking policy variables and prices as given. Similarly, hand-to-mouth households, $j = h$, choose $\{C_{h,t}, h_{h,t}\}^{\infty}_{t=0}$, to maximise (4) subject to (6) and (8), by taking policy variables and prices as given. The optimality conditions for the households are given in Appendix A.

2.3 The government

The government's budget constraint is given by:

$$\begin{aligned} &B_{t+1} + \tau_{s,t} n_s (w_{s,t} h_{st} + r^q_t K^q_{s,t} + r^e_t K^e_{s,t}) + \tau_{u,t} n_u \times \\ &\times (w_{u,t} h_{u,t} + r^q_t K^q_{u,t} + r^e_t K^e_{u,t}) + n_h \tau_{h,t} \left(w_{u,t} h_{h,t}\right) = G^c_t + R^b_t B_t \end{aligned} \tag{12}$$

where, G^c_t is average government expenditure per household and B_t is the stock of government bonds per household. Since we focus on the revenue side of the budget constraint, we assume that G^c_t is wasteful and follows an exogenous AR(1) process. Thus, its fluctuations act as exogenous spending shocks which require a change in the revenue collected (see e.g. Chari *et al.* (1994), Stockman (2001) and Arseneau and Chugh (2012) for a similar approach regarding G^c_t):

$$\log\left(G^c_{t+1}\right) = (1 - \rho_{G^c}) \log\left(G^c\right) + \rho_{G^c} \log\left(G^c_t\right) + \varepsilon^{G^c}_t \tag{13}$$

where $\varepsilon^{G^c}_t \sim iidN(0, \sigma^2_{G^c})$. We consider below policy regimes where the tax rates are exogenously set or they are optimally chosen by the government. When tax rates are exogenous, the residual variable in (12) is B_{t+1}. When we examine optimal policy, the government chooses all tax rates and new debt to maximise aggregate welfare. We also examine the optimal choice of taxes in the case where the government cannot issue debt to balance the budget (see also Stockman (2001), who considers optimal capital and labour taxes with and without access to debt, albeit in a different setup).

2.4 Market clearing conditions

The labour and capital market clearing conditions are given by:

$$h^f_s = n_s h_s \tag{14}$$

$$h^f_u = n_u h_u + n_h h_h \tag{15}$$

$$K^i_t = n_s K^i_{s,t} + n_u K^i_{u,t} \tag{16}$$

$$B_t = n_s B_{s,t} + n_u B_{u,t}. \tag{17}$$

The aggregate resource constraint is:

$$
\begin{aligned}
Y_t = {} & G_t^c + n_s C_{s,t} + n_u C_{u,t} + n_h C_{h,t} + n_s \left(I_{s,t}^q + I_{s,t}^e\right) + \\
& + n_u \left(I_{u,t}^q + I_{u,t}^e\right) + n_s \left\{\psi_s \left[\left(K_{s,t}^q\right)^2 + \left(K_{s,t}^e\right)^2 + (B_{s,t})^2\right]\right\} + \\
& + n_u \left\{\psi_u \left[\left(K_{u,t}^q\right)^2 + \left(K_{u,t}^e\right)^2 + (B_{u,t})^2\right]\right\}.
\end{aligned} \tag{18}
$$

3 Exogenous policy

Prior to studying optimal tax policy, we calibrate the model to examine whether its predictions regarding first and second moments of the endogenous variables are consistent with the data. This analysis is undertaken when the model is driven by exogenous and empirically relevant government spending and income tax rate processes, while government debt is the residual variable in the budget constraint of the government.

3.1 Decentralized competitive equilibrium

Given initial levels of the assets, K_0^q, K_0^e, $B_{s,0}$, $B_{u,0}$, the stationary stochastic processes for four policy instruments $\{\tau_{s,t}, \tau_{u,t}, \tau_{h,t}, G_t^c\}_{t=0}^{\infty}$ and for technology $\{A_t, A_t^e\}_{t=0}^{\infty}$, the DCE system of equations is characterized by a sequence of allocations $\{C_{s,t},\ C_{u,t},\ C_{h,t},\ h_{s,t},\ h_{u,t},\ h_{h,t},\ K_{s,t+1}^q, K_{s,t+1}^e,\ K_{u,t+1}^q, K_{u,t+1}^e,$ $I_{s,t}^q, I_{s,t}^e,\ I_{u,t}^q,\ I_{u,t}^e,\ B_{s,t+1},\ B_{u,t+1}\}_{t=0}^{\infty}$ and prices $\{w_{s,t}, w_{u,t}, r_t^q, r_t^e, r_t^b\}_{t=0}^{\infty}$ such that: (i) households maximize their welfare and firms their profits, taking policy and prices as given; (ii) the government budget constraint is satisfied in each time period and (iii) all markets clear.

3.2 Business cycle statistics

We aim for the exogenous-policy model to replicate the long-run great ratios and key labour and asset market characteristics as well as explaining the cyclical volatilities and correlations with output of key variables in the economy. In Table 1 we report the data volatilities and correlations with output from existing studies for variables which correspond with key endogenous variables in our model. These are taken directly from the results reported in Lindquist (2004) and Pourpourides (2011) who use quarterly data for the period 1979-2002 and 1979-2003 respectively.[11]

[11] To obtain labour supply per skill group at a quarterly frequency, these studies disaggregate the labour force into skilled and unskilled by taking into account the years spent in education (i.e. skilled workers are those with 14 or more years of schooling). This is based on the assumption that college-educated workers are primarily employed in occupa-

As can be seen in Table 1, these studies document some interesting results regarding the labour market statistics. In particular, they point out that the skill premium is effectively uncorrelated with output and smoother than output in business cycle frequencies. Moreover, the cyclical properties of the labour supply of skilled and unskilled workers do not differ qualitatively, both having a positive correlation with output, while being less volatile than output. However, unskilled labour hours seem to have a slightly stronger pro-cyclical relationship with output, compared with skilled labour hours. The statistics regarding consumption and investment are similar to those commonly obtained in other macroeconomic research.

Table 1: Business cycle statistics of main endogenous variables

X_i	Y	C	I	$\frac{w_s}{w_u}$	h_s	h_u
$\widehat{\rho}(X_i, Y)$	1	0.83	0.76-0.91	0.09-0.19	0.69	0.73
$\widehat{\sigma}(X_i)$	0.013-0.014	0.011-0.012	0.037-0.063	0.006-0.013	0.010	0.012

Sources: Data reported in Lindquist (2004) and Pourpourides (2011).

3.3 Calibration

The parameters of the model are calibrated either based directly on data (including existing econometric evidence) or by ensuring that the steady-state and cyclical properties of key endogenous variables are consistent with the data. The calibrated parameters are summarised in Tables 2 and 3.

3.3.1 Population shares

We assume that the population breakdown in our model economy is given as $n_s = 0.4$, $n_u = 0.4$, $n_h = 0.2$.[12] The share of skilled households is roughly consistent with data from the 2010 U.S. Census, which indicates that 43% of the population has a college degree and that the percentage of households

tions that require high skills and have higher returns (see e.g. Acemoglu and Autor (2011) and references therein). Acemoglu and Autor (2011) present annual data for the relative supply of college-educated versus high-school graduates and for the wage premium paid to college educated workers. Although we use the quarterly data for our business cycle analysis, the second moments of the skill premium, using the annual data and the classification in Acemoglu and Autor (2011), gives similar results. In particular, the cyclical relative volatility and correlation of the skill premium with output are respectively given by 0.49 and -0.13.

[12] We will discuss below that a particular advantage of the 40/40/20 percent split in population is that it allows us to approximate the effective income tax rate which applies to each group by using the Piketty and Saez (2007) income tax data per income quintal.

without any assets is 18.7%.[13] It also broadly coheres with the data in Acemoglu and Autor (2011), which implies that the average share of the labour force with a college degree is about 45%. Finally, the split of unskilled households into hand-to-mouth and those who can access the asset market, ties in with empirical evidence from Traum and Yang (2010) and Cogan *et al.* (2010), who estimate the share of the hand-to-mouth population for the U.S. at 18% and 26.5% respectively.

Table 2: Model parameters

Parameter	Value	Definition	Source
$0 \leq \delta^q \leq 1$	0.014	depreciation rate of capital structures	calibration
$0 \leq \delta^e \leq 1$	0.031	depreciation rate of capital equipment	calibration
$0 < \beta < 1$	0.990	time discount factor	calibration
$\frac{1}{1-\gamma}$	1.429	consumption to leisure elasticity	calibration
$0 < \mu < 1$	0.450	weight attached to consumption in utility	calibration
$\sigma > 1$	2.000	coefficient of relative risk aversion	assumption
$0 \leq \alpha \leq 1$	0.130	capital structures share of income	calibration
$\frac{1}{1-\rho}$	0.669	capital equipment to skilled labour elasticity	assumption
$\frac{1}{1-\varphi}$	1.669	capital equipment to unskilled labour elasticity	assumption
$0 < \lambda < 1$	0.557	composite input share of output	calibration
$0 < \nu < 1$	0.581	capital equipment share of composite input	calibration
$0 < \frac{G^c}{Y} < 1$	0.188	government spending	calibration
$0 < \frac{B}{Y} < 1$	2.120	total debt to output ratio	calibration
$\psi_s > 0$	0.0010	asset holding cost for skilled agents	calibration
$\psi_u > 0$	0.0042	asset holding cost for unskilled agents	calibration
τ_s	0.247	average income tax rate, skilled	data
τ_u	0.180	average income tax rate, unskilled	data
τ_h	0.144	average income tax rate, hand-to-mouth	data

3.3.2 Tax-spending policy

We assume that the income tax rates for $j = s, u, h$ follow AR(1) processes:

$$\log(\tau_{j,t+1}) = (1 - \rho_j)\log(\tau_j) + \rho_j \log(\tau_{j,t}) + \varepsilon_t^{\tau_j} \tag{19}$$

where $\varepsilon_t^{\tau_j} \sim Niid(0, \sigma_{\tau_j})$. We obtain income tax data by using the Piketty and Saez (2007) dataset, which reports annual data on income tax rates per income group (in quintals) for the period 1966-2001.[14] We use this data to

[13] This information is obtained from Table 4 of the U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel, Wave 7, updated in July 12, 2013.

[14] These tax rates refer to average tax rates by income groups. To obtain quarterly series from the annual data, we follow the interpolation method in Litterman (1983). We use as

obtain three income tax rates, the first for the lowest quintal, the second as the average for the two middle quintals, and the third as the average for the two top quintals. We then employ these time-series of tax rates as proxies for $\tau_{j,t}$, $j = h, u, s$, respectively, in our model. Therefore, we set the constant terms in the AR processes described above for, τ_j, $j = h, u, s$, to be equal to the data averages for the respective income quintals, i.e. for τ_{0-20}, τ_{20-60} and τ_{60-100}. Moreover, we calculate ρ_j and σ_j by estimating the AR(1) processes in (19), using the cyclical component of the respective tax series in the data. All tax and spending series are logged and HP-filtered with a smoothing parameter of 1600 before estimating the required second moments.

The government spending series is obtained using quarterly data from the BEA for the period 1979 to 2002.[15] Following the same procedure as with the tax rates, we calculate ρ_G and σ_G by estimating the AR(1) processes in (13), using the cyclical component of the public spending series. Quarterly data on total public debt for the period 1979-2002 are obtained from FRED.[16] We calibrate the long-run value of government spending so that the steady-state debt to GDP ratio is (4×0.53). As will be discussed below, this calibration also implies as a by-product that the public spending to output ratio in the model is similar to the average in the data, at about 19%.

3.3.3 Production, asset and labour markets

The elasticities of substitution between skilled labour and capital equipment and between unskilled labour and capital equipment (or skilled labour) have been estimated by Krusell *et al.* (2000). We use their estimates, so that $\varphi = 0.401$ and $\rho = -0.495$. The remaining parameters in the production function are calibrated to ensure that the steady-state predictions of the model in asset and labour markets are consistent with the data (following e.g. Lindquist (2004), He and Liu (2008), Pourpourides (2011) and He (2012)). The income shares λ and ν are calibrated to obtain a skill premium of 1.659 and a labour share of income of about 69%. In particular, the target value for the skill premium is obtained from the dataset of Castro and Coen-Pirani (2008) using quarterly US data for the period 1979-2003.[17] The share of labour income in GDP is obtained from BEA data on personal income for

---

an indicator variable the quarterly time-series of labour income tax rates from Arseneau and Chugh (2012).

[15] This series refers to government consumption expenditures and gross investment as it is reported in NIPA Table 1.1.5.

[16] The series for total public debt refers to GFDEBTN series published by FRED.

[17] Using the data in Acemoglu and Autor (2011) gives a similar value for the average skill premium, namely 1.6.

the period 1970-2011.[18] The productivity of capital structures, α is set at the same value as Lindquist (2004) and helps to bring the model's capital to output ratios close to the data. The calibrated parameters in the production function are generally very similar to those estimated or calibrated in the literature.

The depreciation rates of capital structures and capital equipment are also set to match those in the relevant literature. In particular, we set $\delta^e = 0.031$ and $\delta^q = 0.014$ to be within the range of Lindquist (2004) and Pourpourides (2011).[19]

We set the asset holding cost parameters as $\psi_s = 0.001$ and $\psi_u = 0.0042$. There are two targets for these parameters. The first is that the total asset holdings for skilled households in the deterministic steady-state is 2.5 times higher than for unskilled households. This ensures that the model's steady-state matches data from the US Census[20] which indicates that the wealth of the population with at least a bachelor degree is two and half times more than those without a bachelor degree. The second target is that the model with exogenous policy produces a quarterly capital to output ratio equal to 6.55 in the steady-state. This is consistent with an annual capital to output ratio of about 1.64, obtained using BEA annual data on capital stocks from 1979 to 2011[21]. As will be discussed below, this calibration also implies as a consequence that the output share of asset holding costs is about 7%, consistent with the income share of the finance industry since 1980 (see e.g. Philippon (2014)).

3.3.4 Utility function

The time discount factor, $\beta = 0.99$, is calibrated to match an asset return net of depreciation and taxes equal to 1%. The weight of consumption in utility, $\mu = 0.45$, is set so that in the steady-state the ratio of consumption over output is about 65%.[22] Moreover, $\gamma = 0.30$ so that the households devote about 33% of their time to work. For the base results below, we set

[18] The specific data are obtained from the BEA, Table 2.1, Personal Income and Its Disposition.

[19] For instance, Krusell *et al.* (2000) report $\delta^q = 0.0125$ and $\delta^e = 0.031$; Pourpourides $\delta^q = 0.014$ and $\delta^e = 0.027$; and Lindquist $\delta^q = 0.014$ and $\delta^e = 0.031$.

[20] The specific information is obtained using Table.1 from the U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel, Wave 7, last modified July 12, 2013.

[21] The capital stock is calculated using the following data from BEA: NIPA Table 1.1 (line 3 plus line 21 minus line 7) and Tables 7.1A (line 30) and Table 7.1B (line 38).

[22] The data series used for consumption is obtained from the BEA, NIPA Table 1.1.5 for the period 1979-2011.

the value for the coefficient of relative risk aversion, $\sigma = 2$, equal to the mid-point of the most likely range reported in Gandelman and Hernández-Murillo (2013). We then further investigate the role of a higher σ for optimal policy below. Note that, as is common in optimal taxation analyses (see e.g. Mankiw *et al.* (2009)), we assume common preferences across households. This ensures that differences in optimal income tax rates across households reflect heterogeneity in opportunities and not in preferences.

Table 3: Stochastic processes

Parameter	Value	Definition	Source
σ_A	0.0075	std. dev. of TFP	calibration
ρ_A	0.9500	AR(1) coef. of TFP	assumption
σ_{A^e}	0.0073	std. dev. of inv.-specific tech. change	calibration
ρ_{A^e}	0.8000	AR(1) coef. of inv.-specific tech. change	calibration
σ_{τ_s}	0.0019	std. dev. of income tax, skilled	data
ρ_{τ_s}	0.8700	AR(1) coef. of income tax, skilled	data
σ_{τ_u}	0.0021	std. dev. of income tax, unskilled	data
ρ_{τ_u}	0.9000	AR(1) coef. of income tax, unskilled	data
σ_{τ_h}	0.0019	std. dev. of income tax, hand-to-mouth	data
ρ_{τ_h}	0.9400	AR(1) coef. of income tax, hand-to-mouth	data
σ_{G^c}	0.0090	std. dev. of public spending	data
ρ_{G^c}	0.7700	AR(1) coef. of public spending	data

3.3.5 Technology

The constant terms in the processes for TFP and investment-specific technological change are normalized to unity (i.e. $A = 1$ and $A^e = 1$ respectively). We calibrate the autocorrelation and standard deviation parameters for the process of investment-specific technological change (ρ_{A^e} and σ_{A^e}) to match the correlation of investment with output and the standard deviation of investment in the data as presented in Table 1. The autocorrelation parameter of TFP is set equal to 0.95, following Lindquist (2004) and Pourpourides (2011), while σ^A is calibrated to match the volatility of output observed in the data (see Table 1).

3.4 Solution and results

The steady-state solution of the DCE system for key variables is compared with their corresponding data averages in Table 4. To study dynamics, we compute a second-order approximation of the equilibrium conditions around the deterministic steady-state, by implementing the perturbation methods in Schmitt-Grohé and Uribe (2003). We simulate time paths under shocks to total factor productivity, investment-specific technological change, government spending, and income tax realizations, that are obtained from the distributions specified above. We conduct 10,000 simulations of 96 periods to match the number of observations in the data used by the studies in Table 1, initialised from the steady-state in Table 4. For each simulation, we HP-filter the logged series and then compute the required moments and report the means of these moments across the simulations in Table 5.

Table 4: Steady-state of the exogenous policy model

Variable	Model	Data	Variable	Model	Data
$\frac{K}{Y}$	6.550	6.550	h_s	0.332	0.317
$\frac{I}{Y}$	0.144	0.159	h_u	0.306	-
$\frac{C}{Y}$	0.651	0.659	h_h	0.349	-
$\frac{G^c}{Y}$	0.188	0.195	$\widetilde{h}_u$	0.320	0.348
r_{net}	0.010	0.010	$\frac{w_s}{w_u}$	1.659	1.659
$\frac{b}{Y}$	2.120	2.120	$\frac{wh}{Y}$	0.690	0.686

The results in Tables 4 and 5 suggest that the predictions of the model with respect to both the steady-state and business cycle properties of the series cohere well with the data.[23] Starting with the steady-state, Table 4 shows that all model predictions are quantitatively similar to the long-run averages in the data, including several dimensions that were not explicit targets of the calibration.

Regarding the labour markets, the utility function was calibrated so that households work on average about one third of their available time. The model predicts similar work hours for skilled and unskilled workers. Note that $\widetilde{h}_u$ is average work hours by households providing unskilled labour in our model.[24] Table 4 also shows that the labour's share of income in the model, $\frac{w_s n_s \bar{h_s} + w_u n_u h_u + w_u n_h h_h}{Y} = 0.69$ is very close to the value (i.e. 0.686) obtained from the BEA Table 2.1 for 1979-2002.

The work-time allocations also imply Frisch (or λ-constant) labour supply elasticities of 1.43 for skilled, 1.51 for unskilled and 1.24 for hand-to-mouth

[23]Note that the data sources for the series in Table 4 include: (i) BEA, NIPA Table 1.1.5 for output, investment and consumption; (ii) BEA, NIPA Table 1.1 (line 3 plus line 21 minus line 7) and Tables 7.1A (line 30) and Table 7.1B (line 38) for the capital stock; (iii) BLS, Current Employment Statistics survey for hours worked; (iv) World Bank for the real rate of return; (v) BEA, NIPA Table 2.1 for labour's share in income; and (vi) U.S. data from Castro and Coen-Pirani (2008) for the skill premium. Comparable averages are obtained using the dataset in Lindquist (2004), for those variables that are similar in both studies.

[24]The specific table used from the BLS is the Employment, Hours, and Earnings from the Current Employment Statistics survey (National).

workers, which are generally consistent with the literature (see e.g. Chetty *et al.* (2011), and Keane and Rogerson (2012)). These Frisch elasticities cohere with two different sources of empirical evidence regarding labour supply decisions of heterogeneous workers. In particular, Domeij and Flodden (2006) find that credit-constrained households have a more inelastic labour supply, while Blau and Khan (2007) and Kimball and Shapiro (2008) find that workers with a college degree have lower labour supply elasticity compared with workers with some or no college education. Note that the differences between the labour supply elasticities of the different types of households predicted by the model are not due to assumed differences in preferences. Instead, they are driven by the different opportunities that the households face in the labour and asset markets as well as the production structure of the economy.

Regarding the asset markets, while the transaction costs were calibrated to match wealth inequality and capital as a share of GDP, the model also predicts that asset holding costs as a share of output are about 7%. Note that this is consistent with the share of the financial sector in GDP since 1980 (see e.g. Philippon (2014)). Finally, regarding the government budget, while G^c was calibrated to match $\frac{B}{Y}$, the model also predicts that $\frac{G^c}{Y}$ is consistent with the data.

Turning to the business cycle statistics in Table 5, the overall fit is comparable to the data reported in Table 1 and to the results from existing research on business cycle models with wage inequality under capital-skill complementarity (see e.g. Lindquist (2004) and Pourpourides (2011)). The model matches the key stylised facts regarding the skill premium in the data, i.e. it is effectively not correlated with output and its volatility is less than that of output. In addition, the model predictions regarding the second moments of hours worked are generally consistent with the data both qualitatively and quantitatively. In particular, the model predicts a higher correlation of unskilled hours with output, compared with the correlation between skilled hours and output. Overall, we conclude that the model's predictions regarding the key endogenous variables are empirically relevant.

Table 5: Business cycle statistics of main endogenous variables (model)

X_i	Y	C	I	$\frac{w_s}{w_u}$	h_s	h_u	h_h
$\widehat{\rho}(X_i, Y)$	1	0.95	0.95	0.19	0.83	0.95	0.96
$\widehat{\sigma}(X_i)$	0.014	0.009	0.057	0.003	0.008	0.008	0.002

4 Optimal distribution of tax burden

Having established the empirical relevance of the calibration, we now discard the exogenous processes for the income tax rates in (19) and instead assume that the paths of these tax rates are optimally chosen by a government that seeks to maximise a Utilitarian objective function under commitment. However, government consumption spending continues to follow the exogenous process in (13), as commonly assumed in the analysis of optimal taxation.

4.1 The problem of the government

The government chooses the paths of the three income tax rates and of state-uncontingent bonds to maximise aggregate welfare subject to the optimality conditions of households and firms and the appropriate feasibility and government budget constraints. The literature has examined the importance of restrictions on government debt, either directly, in the form of bounds on debt or constraints that fix its quantity over the business cycle, or via frictions in asset markets (see e.g. Aiyagari *et al.* (2000), Stockman (2001), Fahri (2006)). These restrictions are motivated by a political-economic reality where imperfections in asset markets and institutional debt limits restrict the use of this instrument to smooth economic fluctuations. In our model, we first examine the case where asset market imperfections imply that the government can issue state-uncontingent bonds subject to the frictions created by the asset holding costs discussed above. In addition to this benchmark case, we examine in Section 4.4 the importance of a constraint that implies that the government cannot issue debt over the business cycle, by comparing optimal income taxes with and without this constraint.

Our interest is on income taxation at the household level, which can be progressive or regressive, rather than on optimal factor input taxation as in e.g. the representative agent models in the standard Ramsey analysis of capital and labour taxes.[25] The assumptions regarding the policy instruments imply that we examine an optimal taxation problem with an incomplete set of tax instruments, since tax rates are not available for all pairs of goods in the economy (see e.g. Chari and Kehoe (1999), for the definition of a complete tax system).[26]

[25] Income taxes in practice apply uniformly to the income that the household accrues due to employment, self-employment, business and rental activities. Moreover, dividend and savings taxes are also effectively linked to the income of the household. Although the reality of the tax systems is significantly more complicated, due to tax exemptions and other (potentially targeted) tax-transfer schemes, we assume that the households pay a single income tax for the sum of their income other than payments from government bonds.

[26] In Angelopoulos *et al.* (2015), we examine second-best policy for a Ramsey planner that has access to a complete set of tax instruments and state-contingent debt, under capital-skill complementarity. However, that analysis focuses on a representative household, so that we do not address issues pertaining to income inequality between heteroge-

Our aim is to identify the optimal distribution of the income tax burden over the business cycle in a setup with realistic income inequality and to evaluate the implications of flexible versus fixed government budgets. Since the government can choose different tax rates for each income group, it has to decide whether these tax rates should have the same volatility and co–movement with output over the business cycle and, if not, how to set these cyclical properties for each tax rate. We examine the problem of a government that has Utilitarian preferences, so that its objective function is given by the expected lifetime utility of the weighted average of the welfare of the three types of households, where the weights attached to each type are equal to the population share of that type, n_j (see e.g. Mankiw *et al.* (2009) on the common use of Utilitarian social preferences for optimal taxation analyses). Under a flexible public budget, the government chooses $\{\tau_{s,t}, \tau_{u,t}, \tau_{h,t}, C_{s,t}, C_{u,t}, C_{h,t}, h_{s,t}, h_{u,t}, h_{h,t}, K^q_{s,t+1}, K^e_{s,t+1}, K^q_{u,t+1}, K^e_{u,t+1}, I^q_{s,t}, I^e_{s,t}, I^q_{u,t}, I^e_{u,t}, B_{s,t+1}, B_{u,t+1}, w_{s,t}, w_{u,t}, r^q_t, r^e_t, r^b_t\}^\infty_{t=0}$, to maximise:

$$U^g = E_t \sum\nolimits_{j=s,u,h} \left[n_j \sum_{t=0}^{\infty} \beta^t u(C_{j,t}, l_{j,t}) \right] \tag{20}$$

subject to the budget constraints (7), (8), (12), market clearing (14)-(17), and optimality conditions (A1)-(A13) in Appendix A. We assume that the government can commit to the optimal paths. To preserve space, we only discuss the first-order conditions from the government's problem for the policy instruments (income taxes and bonds). The equilibrium conditions for the policy instruments encapsulate key incentives that the Utilitarian government needs to consider when setting taxes and debt optimally over the business cycle. These are summarised here and then further discussed in the following sub-sections, when analysing how the government resolves the trade-off implied by setting its policy instruments. In particular, as we shall see, differences between optimal income tax rates arise from a quantitative evaluation of the costs and benefits associated with income taxes and a comparison of those costs and benefits *between* the taxes.

4.1.1 Optimality conditions for taxes

The first-order condition for $\tau_{s,t}$ is given by:

$$\begin{aligned} &\tfrac{\partial U^g}{\partial \tau_{s,t}} = \lambda^{gbc}_t \left[n_s \left(K^e_{s,t} r^e_t + K^q_{s,t} r^q_t \right) + n_s h_{s,t} w_{s,t} \right] - \\ &-\lambda^{bcs}_t \left[K^e_{s,t} r^e_t + K^q_{s,t} r^q_t + h_{s,t} w_{s,t} \right] - \lambda^{lss}_t U_{c_s}(t) \, w_{s,t} - \\ &-\lambda^{kqs}_{t-1} U_{c_s}(t) \, r^q_t - \lambda^{kes}_{t-1} U_{c_s}(t) \, r^e_t = 0 \end{aligned} \tag{21}$$

neous households and differences in income taxes.

where λ_t^{lss}, λ_{t-1}^{kqs}, λ_{t-1}^{kes}, λ_t^{bcs} and λ_t^{gbc} refer to the Lagrange multipliers attached to the skilled household's FOCs for hours worked, capital structures, capital equipment, the household's budget constraint and the government budget constraint respectively.[27] The first-order condition for the choice of $\tau_{s,t}$ dictates that this should be chosen so that the net marginal social benefits equals the marginal social costs. The net marginal benefits, MB^s, arising from an increase in the tax burden to skilled households are given by: $\lambda_t^{gbc}\left[n_s\left(K_{s,t}^e r_t^e + K_{s,t}^q r_t^q\right) + n_s h_{s,t} w_{s,t}\right] - \lambda_t^{bcs}\left[K_{s,t}^e r_t^e + K_{s,t}^q r_t^q + h_{s,t} w_{s,t}\right]$. The first term in MB^s captures the increase in the tax revenue collected by the government. This works to increase social welfare by leading to a reduction in the total revenue requirements of the government and thus the need to use the remaining distorting policy instruments. The increase in tax revenue is transformed into units of social welfare via the marginal social valuation of an additional unit of tax revenue, i.e. λ_t^{gbc}. The second term in MB^s captures the decrease in the skilled household's income, which works to reduce household and thus social welfare. This is again transformed into social welfare terms by the marginal social valuation of relaxing the skilled household's budget constraint, i.e. λ_t^{bcs}. The marginal costs, MC^s, in turn are given by: $\lambda_t^{lss} U_{c_s}(t) w_{s,t} + \lambda_{t-1}^{kqs} U_{c_s}(t) r_t^q + \lambda_{t-1}^{kes} U_{c_s}(t) r_t^e$. These capture the distortions in labour supply and investment in both types of capital, associated with the fall in the marginal net returns to these three activities. These are first transformed into household utility terms by U_{c_s} and second into units of social welfare by the relevant shadow social price of the distortion, i.e. λ_t^{lss}, λ_{t-1}^{kqs} and λ_{t-1}^{kes}.

The first-order conditions for the remaining two income tax rates are given by:

$$\begin{aligned}
&\tfrac{\partial U^g}{\partial \tau_{u,t}} = \lambda_t^{gbc}\left[n_u\left(K_{u,t}^e r_t^e + K_{u,t}^q r_t^q\right) + n_u h_{u,t} w_{u,t}\right] - \\
&-\lambda_t^{bcu}\left[K_{u,t}^e r_t^e + K_{u,t}^q r_t^q + h_{u,t} w_{u,t}\right] - \lambda_t^{lsu} U_{c_u}(t) w_{u,t} - \\
&-\lambda_{t-1}^{kqu} U_{c_u}(t) r_t^q - \lambda_{t-1}^{keu} U_{c_u}(t) r_t^e = 0
\end{aligned} \tag{22}$$

$$\begin{aligned}
&\tfrac{\partial U^g}{\partial \tau_{h,t}} = \lambda_t^{gbc}\left[n_h h_{h,t} w_{u,t}\right] - \lambda_t^{bch}\left[h_{h,t} w_{u,t}\right] - \\
&-\lambda_t^{lsh} U_{c_h}(t) w_{u,t} = 0
\end{aligned} \tag{23}$$

where λ_t^{lss}, λ_{t-1}^{kqs}, λ_{t-1}^{kes} and λ_t^{bcs} refer to the Lagrange multipliers attached to the unskilled household's FOCs for hours worked, capital structures, capital

[27] Note that the first-order conditions can be obtained by dropping any of the budget constraints or the aggregate resource constraint, given that these are linearly combined. To facilitate presentation, we present the FOCs for the government here assuming that the aggregate resource constraint is dropped, since this allows for a symmetric treatment of the arguments entering the FOCs.

equipment and the household's budget constraint respectively; and λ_t^{lsh}, and λ_t^{bch} refer to the Lagrange multipliers attached the hand-to-mouth FOC for hours worked, and the hand-to-mouth budget constraint respectively. The interpretation of these two first-order conditions follows the same reasoning as for the condition for the skilled income tax.

The above conditions for the income tax rates demonstrate that there are two main channels for the impact of income taxes on social welfare. First, they increase total revenue while decreasing disposable income of the households. Second, they distort the households' factor supply decisions. The magnitude of these channels may differ between households, given heterogeneity in opportunities, so that a quantitative evaluation is required to determine optimal taxes. In particular, the net benefit is expected to be increasing with the income level, because higher income implies both a higher tax base (and thus higher increase in the tax revenue) and a lower social welfare cost by reducing income and consumption for the wealthier households. The latter effect is driven by the concavity of the utility functions at the household and aggregate level. On the other hand, as explained below in more detail, the negative implications of the distortions in factor supplies (the "tax wedges") for aggregate productivity are stronger for income taxes on skilled households, followed by unskilled and hand-to-mouth.

4.1.2 Optimality conditions for bonds

The first-order condition for $B_{s,t+1}$ is given by:

$$\begin{aligned}\tfrac{\partial U^g}{\partial B_{s,t+1}} &= \lambda_t^{gbc} n_s + \beta \lambda_{t+1}^{bcs} \left(R_{t+1}^b - 2\psi_s B_{s,t+1}\right) - \\ &-\beta \lambda_{t+1}^{gbc} n_s R_{t+1}^b - \lambda_t^{bcs} - 2\beta \psi_s \lambda_t^{bs} U_{c_s}(t+1) = 0\end{aligned} \tag{24}$$

where λ_t^{bs} is the Lagrange multiplier attached to the skilled household's FOC for bonds. The first-order condition for the choice of $B_{s,t+1}$ dictates that this should be chosen so that the benefits from a marginal increase in $B_{s,t+1}$ must equal marginal costs. The marginal benefits are given by (i) the current increase in government revenue, n_s, transformed into social welfare units by λ_t^{gbc}; and (ii) the future net discounted increase in household income, $\beta\left(R_{t+1}^b - 2\psi_s B_{s,t+1}\right)$, valued in social welfare terms via λ_{t+1}^{bcs}. Note that $-2\psi_s B_{s,t+1}$ refers to the asset holding costs that need to be paid in the next period. The marginal costs are given by (i) the future discounted social welfare cost of the increased requirement for tax revenue to finance interest payments on debt, $\beta \lambda_{t+1}^{gbc} n_s R_{t+1}^b$; (ii) the current reduction in household disposable income via the purchase of bonds, λ_t^{bcs}; and (iii) the distortion due to asset holding costs, $2\beta\psi_s \lambda_t^{bs} U_{c_s}(t+1)$. As above, social welfare units are obtained via the shadow social price of the distortion, λ_t^{bs}.

Similarly, the first-order condition for $B_{u,t+1}$ is given by:

$$\begin{aligned} \tfrac{\partial U^g}{\partial B_{u,t+1}} &= \lambda_t^{gbc} n_u + \beta \lambda_{t+1}^{bcu} \left(R_{t+1}^b - 2\psi_u B_{u,t+1} \right) - \\ &- \beta \lambda_{t+1}^{gbc} n_u R_{t+1}^b - \lambda_t^{bcu} - 2\beta \psi_u \lambda_t^{bu} U_{c_u} \left(t+1 \right) = 0 \end{aligned} \tag{25}$$

where λ_t^{bu} is the Lagrange multiplier attached to the unskilled household's FOC for bonds.

4.1.3 Trade-off for optimal policy

Under perfect markets and a representative household, a smooth income tax creates fewer policy distortions in the household's optimisation problem and has been shown in the literature to be optimal (see e.g. Lucas and Stokey (1984) and Ljungvist and Sargent (2012)). However, under imperfect markets, a tax rate may need to be used to address market distortions, so that there can be an incentive to increase the volatility of a tax rate over the business cycle (see e.g. the search frictions example in Arseneau and Chugh (2012) within a representative household model). Furthermore, the literature has demonstrated that market frictions matter for the optimal degree of cyclicality of income taxes within a representative agent framework. For instance, in the neoclassical model the labour income tax is generally positively related with the TFP shocks and thus tends to be pro-cyclical. In contrast, in the labour market search frictions model of Arseneau and Chugh (2012), the labour income tax is optimally counter-cyclical.

Optimal policy in our framework requires the government to decide if all income taxes should have similar stochastic properties, in terms of volatility and cyclicality, or, if not, how they should differ for the different levels of income across households. The imperfections that limit participation in asset and labour markets as well as the production structure in our model economy create different channels for the stochastic properties of tax rates to affect aggregate welfare.

To decide how to set the three income taxes over the business cycle, the government must evaluate a key trade-off. On one hand, unfavourable income taxes, in the form of higher volatility and counter-cyclicality, imply higher marginal social welfare losses when applied to the lower income/consumption groups. This negative effect is particularly strong for hand-to-mouth households, since asset market exclusions imply that they lack the means to smooth consumption over time. At the same time, the tax revenue gains from counter-cyclical and more volatile taxes increase with the wealth of the tax base to which they apply. Therefore, the government has an incentive, for both reasons, to keep the taxes that apply to households with less wealth smoother and less counter-cyclical.

On the other hand, unfavourable taxes directed at the skilled households propagate more in the economy given that the factor supply choices of these households have a greater effect on capital accumulation. This is because the high income skilled households supply more capital than either of the unskilled households, and because of the complementarity between capital and skill. Thus, the skilled income tax implies higher distortions at the aggregate level than the other income taxes. This creates a strong incentive for the government to smooth the tax for the skilled household and to set their taxes to be the least counter-cyclical. By extension, income taxes to the middle income group of unskilled workers are less distorting than the skilled but more distorting than taxes to hand-to-mouth workers whose factor supply choices affect the propagation mechanism the least.

4.1.4 Solution

To numerically solve for the outcomes under optimal policy, we follow Arseneau and Chugh (2012) and first compute the deterministic steady-state equilibrium under optimal policy. We then next approximate the dynamic equilibrium paths using a second-order approximation of the equilibrium conditions under optimal policy for time $t > 0$ around the deterministic steady-state of these conditions.[28] Compared with Arseneau and Chugh (2012) who use a first-order approximation in a representative agent context, it is useful in our heterogeneous household setup that includes hand-to-mouth households not to impose certainty equivalence on decision making. This is because the effects of exogenous uncertainty on households depend on their ability to smooth consumption, thus affecting hand-to-mouth households differently from the remaining households.

4.2 Optimal policy in the long-run

The optimal steady-state tax rates and the debt-to-output ratio are presented in Table 6. For convenience, we repeat the data averages discussed earlier.

Table 6: Optimal policy in the steady-state

	τ_{60-100}	τ_{20-60}	τ_{0-20}	$\frac{B}{Y}$
Data average	0.247	0.180	0.144	2.120
	τ_s	τ_u	τ_h	$\frac{B}{Y}$
Steady-state	0.261	0.125	-0.017	0.000

[28] The presence of holding costs for debt in our setup implies that an asymptotic steady-state is well defined and is independent of initial conditions (see Appendices A and B for further details on the steady-state).

As can be seen, optimal income taxation is more progressive in the long-run than in the exogenous policy model. This implies that the government can increase aggregate welfare by using distorting taxes to redistribute some income from higher to lower income households. This result reflects the optimal resolution of the following trade-off. On one hand, since the Utilitarian government cares equally about all households and the social welfare function is concave with respect to the consumption of the households, aggregate welfare gains can be achieved by increasing consumption of lower income households and decreasing consumption of higher income households. On the other hand, however, since this redistribution requires the use of distorting income taxes, there is an efficiency loss that reduces average income and hence aggregate welfare for all households.

Further note that optimal debt in the long-run is zero in this setup. Ramsey optimal policy in a frictionless model usually implies that it is optimal for the government to hold non-zero assets in the long-run, since this allows it to change its tax revenue requirements to minimise tax distortions (see e.g. Chamley (1986)). However, the existence of imperfections in the asset markets, captured by the transaction costs for holding bonds, create a distortion in the intertemporal allocations of the household resulting in efficiency losses. The government finds it optimal to eliminate the intertemporal wedges associated with the holding costs of government assets in the steady-state, since they create permanent efficiency losses in the form of reductions in disposable income. This result bears similarities to the optimal elimination of the intertemporal capital tax wedge in the steady-state in the frictionless neoclassical model with a complete instrument set out in Chamley (1986).[29] Note however that, as we shall see below, the government does use debt over the business cycle to respond to economic fluctuations.

4.3 Optimal taxes and debt over the business cycle

We next compute fluctuations around the deterministic optimal steady-state by simulating the optimal-policy equilibrium under the same exogenous realisations to TFP, investment-specific technological change and government consumption spending as in the exogenous-policy simulations in Section 3 (see also Arseneau and Chugh (2012) for a similar approach).[30] We present

[29] Appendix B provides an analytical demonstration that optimally government debt is zero in the steady-state.

[30] To calculate the required statistics for optimal policy, we conduct simulations under shocks to the exogenous processes, which are initialised from the steady-state (in this case of the economy under optimal policy), obtain the required statistics and calculate their mean value across the simulations. We conduct 10,000 simulations of 96 periods.

the results for the optimal properties of the tax system for the benchmark model under all exogenous processes in Table 7 and under one process at a time in Table 8.

4.3.1 Volatility and cyclicality of taxes

The results in Table 7 suggest that optimal income taxes have volatilities that are about half to two-thirds as volatile as output. In the neoclassical model analysed in Chari *et al.* (1994), optimal labour income taxes are effectively flat over the business cycle, whereas taxes on capital income are significantly more volatile. Since income taxes in our setup apply to both capital and labour income, their volatility is between that found in studies concentrating on factor income taxation. The standard deviations reported in Table 7 lead to confidence intervals which indicate that 95% of the time the income taxes optimally vary over the cycle within a range of about 2.5 percentage points. The optimal standard deviation for the tax rates is quantitatively similar for the three tax rates, highest for τ_u and lowest for τ_h. Finally, government debt is more volatile than income taxes.

The quantitative resolution of the key trade-off faced by the policy maker discussed above leads to optimal income taxes which have similar business cycle properties. In particular, income taxes that apply to the middle-income households are more volatile than the remaining taxes, but the differences are not significant. All taxes are less volatile than output but exhibit sizeable variation over the business cycle compared to models without wage and asset inequality. Similarly, the output correlations of the taxes resulting from the optimal resolution of the trade-off set out above are all negative and similar across the household types.

Table 7: Optimal tax policy (business cycle)

		X_i		
	τ_s	τ_u	τ_h	B
		All shocks		
$\mu(X_i)$	0.261	0.125	-0.016	0.0016
95% $CI(X_i)$	(0.248, 0.273)	(0.110, 0.139)	(-0.027, -0.005)	(-0.032, 0.036)
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.613	0.698	0.519	1.632
$\rho(X_{i,t}, X_{i,t-1})$	0.464	0.722	0.902	0.995
$\rho(X_i, Y)$	-0.466	-0.478	-0.489	-0.292

Finally, the results in Table 7 show that debt responds countercyclically and is more volatile than output. This is despite the asset market frictions which imply that holding assets (including government bonds) has resource costs for the economy. Therefore, absorbing the fiscal effects of temporary exogenous shocks by changes in government bonds leads to changes in the aggregate disposable income, which needs to be traded-off against the distortions introduced by the volatility of the income taxes. The social cost of issuing government debt is over and above the burden imposed in future periods by the requirement to raise taxes to service the higher debt. The outcome of the trade-off in choosing debt versus taxes to respond to exogenous financing shocks suggests that, as commonly found in the literature, debt has an important shock-absorbing role, since it is significantly more volatile than the income taxes.

4.3.2 Optimal response to economic fluctuations

We next analyse the optimal reaction of tax policy to different exogenous shocks by examining the impulse responses (IRs) of the key economic variables after a temporary, output-reducing standard deviation shock to each of the exogenous processes. These are plotted in Figures 1 for all three exogenous shocks in the model. This allows us to evaluate how the government optimally changes the tax rates in the short- and medium-run in response to each of the exogenous shocks.

[**Figure 1 here**]

To complement the impulse response analysis, Table 8 presents the optimal properties of the tax system under separate sources of economic fluctuations. Isolating the effects of the processes can help to shed more light on the properties of optimal taxation analysed above.

Table 8: Optimal tax policy under individual shocks

	X_i			
	τ_s	τ_u	τ_h	B
			A shock	
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.476	0.619	0.495	1.505
$\rho(X_i, Y)$	-0.730	-0.467	-0.521	-0.303
			A^e shock	
$\frac{\sigma(X_i)}{\sigma(Y)}$	2.353	2.000	1.529	3.294
$\rho(X_i, Y)$	0.809	-0.950	-0.600	-0.362
			G^c shock	
$\frac{\sigma(X_i)}{\sigma(Y)}$	2.500	2.750	4.750	11.500
$\rho(X_i, Y)$	-0.416	-0.854	-0.798	-0.990

TFP shocks We start with TFP shocks. As can be seen in the first panel of Table 8, fluctuations driven by TFP shocks do not qualitatively change the relative volatilities of optimal taxes, compared with the base case in Table 7. The main difference is that τ^s becomes the most counter-cyclical tax, followed by τ_h and τ_u. These results are consistent with the impulse responses for taxes and key endogenous variables following a negative TFP shock, in Figure 1. As can be seen, as output is reduced, all taxes are increased on impact. However, while τ^h and τ^u continue to increase for about 20 quarters, before they start to decrease, τ^s is reduced immediately after the initial increase. Thus, its response mirrors that of income more closely. These tax patterns allow the government to induce similar fluctuations in factor inputs and in the consumption of the three types of household, despite the fact that these households have different access to the asset markets and thus to a smoothing mechanism.

Investment specific shocks The second moments of the optimal taxes under shocks to investment specific technological change, A_t^e, are presented in Table 8. These results suggest that τ^u and τ^h are strongly counter-cyclical, whereas τ^s is strongly pro-cyclical. Moreover, the standard deviations relative to output increase relative to Table 7 and are falling with respect to income.

Compared with the neutral TFP shocks, under skill-biased shocks to A_t^e, the incentive to dampen fluctuations in the endogenous propagation mechanism is stronger than the incentive to equalise the fluctuations in post-tax income and consumption for the different households. As can be seen in Figure 1, as output is reduced following a negative shock to A_t^e; τ^h and τ^u are increased, whereas τ^s is reduced. Given the potential large reduction in equipment capital accumulation, which significantly affects aggregate productivity, the government finds it optimal to reduce τ^s. Thus, increasing the complement to equipment capital in production and encouraging accumulation of equipment capital.

However, given that unskilled labour productivity depends positively on capital and skilled labour, the fall in τ^s also leads to increases in the return to unskilled labour (see the response of w^u in Figure 1). In contrast, the increase in the skill supply leads to a fall in w^s. Hence, pro-cyclical τ^s taxes act to indirectly support the return to unskilled labour. This allows the government to impose counter-cyclical τ^u and τ^h taxes, and thus insulate part of the revenue implications of the negative shock to A_t^e.[31]

[31]Obviously, the results are reversed under positive shocks to A_t^e. In this case, the government finds it optimal to distribute some of the welfare gains associated with the

Government spending shocks Under shocks to government spending, both the standard deviation and the correlation of debt increase. This indicates that the government finds it optimal to absorb more of pure temporary public finance shocks by changing its asset position, compared with technology shocks. Temporary public finance shocks do not have direct productivity implications, therefore their impact on households' income (and thus the importance of the first channel in the optimal policy trade-off) is second-order relative to the distortions caused by the taxes on economic choices. This further implies that the incentive to minimise distortions that propagate more, when taxes change, becomes stronger. As a result, the income tax rates that apply to skilled households are less volatile and less counter-cyclical, compared with households which offer unskilled labour.

4.4 Balanced budget restriction

We next explore the implications for optimal tax policy of a restriction imposed on economic policy which requires the government to run zero deficits over the business cycle. In particular, we assume that the government cannot issue new debt, but instead debt is restricted at the level required so that the steady-state debt-to-output ratio is equal to the data average. The results from this experiment are reported in Table 9. The inability to use debt to absorb the fiscal implications of exogenous shocks implies that income taxation needs to be, on average, more volatile and more counter-cyclical compared with the benchmark model in Table 7. Our interest here is in whether these increases would be proportional for all income taxes, or, if not, which type of household is affected more in terms of income taxation by this policy restriction.

Table 9: Balanced budget

	X_i			
	τ_s	τ_u	τ_h	B
$\mu(X_i)$	0.282	0.163	0.030	0.672
95% $CI(X_i)$	(0.267, 0.296)	(0.137, 0.190)	(0.008, 0.053)	N/A
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.619	1.153	0.983	N/A
$\rho(X_i, Y)$	-0.698	-0.826	-0.881	N/A

For this experiment, we find a large increase in the volatility and countercyclicality of the two unskilled (low and middle income) households and effectively no change in the volatility and countercyclicality of the tax to

skill biased change to the rest of the economy.

skilled. The changes in τ_u and τ_h are quantitatively significant. In particular, the 95% confidence intervals imply that under a balanced budget, τ_u and τ_h vary in ranges that cover 5.3 and 4.5 percentage units respectively. The difference with τ_s, the range of which remains at about 2.9 percentage units, is appreciable over the business cycle. Therefore, households providing unskilled labour bear the cost, in terms of higher income tax fluctuations, of the balanced budget restriction. In this case, the inevitable increase in economic volatility, brought about by the constraint to meet all public finance requirements by changes in taxes, implies that the quantitative importance of the factor supply distortions of skilled households increases.

4.5 The importance of risk-aversion

The risk aversion parameter σ was not calibrated to hit a target in Section 3.3.4, but instead was set to a commonly employed value in macroeconomic models. The extent of risk aversion characterising households' preferences may, however, matter for optimal policy over the business cycle. This is because it reflects the extent to which fluctuations in consumption and leisure affect welfare directly.[32] This may be relevant in our setup, since hand-to-mouth households are particularly exposed to fluctuations in income. We therefore present in Table 10 the results for flexible and fixed debt when $\sigma = 4$.[33]

As expected, when debt is optimally chosen, the higher costs attached to income fluctuations for hand-to-mouth households implies that these households experience the least volatile and countercyclical income taxes . However, the required increase in standard deviations, due to the balanced budget restriction, is again disproportionately allocated to the three households. As under $\sigma = 2$, the standard deviation of τ_u increases significantly, while that of τ_s effectively does not change. The change in the standard deviation of τ_s is smaller compared with the results $\sigma = 2$. On the other hand, the countercyclicalities increase more for households providing unskilled labour services, compared with skilled households, reflecting the incentive to keep the factor supply choices of skilled households smoother.

Overall, higher degrees of risk-aversion imply an increased incentive to smooth income for hand-to-mouth and dampen quantitatively the differences

[32] Analyses on optimal taxes over the business cycle generally consider the importance of higher values of risk aversion (see e.g. Chari *et al.* (1994) and Stockman (2001)).

[33] Note that increasing σ does not significantly affect most second moments of the model under exogenous policy reported in Table 5. However, it does lead to bigger increases in the output correlation with the skill premium. Thus it worsens the overall fit of the model relative to the benchmark case under $\sigma = 2$ presented in Section 3.3.4.

between the volatilities of optimal income taxes. This is because the direct effects of fluctuations on households' welfare become quantitatively more important relative to the effects of market exclusions on their income.

Table 10: Optimal tax policy with higher RRA

	X_i			
	τ_s	τ_u	τ_h	B
		Flexible debt		
$\mu(X_i)$	0.296	0.118	-0.069	0.0026
95% $CI(X_i)$	(0.279, 0.312)	(0.100, 0.136)	(-0.084, -0.054)	(-0.030, 0.035)
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.800	0.867	0.724	1.571
$\rho(X_i, Y)$	-0.549	-0.422	-0.264	-0.340
		Balanced budget		
$\mu(X_i)$	0.320	0.160	-0.025	0.672
95% $CI(X_i)$	(0.303, 0.337)	(0.129, 0.190)	(-0.042, -0.009)	N/A
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.770	1.398	0.752	N/A
$\rho(X_i, Y)$	-0.598	-0.771	-0.737	N/A

5 Conclusions

This paper analysed the optimal distribution of the tax burden over the business cycle for households differentiated by unequal access to labour and asset markets. In particular, we assumed that a subset of households provided skilled labour services, whereas the rest worked as unskilled. Moreover, participation premia in the asset markets varied in such a way that a subset of the unskilled households was excluded from investing in the capital stock and government bonds.

The model was shown to capture the empirical regularities of macroeconomic variables over the business cycle and was consistent with key features of the labour markets and wealth ownership that we considered. Our analysis considered the problem of a government that chose the paths of income tax rates and debt to maximise aggregate welfare. In doing so we constrained the policy menu of the government to focus on income taxes that applied to total household income. We then examined the implications for optimal government policy under fixed public debt.

We found that, under a flexible budget, the income taxation burden was spread roughly evenly across the three households. However, under a balanced budget restriction, the burden was distributed least favourably to the middle-income and most favourably to the high income households.

References

[1] Acemoglu, D. and D. Autor (2011). 'Skills, tasks and technologies: implications for employment and earnings', in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, 4b, 1043-1171, Amsterdam: Elsevier.

[2] Aghion, P. and P. Howitt (2009). 'The economics of growth', MIT Press, Cambridge, Mass.

[3] Aiyagari, S., Marcet, A., Sargent, T. and J. Seppala (2002). 'Optimal taxation without state-contingent debt', *Journal of Political Economy*, 110 (6). 1220–1254.

[4] Angelopoulos, K., Asimakopoulos, S., and J. Malley (2014). 'Tax smoothing in a business cycle model with capital-skill complementarity', *Journal of Economic Dynamics and Control*, ISSN 0165-1889, (doi:10.1016/j.jedc.2014.11.002) (In Press).

[5] Angelopoulos, K., Fernandez, B. and J. Malley (2014). 'The distributional consequences of tax reforms under capital-skill complementarity', *Economica*, 81, 747-767.

[6] Angelopoulos, K., Malley, J. and A. Philippopoulos (2013). 'Human capital, social mobility and the skill premium', *CESifo Working Paper 4388*.

[7] Arseneau, D. and S. Chugh (2012). 'Tax smoothing in frictional labor markets', *Journal of Political Economy*, 120, 926-985.

[8] Bernheim, B. and D. Garrett (2003). 'The effects of financial education in the workplace: evidence from a survey of households', *Journal of Public Economics*, 87, 1487–1519.

[9] Benigno, P. (2009). 'Price stability with imperfect financial integration', *Journal of Money, Credit and Banking*, 41, 121-49.

[10] Blau, F. and L. Kahn (2007). 'Changes in the labor supply behavior of married women: 1980–2000', *Journal of Labor Economics*, 25, 393-438.

[11] Campbell, J. and G. Mankiw (1989). 'Consumption, income, and interest rates: Reinterpreting the time-series evidence', in O.J. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*, 185-216, MIT Press.

[12] Castro R. and D. Coen-Pirani (2008). 'Why have aggregate skilled hours become so cyclical since the mid-1980s?', *International Economic Review*, 49, 135-184.

[13] Chamley, C. (1986). 'Optimal taxation of capital income in general equilibrium with infinite lives', *Econometrica*, 54 (3), 607-622.

[14] Chari, V., Lawrence, C. and P. Kehoe (1994). 'Optimal fiscal policy in a business cycle model', *The Journal of Political Economy*, 102, 617-652.

[15] Chari, V. and P. Kehoe (1999). 'Optimal fiscal and monetary policy', in J. Taylor and M. Woodford (eds.), Handbook of Macroeconomics, 1, 1671-1745, Amsterdam: Elsevier.

[16] Chetty, R., Guren, A., Manoli, D. and A. Weber (2011). 'Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins', *American Economic Review*, 101, 471-75.

[17] Cogan, J., Cwik, T., Taylor, J. and V. Wieland (2010). 'New Keynesian versus old Keynesian government spending multipliers', *Journal of Economic Dynamics and Control*, 34, 281-295.

[18] Domeij, D. and M. Floden (2006). 'The labor-supply elasticity and borrowing constraints: Why estimates are biased', *Review of Economic Dynamics*, 9(2), 242-262.

[19] Farhi, E. (2010). 'Capital taxation and ownership when markets are incomplete', *Journal of Political Economy*, 118 (5), 908-948.

[20] Galí, J., Lopez-Salido, D. and J. Valles (2007). 'Understanding the effects of government spending on consumption', *Journal of the European Economic Association*, 5, 227-270.

[21] Gandelman, N. and R. Hernández-Murillo (2013). 'What do happiness and health satisfaction data tell us about relative risk aversion?', *Journal of Economic Psychology*, 39, 301-312.

[22] Goldin, C. and L. Katz (2008). 'The race between education and technology', The Belknap Press of Harvard University Press.

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

[24] He, H. (2012). 'What Drives the Skill Premium: Technological change or demographic variation?', *European Economic Review*, 56, 1546-1572.

[25] He, H and Z. Liu (2008). 'Investment-specific technological change, skill accumulation, and wage inequality', *Review of Economic Dynamics*, 11, 314-334.

[26] Hornstein, A., Krusell, P. and G. Violante (2005). 'The effects of technical change on labor market inequalities', in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, vol. 1, 1275-1370. Amsterdam: Elsevier.

[27] Katz, L. and K. Murphy (1992). 'Changes in relative wages, 1963-1987: supply and demand factors', *The Quarterly Journal of Economics*, 107, 35-78.

[28] Keane, M. and R. Rogerson (2012). 'Micro and macro labor supply elasticities: A reassessment of conventional wisdom', *Journal of Economic Literature*, 50, 464-476.

[29] Kocherlakota, N. (2010). 'The new dynamic public finance', Princeton University Press, Princeton, New Jersey.

[30] Krusell, P., Ohanian, L., Ríos-Rull J. and G. Violante (2000). 'Capital–skill complementarity and inequality: a macroeconomic analysis', *Econometrica*, 68, 1029–1053.

[31] Lindquist, M. (2004). 'Capital–skill complementarity and inequality over the business cycle', *Review of Economic Dynamics*, 7, 519-540.

[32] Litterman, R. (1983). 'A random walk, Markov model for the distribution of time-series', *Journal of Business & Economic Statistics*, 1, 169-173.

[33] Ljungqvist L. and T. Sargent (2012). 'Recursive macroeconomic theory', 3rd edition, The MIT Press, Cambridge, Massachusetts.

[34] Lucas, R. and N. Stokey (1983). 'Optimal fiscal and monetary policy in an economy without capital', *Journal of Monetary Economics*, 12 (1). 55-93.

[35] Mankiw, G. (2000). 'The savers-spenders theory of fiscal policy', *American Economic Review*, 90, 120-125.

[36] Mankiw, G., Weinzierl, M. and D. Yagan (2009). 'Optimal taxation in theory and practice', *Journal of Economic Perspectives*, 23, 147-174.

[37] Matsuyama K. (2006). 'The 2005 Lawrence R. Klein Lecture: emergent class structure', *International Economic Review*, 47, 327-360.

[38] Persson, T. and G. Tabellini (1992). 'The politics of 1992: fiscal policy and European integration', *Review of Economic Studies*, 59, 689-701.

[39] Philippon, T. (2014). 'Has the American finance industry become less efficient?', *The American Economic Review* (forthcoming).

[40] Piketty, T. and E. Saez (2007). 'How progressive is the U.S. federal tax system? A historical and international perspective', *Journal of Economic Perspectives*, 21, 3-24.

[41] Pourpourides, P. (2011). 'Implicit contracts and the cyclicality of the skill-premium', *Journal of Economic Dynamics and Control*, 35, 963-979.

[42] Schmitt-Grohé, S. and M. Uribe (2003). 'Closing small open economy models', *Journal of International Economics*, 61, 163-185.

[43] Stockman, D. (2001). 'Balanced-budget rules: Welfare loss and optimal policies', *Review of Economic Dynamics*, 4, 438-459.

[44] Traum, N. and S. Yang (2010). ' Does government debt crowd out investment? A Bayesian DSGE approach', Congressional Budget Office Working Paper 2010-02.

[45] Werning, I. (2007). 'Optimal fiscal policy with redistribution', *The Quarterly Journal of Economics*, 122, 925-967

Appendix A: FOCs for households and firms

Skilled households

The first-order conditions of skilled households with respect to their choice variables are:

Hours worked

$$U_{c_s}(t)(1-\tau_{s,t})w_{s,t}+U_{h_s}(t)=0 \tag{A1}$$

Capital structures

$$\begin{aligned}&\beta\left\{E_tU_{c_s}(t+1)\left[(1-\tau_{s,t+1})r^q_{t+1}+(1-\delta^q)-2\psi_sK^q_{s,t+1}\right]\right\}-\\&-U_{c_s}(t)=0\end{aligned} \tag{A2}$$

Capital equipment

$$\begin{aligned}&\beta\left\{E_tU_{c_s}(t+1)\left[(1-\tau_{s,t+1})r^e_{t+1}+\frac{(1-\delta^e)}{A^e_{t+1}}-2\psi_sK^e_{s,t+1}\right]\right\}-\\&-\frac{U_{c_s}(t)}{A^e_t}=0\end{aligned} \tag{A3}$$

Bonds

$$\beta\left\{E_tU_{c_s}(t+1)\left[R^b_{t+1}-2\psi_sB_{s,t+1}\right]\right\}-U_{c_s}(t)=0 \tag{A4}$$

where $U_{c_j}(t)=\left[\left(\mu C^{\gamma}_{j,t}+(1-\mu)(1-h_{j,t})^{\gamma}\right)^{\frac{1}{\gamma}}\right]^{-\sigma}\mu C^{\gamma-1}_{j,t}$ and $U_{h_j}(t)=-\left[\left(\mu C^{\gamma}_{j,t}+(1-\mu)(1-h_{j,t})^{\gamma}\right)^{\frac{1}{\gamma}}\right]^{-\sigma}(1-\mu)(1-h_{j,t})^{\gamma-1}$ for $j=s,u,h$. These first-order conditions equate the marginal benefits from labour hours and investment in the two types of capital and bonds to their respective marginal costs.

Unskilled households

Hours worked

$$U_{c_u}(t)(1-\tau_{u,t})w_{u,t}+U_{h_u}(t)=0 \tag{A5}$$

Capital structures

$$\begin{aligned}&\beta\left\{E_tU_{c_u}(t+1)\left[(1-\tau_{u,t+1})r^q_{t+1}+(1-\delta^q)-2\psi_uK^q_{u,t+1}\right]\right\}-\\&-U_{c_u}(t)=0\end{aligned} \tag{A6}$$

Capital equipment

$$\beta\left\{E_t U_{c_u}(t+1)\left[(1-\tau_{u,t+1})\, r^e_{t+1}+(1-\delta^e)-2\psi_u K^e_{u,t+1}\right]\right\}- \\ -U_{c_u}(t)=0 \tag{A7}$$

Bonds

$$\beta\left\{E_t U_{c_u}(t+1)\left[R^b_{t+1}-2\psi_u B_{u,t+1}\right]\right\}-U_{c_u}(t)=0 \tag{A8}$$

These first-order conditions equate the marginal benefits from labour hours and investment in the two types of capital and bonds to their respective marginal costs. Note that the presence of the asset holding costs, ψ_s and ψ_u ensures that (A4) and (A8) differ in the steady-state, which is required so that the latter is uniquely determined. Moreover, given the assumption of different depreciation and tax rates, (A2), (A3), (A6) and (A7) also differ from each other in the steady-state, irrespective of ψ_s and ψ_u.

Hand-to-mouth households

Hours worked

$$U_{c_h}(t)\,(1-\tau_{h,t})\, w_{u,t}+U_{h_h}(t)=0 \tag{A9}$$

The first-order condition equates the marginal benefit from labour hours to its respective marginal cost.

Firms

Profit maximisation leads to the usual first-order conditions equating marginal products of factor inputs to their prices. In particular, the FOCs of the firm include:

$$MPK^q=r^q_t \tag{A10}$$

where $MPK^q=\frac{\alpha Y_t}{K^{f,q}_t}$ and Y_t is given by (1) in the main text.

$$MPK^e=r^e_t \tag{A11}$$

where

$$MPK^e=A_t\left(K^{f,q}_t\right)^{\alpha}\{\lambda\left[\nu\left(K^{f,e}_t\right)^{\rho}+(1-\nu)\left(h^f_{s,t}\right)^{\rho}\right]^{\varphi/\rho}+ \\ +(1-\lambda)\left(h^f_{u,t}\right)^{\varphi}\}^{\frac{1-a-\varphi}{\varphi}}\left[\nu\left(K^{f,e}_t\right)^{\rho}+(1-\nu)\left(h^f_{s,t}\right)^{\rho}\right]^{\frac{\varphi-\rho}{\rho}}\times \\ \times(1-a)\,\lambda\nu\left(K^{f,e}_t\right)^{\rho-1}$$

$$MPL^s = w_{s,t} \tag{A12}$$

where:

$$\begin{aligned} MPL^s = {} & A_t \left(K_t^{f,q}\right)^{\alpha} \{\lambda \left[\nu \left(K_t^{f,e}\right)^{\rho} + (1-\nu) \left(h_{s,t}^f\right)^{\rho}\right]^{\varphi/\rho} + \\ & + (1-\lambda) \left(h_{u,t}^f\right)^{\varphi} \}^{\frac{1-a-\varphi}{\varphi}} \left[\nu \left(K_t^{f,e}\right)^{\rho} + (1-\nu) \left(h_{s,t}^f\right)^{\rho}\right]^{\frac{\varphi-\rho}{\rho}} \times \\ & \times (1-a) \lambda (1-\nu) \left(h_{s,t}^f\right)^{\rho-1} \end{aligned}$$

$$MPL^u = w_{u,t} \tag{A13}$$

where:

$$\begin{aligned} MPL^u = {} & A_t \left(K_t^{f,q}\right)^{\alpha} \{\lambda \left[\nu \left(K_t^{f,e}\right)^{\rho} + (1-\nu) \left(h_{s,t}^f\right)^{\rho}\right]^{\varphi/\rho} + \\ & + (1-\lambda) \left(h_{u,t}^f\right)^{\varphi} \}^{\frac{1-a-\varphi}{\varphi}} (1-a) (1-\lambda) \left(h_{u,t}^f\right)^{\varphi-1} \end{aligned}$$

Note that when $1 > a+\varphi$ and $\varphi > \rho$, as is the case in our calibration, MPL_s increases in $K_t^{f,q}$, $K_t^{f,e}$and $h_{u,t}^f$, while MPL_u increases in $K_t^{f,q}$, $K_t^{f,e}$and $h_{s,t}^f$.

Appendix B: Optimal steady-state debt

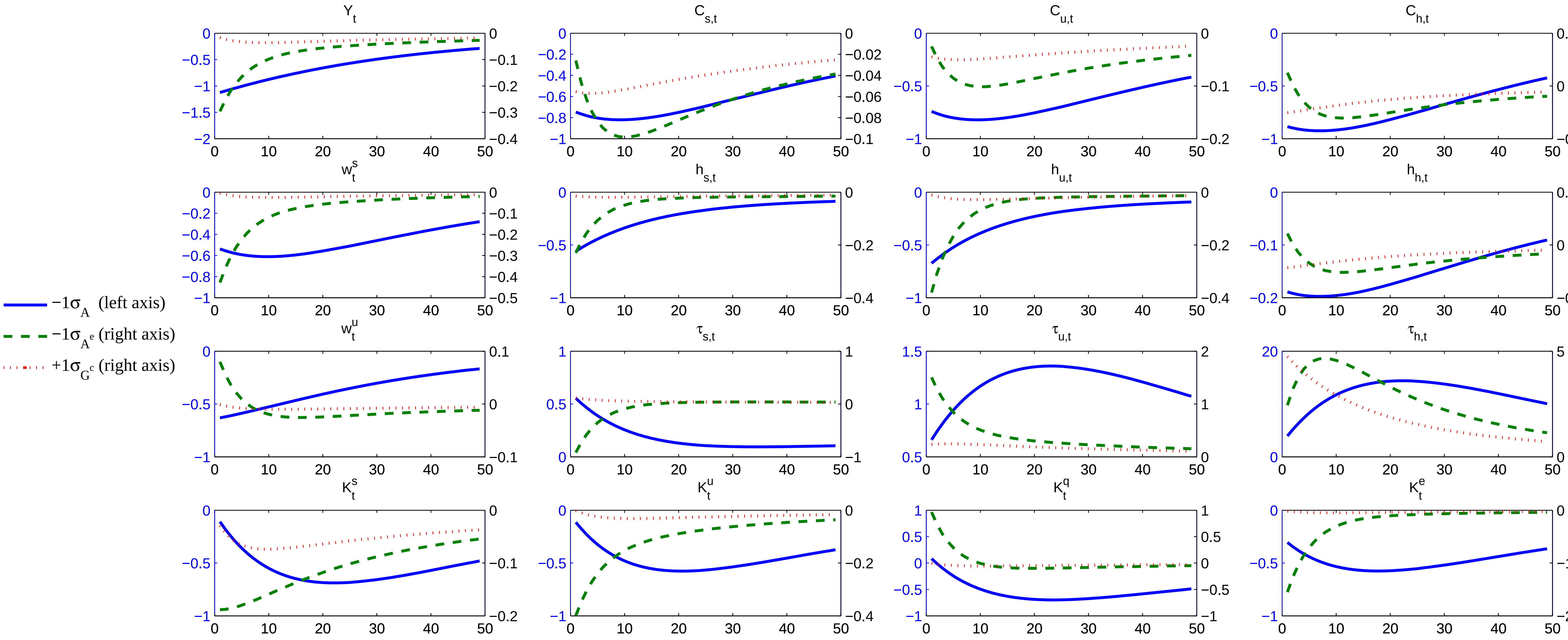
Conditions (24) and (25) imply that in the steady-state $B_s = B_u = 0$, if ψ_s and ψ_u are different from zero. To see this, e.g. for (24), note that its steady-state solution, if $\psi_s \neq 0$ and using (A4) in the steady-state for R^b, implies that:

$$B_s = -\frac{\lambda^{bs} U_{c_s}}{\lambda^{gbc}} \tag{B1}$$

where $U_{c_s} > 0$ and $\lambda^{gbc} > 0$. If the household holds government bonds in the steady-state, so that $B_s > 0$, then $\lambda^{bs} > 0$, since the latter represents marginal valuation of additional assets for the household. But then, the right hand side of (B1) is negative, hence $B_s > 0$ cannot be a solution. Conversely, if $B_s < 0$, so that $\lambda^{bs} < 0$, the right hand side of (B1) is positive, hence $B_s < 0$ cannot be a solution. Therefore, the only admissible solution is $B_s = 0$, which implies $\lambda^{bs} = 0$. Off steady-state, nevertheless, the optimal solution for $B_{s,t+1}$ may be positive or negative.

If $\psi_s = \psi_u = 0$, (24) and (25) in the steady-state are always satisfied when (A4) and (A8) are and cannot be used to pin down debt. For instance, (24) becomes $\left(\lambda^{gbc} n_s - \lambda^{bcs}\right)\left(\beta R^b - 1\right) = 0$, or, using (A4), $0 = 0$. In this case, the steady-state conditions are not sufficient to pin down an optimal allocation.

Figure 1: Impulse responses to temporary shocks



All plots are in percent deviations from the steady−state