



Citation for published version:

Angelopoulos, K, Asimakopoulos, S & Malley, J 2017, 'The optimal distribution of the tax burden over the business cycle', *Macroeconomic Dynamics*. <https://doi.org/10.1017/S1365100517000700>

DOI:

[10.1017/S1365100517000700](https://doi.org/10.1017/S1365100517000700)

Publication date:

2017

Document Version

Peer reviewed version

[Link to publication](https://doi.org/10.1017/S1365100517000700)

This article has been published in *Macroeconomic Dynamics* <https://doi.org/10.1017/S1365100517000700>. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © copyright holder.

University of Bath

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

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

Konstantinos Angelopoulos and CESifo
University of Glasgow

Stylianos Asimakopoulos
University of Bath

James Malley
University of Glasgow and CESifo

June 9, 2017

Abstract

This paper analyses optimal capital and labour income taxation for households differentiated by labour skill, income and wealth, under a balanced government budget, over the business cycle. 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 U.S. economy, as well as the empirical observations on wage (skill premium) and wealth inequality. We find that optimal taxes for middle-income households are more volatile than the remaining taxes. Moreover, the government re-allocates the total tax burden in bad times so that the share of total tax revenue paid by middle-income households rises. This share also rises for low-income households but by significantly less, while the tax share for skilled households falls.

Keywords: optimal taxation, business cycle, skill premium, income distribution
JEL Classification: E24, E32, E62

*We would like to thank the editor, William Barnett, two anonymous referees and an associate editor, 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 CESifo Macro Area Conference 2015, Royal Economic Society 2014 Annual Conference, the Universities of Kent, 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.

1 Introduction

There is a considerable literature that aims to characterize the properties of optimal capital and labour income tax policy over the business cycle. This research is undertaken 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, whereas capital income taxes should fluctuate around zero. Werning (2007) 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.* (2015), in a representative household business cycle model with state contingent debt, capital-skill complementarity and endogenous skill acquisition finds that optimal skilled and unskilled labour tax smoothing holds. In contrast, Arseneau and Chugh (2012) find that under search frictions in the labour market, the optimal labour income tax becomes very volatile and counter-cyclical. Additionally, Stockman (2001) shows that a balanced-budget restriction in a representative-agent model implies that although labour income taxes are still smoother than capital taxes, the volatility of labour taxes relative to that of capital taxes increases.

However, the literature has not examined optimal capital and labour 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 production¹ and the importance of "hand-to-mouth" consumers for macroeconomic stabilisation policy.² Moreover, the business cycle properties of optimal income taxes when the government cannot issue debt to respond to exogenous shocks have not yet been examined in this environment. This 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 use the tax rates to meet its spending requirements in response to exogenous shocks over the business cycle.

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

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

In light of the above, we aim to analyse optimal capital and labour income taxation over the business cycle when the government cannot issue debt to respond to exogenous shocks in setup 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 achieve this, we develop a model economy characterised by capital-skill complementarity in the production sector, 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 skilled and unskilled labour markets. We assume that skilled households are college-educated. 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 result in households which also differ with respect to their participation in the asset markets. 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 household heterogeneity associated with differences in opportunities to participate in markets for labour and assets. In capital markets, frictions lead to different returns and participation. In labour markets, socio-economic barriers, in the form of limited access to funding for training, family background and neighbourhood and other peer effects, exclude certain socio-economic groups from the skilled work. 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 capital and labour income taxes for three key socio-economic groups.

We calibrate a version of the model with exogenous tax policy to the U.S. quarterly data and find that the model fits the key long-run stylized facts

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

as well as the main cyclical properties of the data, i.e. the skill premium is effectively a-cyclical and not volatile.⁴ We then characterize optimal policy by allowing a Utilitarian government to choose capital and labour income tax rates for different households over the business cycle to maximise aggregate welfare, given its revenue requirements.

To decide how to set taxes over the business cycle, the government must evaluate a key trade-off. On one hand, unfavourable income taxes, in the form of higher levels, higher volatility and counter-cyclical, imply higher marginal social welfare losses when applied to the lower income groups. This negative effect is especially 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 tax base to which they apply. Therefore, the government has an incentive to keep the taxes that apply to households with less income smoother and less counter cyclical. On the other hand, unfavourable labour income taxes for skilled households propagate more in the economy given the complementarity between capital and skill labour. This creates a strong incentive for the government to keep the labour tax for the skilled household smooth and pro-cyclical.

We find that optimal labour taxes are less volatile than output and less volatile than capital taxes, similar to Stockman (2001). Our main results pertain to the comparisons of capital and labour income taxes between households with different income levels. Overall, we find that optimal labour and effective tax rates for middle-income households are more volatile than the remaining taxes. We also find that the tax on skilled labour income is pro-cyclical, whereas the labour income tax is a-cyclical for the middle and low income households. Finally, effective tax rates for the top two income groups are counter-cyclical whereas they are again a-cyclical for the low income households.

Given the above trade-off for optimal tax policy, middle-income households are relatively less exposed to economic fluctuations than hand-to-mouth households. Moreover, their factor supply choices are relatively less critical for amplifying exogenous shocks than skilled households. The labour tax to skilled households, in particular, is the least volatile income tax under productivity shocks. The protection afforded by the optimal tax system to skilled labour supply is greater under productivity shocks where the role of skilled hours in counteracting aggregate productivity losses is more important. Overall, due to exclusions in asset and labour markets, under

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

capital-skill complementarity, the importance that skilled workers enjoy in production and the vulnerability of hand-to-mouth consumers to income fluctuations imply an incentive to change taxes least favourably for the middle income group over the business cycle.

Our analysis shows that in bad times, associated with negative productivity shocks and increased expenditure requirements, the different tax rates do not change proportionately. Therefore, there is a redistribution of the tax burden between the three income groups in response to such negative shocks. We quantify this redistribution by calculating the change in the share of the total tax revenue paid by each of the three income groups. We find that the relative tax contributions paid by middle-income households rise, followed by a significantly smaller increase for low-income households, while the relative tax share for skilled households falls.

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 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

⁵When 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.* (2017) 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 (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.

correlation with output is -0.18.⁶ These findings suggest that the proportions of skilled (college educated) and unskilled workers do not change significantly over the business cycle. 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 and in particular that skilled labour complements capital relatively more than unskilled labour.

We also allow for asset holding costs when participating in capital 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 consume all their (labour) income (see e.g. Campbell and Mankiw (1989), Mankiw (2000) and 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 con-

⁶This 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.

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

⁸Data 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.

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

stant. 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 employing labour and capital inputs to produce output, Y_t . The production technology is characterised by capital-skill complementarity (see e.g. Goldin and Katz (2008) for historical evidence on the empirical relevance of this technology in the 20th century). In particular, the constant returns to scale (CRS) production function is given by a constant elasticity of substitution (CES) specification, similar to e.g. Krusell *et al.* (2000):¹⁰

$$Y_t = A_t \left\{ \lambda \left[\nu \left(k_t^f \right)^\rho + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right]^{\frac{\varphi}{\rho}} + (1 - \lambda) \left(h_{u,t}^f \right)^\varphi \right\}^{\frac{1}{\varphi}} \quad (1)$$

where k_t^f is the quantity of capital used by the firm, whereas $h_{s,t}^f$ and $h_{u,t}^f$ denote the quantities of skilled and unskilled labour respectively. In this specification, $A_t > 0$ is the level of total factor productivity (TFP); $\varphi, \rho < 1$ are the parameters determining the factor elasticities, i.e. $1/(1 - \varphi)$ is the elasticity of substitution between capital and unskilled labour and between skilled and unskilled labour, whereas $1/(1 - \rho)$ is the elasticity of substitution between capital and skilled labour; and $0 < \lambda, \nu < 1$ are the factor share parameters. 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:

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

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:

$$\Pi_t = Y_t - w_{s,t} h_{s,t}^f - w_{u,t} h_{u,t}^f - r_t K_t^f \quad (3)$$

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

¹⁰Recent studies in the dynamic general equilibrium (DGE) literature which employ a production function with different degrees of complementarity of capital with skilled and unskilled labour include e.g. Lindquist (2004), Pourpourides (2011) and He (2012).

2.2 Households

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

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

where E denotes expectations; $0 < \beta < 1$ 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:

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

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:

$$1 = l_{j,t} + h_{j,t} \quad (6)$$

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

$$(1 - \tau_{m,t}^l) w_{m,t} h_{m,t} + (1 - \tau_{m,t}^k) r_t K_{m,t} = C_{m,t} + I_{m,t} + \psi_m (K_{m,t})^2 \quad (7)$$

while hand-to-mouth households face the constraint:

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

where $I_{m,t}$ is investment in new capital, $K_{m,t}$; $w_{m,t}$ is the wage rate; r_t , is the net return to capital; and $\psi_m > 0$ measure the holding costs for capital.

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 income, since they face different holding costs. Skilled and unskilled households face finite holding costs, modelled here as quadratic functions of the capital 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, the households face different capital and labour income taxes.

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

$$K_{m,t+1} = (1 - \delta)K_{m,t} + A_t^e I_{m,t} \quad (9)$$

where, $0 \leq \delta \leq 1$. The capital evolution equation allows for an exogenous process, A_t^e , 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_t^e , is assumed to follow a stochastic exogenous AR(1) process:

$$\log(A_{t+1}^e) = (1 - \rho_{A^e}) \log(A^e) + \rho_{A^e} \log(A_t^e) + \varepsilon_t^{A^e} \quad (10)$$

where $\varepsilon_t^{A^e}$ 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, A^e 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 a sequence of state-dependent actions $\{C_{m,t}, h_{m,t}, I_{m,t}, K_{m,t+1}\}_{t=0}^{\infty}$, to maximise (4) subject to (6), (7) and (9), by taking initial conditions, policy variables and prices as given. Similarly, hand-to-mouth households, $j = h$, choose a sequence of state-dependent actions $\{C_{h,t}, h_{h,t}\}_{t=0}^{\infty}$, to maximise (4) subject to (6) and (8), by taking initial conditions, policy variables and prices as given. Exogenous processes are independent of one another. The optimality conditions for the households are given in the Appendix.

2.3 The government

The government's budget constraint is given by:

$$G_t = \tau_{s,t}^l n_s w_{s,t} h_{s,t} + \tau_{s,t}^k n_s r_t K_{s,t} + \tau_{u,t}^l n_u w_{u,t} h_{u,t} + \tau_{u,t}^k n_u r_t K_{u,t} + \tau_{h,t}^l n_h w_{u,t} h_{h,t} \quad (11)$$

where, G_t is average government expenditure per household.

The government budget in (11) is constrained to be balanced in each period, hence the government cannot issue debt to finance G_t .¹¹ In Section 3 below, as part of the model calibration, we consider a policy regime where the tax rates are set exogenously. To avoid public financing issues, we assume that G_t is the residual instrument in equation (11).¹² In contrast, when we examine optimal policy in Section 4, the government chooses all tax rates to maximise aggregate welfare under the constraint that (11) is satisfied period by period and G_t follows an exogenous AR(1) process:

$$\log(G_{t+1}) = (1 - \rho_G) \log(G) + \rho_G \log(G_t) + \varepsilon_t^G \quad (12)$$

where $\varepsilon_t^G \sim iidN(0, \sigma_G^2)$. Since we focus on the revenue side of the budget constraint for optimal policy, we assume that G_t is wasteful. Thus, innovations to G_t act as exogenous shocks to the requirement for tax revenue that needs to be generated (see e.g. Chari *et al.* (1994), Stockman (2001) and Arseneau and Chugh (2012) for a similar approach regarding G_t in the optimal taxation literature).

2.4 Market clearing conditions

The labour and capital market clearing conditions are given by:

$$h_s^f = n_s h_s \quad (13)$$

$$h_u^f = n_u h_u + n_h h_h \quad (14)$$

$$K_t = n_s K_{s,t} + n_u K_{u,t}. \quad (15)$$

The aggregate resource constraint is:

$$Y_t = G_t + n_s C_{s,t} + n_u C_{u,t} + n_h C_{h,t} + n_s I_{s,t} + n_u I_{u,t} + n_s \psi_s (K_{s,t})^2 + n_u \psi_u (K_{u,t})^2. \quad (16)$$

¹¹Implicitly, the amount of government debt and its repayment are a fixed quantity, which is reflected in the calibration of the level of G below, in that we allow the level of G in the steady-state to be determined by tax revenue collected by the government. Stockman (2001) considers optimal capital and labour taxes under a balanced-budget restriction, in a setup with a representative household, by fixing government debt and requiring the government to generate revenue for debt repayments, in addition to financing government spending. Here we focus only on the latter objective, as recent experience has shown that interest payments on debt can depend critically on the type of debt held by the government and on financial market conditions not modelled here.

¹²A non-distorting policy instrument which only serves to balance the government budget under exogenous policy is commonly used in the analysis of optimal taxation. For instance, Farhi (2010) and Arseneau and Chugh (2012) use a lump-sum transfer for this purpose. In our setup, household heterogeneity implies that lump-sum transfers to households can affect allocations. Hence, we let government spending to be residually determined when studying exogenous policy.

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 technology processes and labour and capital income tax rate processes.

Given stationary stochastic processes for the tax rates $\{\tau_{s,t}^l, \tau_{u,t}^l, \tau_{h,t}^l, \tau_{s,t}^k, \tau_{u,t}^k\}_{t=1}^{\infty}$ and for technology $\{A_t, A_t^e\}_{t=1}^{\infty}$, initial values for the exogenous processes and initial levels for assets, $K_{s,0}$ and $K_{u,0}$, the DCE system of equations is characterized by stochastic processes for allocations $\{C_{s,t}, C_{u,t}, C_{h,t}, h_{s,t}, h_{u,t}, h_{h,t}, K_{s,t+1}, K_{u,t+1}, I_{s,t}, I_{u,t}, G_t\}_{t=0}^{\infty}$, and for prices $\{w_{s,t}, w_{u,t}, r_t\}_{t=0}^{\infty}$ 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.

3.1 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 to key endogenous variables in our model. These are taken directly from the results reported in Lindquist (2004) and Pourpourides (2011) and refer to quarterly data for the period 1979-2002 and 1979-2003 respectively.¹³

As can be seen in Table 1, these studies document some interesting results regarding the labour market data. 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

¹³To 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 occupations 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 given by 0.49 and -0.13 respectively.

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.¹⁴ The findings 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
$\hat{\rho}(X_i, Y)$	1	0.83	0.76-0.91	0.09-0.19	0.69	0.73
$\hat{\sigma}(X_i)$	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).

3.2 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, 3A and 3B. Table 2 reports the parameters referring to functional forms and constraints associated with the households' and firms' problems. Table 3A reports the parameters for the exogenous processes required to solve and simulate the model under exogenous policy. In contrast, to solve and simulate the model under optimal policy requires the two productivity processes from Table 3A plus an additional process for government spending whose parameters are reported in Table 3B.

3.2.1 Population shares

We assume that the population breakdown in our model economy is given as $n_s = 0.4$, $n_u = 0.4$, $n_h = 0.2$.¹⁵ 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 without any assets is 18.7%.¹⁶ It also broadly coheres with the data in

¹⁴The statistics for labour supply in Table 1 are reported in Pourpourides (2011) and the underlying data source is Castro and Coen-Pirani (2008). Labour supply data in Lindquist (2004) and output data by US National Income and Product Accounts also generate labour statistics for skilled labour that are comparable with those for unskilled labour, with marginally higher output correlation for skilled relative to unskilled labour. We thank Matthew Lindquist for his help with the labour supply data.

¹⁵We 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.

¹⁶This 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.

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
$0 \leq \delta \leq 1$	0.022	depreciation rate of capital	calibration
$0 < \beta < 1$	0.990	time discount factor	calibration
$\frac{1}{1-\gamma}$	1.429	consumption to leisure elasticity	calibration
$0 < \mu < 1$	0.415	weight attached to consumption in utility	calibration
$\sigma > 1$	2.000	coefficient of relative risk aversion	assumption
$\frac{1}{1-\rho}$	0.669	capital equipment to skilled labour elasticity	assumption
$\frac{1}{1-\varphi}$	1.669	capital equipment to unskilled labour elasticity	assumption
$0 < \lambda < 1$	0.515	composite input share of output	calibration
$0 < \nu < 1$	0.800	capital equipment share of composite input	calibration
$\psi_s > 0$	$1 \times e^{-5}$	asset holding cost for skilled agents	calibration
$\psi_u > 0$	$2.5 \times e^{-5}$	asset holding cost for unskilled agents	calibration

3.2.2 Tax-spending policy

We use two data sources for income taxation. We have first constructed annual data on effective average capital and labour income tax rates from ECFIN using Martinez-Mongay (2000) approach and extending the original dataset.¹⁷ However, this dataset does not provide information on the progressivity of the tax system. Therefore, we also use income tax data from the Piketty and Saez (2007) dataset, which reports annual data on income tax rates per income group (in quintals) for the period 1966-2001. We use this data to obtain a measure of the progressivity of the tax system. In particular, we construct 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 work as follows. Given that in practice a large part of capital income (e.g. firm profits and income from asset holdings) is taxed separately from household income, at a fixed tax rate, we assume for the purposes of calibration that $\tau_{s,t}^k = \tau_{u,t}^k$. We then normalise the three income tax rates

¹⁷In particular, we use the LITR and KITN rates for effective average labour and capital taxes respectively for 1970-2011. The exact data series required for this construction are described in Martinez-Mongay (2000).

constructed using the Piketty and Saez (2007) dataset so that their weighted average (using the relative population weights) equals the average of the effective labour income tax burden using the ECFIN data and Martinez-Mongay (2000) approach. Working like this we are able to broadly calibrate the model to tax rates that capture both the progressivity of household income taxation and the distribution of the tax burden between capital and labour income.

Regarding the labour income taxes, we assume that for $j = s, u, h$:

$$\log(\tau_{j,t+1}^l) = (1 - \rho_j^l) \log(\tau_j^l) + \rho_j^l \log(\tau_{j,t}^l) + \varepsilon_t^{\tau_j^l} \quad (17)$$

where $\varepsilon_t^{\tau_j^l} \sim Niid(0, \sigma_{\tau_j^l})$. We use the time-series of the normalised, income tax rates from the Piketty and Saez (2007) dataset, constructed as described above as proxies for $\tau_{j,t}^l$, $j = h, u, s$, respectively.¹⁸ Therefore, we set the constant terms in the AR processes described above for, τ_j^l , $j = h, u, s$, to be equal to the data averages for the normalised tax rate for the respective income quintals, i.e. for τ_{0-20} , τ_{20-60} and τ_{60-100} . Moreover, we calculate ρ_j^l and σ_j^l by estimating the AR(1) processes in (17), 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.

Regarding the capital income taxes, we assume that for $\tau_{s,t}^k = \tau_{u,t}^k \equiv \tau_t^k$:

$$\log(\tau_{t+1}^k) = (1 - \rho^k) \log(\tau_t^k) + \rho^k \log(\tau_t^k) + \varepsilon_t^{\tau_t^k} \quad (18)$$

where $\varepsilon_t^{\tau_t^k} \sim Niid(0, \sigma_{\tau^k})$. We use the time-series for effective capital tax rates, constructed following Martinez-Mongay (2000) approach, and working as above for the labour income tax rates we estimate ρ^k and σ_{τ^k} .

The government spending series is obtained using quarterly data from the BEA for the period 1979 to 2002.¹⁹ Following the same procedure as with the tax rates, we calculate ρ_G and σ_G by estimating the AR(1) processes in (12), using the cyclical component of the public spending series.

3.2.3 Production, asset and labour markets

The elasticities of substitution between skilled labour and capital and between unskilled labour and capital (or skilled labour) have been estimated

¹⁸These 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 an indicator variable the quarterly time-series of labour income tax rates from Arseneau and Chugh (2012).

¹⁹This series refers to government consumption expenditures and gross investment as it is reported in NIPA Table 1.1.5.

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 U.S. data for the period 1979-2003.²⁰ The share of labour income in GDP is obtained from BEA data on personal income for the period 1970-2011.²¹ The calibrated parameters in the production function are generally very similar to those estimated or calibrated in the literature.

The depreciation rate of capital is set as $\delta = 0.022$ to obtain an investment to output ratio of about 15%. This is within the range of Krusell *et al.* (2000), Lindquist (2004) and Pourpourides (2011).

We set the asset holding cost parameters as $\psi_s = 1 \times e^{-5}$ and $\psi_u = 2.5 \times e^{-5}$. 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 U.S. Census²² 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 2011²³.

3.2.4 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, $\mu = 0.415$ is set so that the investment to consumption ratio from the model is close to the data average of 24%.²⁴ Moreover, $\gamma = 0.30$ so that

²⁰Using the data in Acemoglu and Autor (2011) gives a similar value for the average skill premium, namely 1.6.

²¹The data are obtained from the BEA, Table 2.1, Personal Income and Its Disposition.

²²The 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.

²³The 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).

²⁴The data series used for investment and consumption are obtained from the BEA, NIPA Table 1.1.5 for the period 1979-2011.

the households devote about 33% of their time to work. For the base results below, we set 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 opportunity and not in preferences.

Table 3A: Stochastic processes

Parameter	Value	Definition	Source
σ_A	0.006	std. dev. of TFP	calibration
ρ_A	0.950	AR(1) coef. of TFP	assumption
A	1.000	mean of TFP	set
σ_{A^e}	0.008	std. dev. of inv.-specific tech. change	calibration
ρ_{A^e}	0.975	AR(1) coef. of inv.-specific tech. change	calibration
A^e	1.000	mean of inv.-specific tech. change	set
$\sigma_{\tau_s^l}$	0.002	std. dev. of labour income tax, skilled	data
$\rho_{\tau_s^l}$	0.870	AR(1) coef. of labour income tax, skilled	data
τ_s^l	0.260	mean of labour income tax, skilled	data
$\sigma_{\tau_u^l}$	0.002	std. dev. of labour income tax, unskilled	data
$\rho_{\tau_u^l}$	0.900	AR(1) coef. of labour income tax, unskilled	data
τ_u^l	0.210	mean of labour income tax, unskilled	data
$\sigma_{\tau_h^l}$	0.002	std. dev. of labour income tax, hand-to-mouth	data
$\rho_{\tau_h^l}$	0.940	AR(1) coef. of labour income tax, hand-to-mouth	data
τ_h^l	0.170	mean of labour income tax, hand-to-mouth	data
σ_{τ^k}	0.006	std. dev. of capital income tax	data
ρ_{τ^k}	0.900	AR(1) coef. of capital income tax	data
τ^k	0.310	mean of capital income tax	data

3.2.5 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).

Table 3B: Additional stochastic processes (optimal policy)

Parameter	Value	Definition	Source
σ_G	0.009	std. dev. of public spending	data
ρ_G	0.770	AR(1) coef. of public spending	data
G	0.142	mean of public spending	calibration

3.3 Solution and results

The deterministic 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 and income tax realizations, that are obtained from the distributions specified above (see Table 3A). 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	0.353	0.359
$\frac{I}{Y}$	0.141	0.159	h_u	0.322	-
$\frac{C}{Y}$	0.604	0.659	h_h	0.351	-
r_{net}	0.010	0.010	\tilde{h}_u	0.331	0.348
$\frac{w_s}{w_u}$	1.659	1.659	$\frac{wh}{Y}$	0.698	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.²⁵ Starting with the steady-state, Table 4

²⁵Note 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) Lindquist (2004) for hours worked (obtained by dividing hours worked per week by 7*16 hours available for work and leisure); (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.

shows that the model predictions are quantitatively similar to the long-run averages in the data.

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 work hours for skilled and unskilled workers that are similar to the data. Note that \tilde{h}_u is average work hours by households providing unskilled labour in our model. Table 4 also shows that the labour's share of income in the model, $\frac{w_s n_s h_s + w_u n_u h_u + w_h n_h h_h}{Y} = 0.698$ 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 1.39 for skilled, 1.45 for unskilled and 1.27 for hand-to-mouth 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, the transaction costs are calibrated to match wealth inequality and capital as a share of GDP, and indeed the model predicts $K_s/K_u = 2.5$. Moreover, government spending is residually determined given the data averages for the tax rates and the model's steady-state predicts a value for total tax revenue as a share of GDP of about 25%, which is consistent with the U.S. data (see e.g. summary statistics and data sources in Hindriks and Myles (2013, ch. 4)).

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 and with the predictions from the dynamic general equilibrium models in Lindquist (2004) and Pourpourides (2011). 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
$\hat{\rho}(X_i, Y)$	1	0.91	0.84	0.18	0.94	0.72	0.98
$\hat{\sigma}(X_i)$	0.014	0.007	0.051	0.009	0.004	0.005	0.002

4 Optimal distribution of tax burden

We now discard the exogenous processes for the income tax rates in (17) and (18) and instead assume that the processes for these tax rates are optimally chosen by a government that seeks to maximise a Utilitarian objective function under commitment. Government spending follows the exogenous process in (12) with the parameters given in Table 3B.

4.1 The problem of the government

The government chooses state-dependent tax rates to maximise aggregate welfare subject to the optimality conditions of households and firms and the government budget constraint.²⁶ The government can choose different tax rates for each different source of income, implying that it can choose an effective income tax rate for each household, i.e. for each income group in the economy. Different responses between the taxes after exogenous technology and revenue requirement shocks imply a redistribution of the tax burden over the business cycle. The government 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. Thus, its objective function is given by the expected lifetime utility of the weighted average of the welfare of the three types of households:

$$U^g = E \sum_{j=s,u,h} \left[n_j \sum_{t=0}^{\infty} \beta^t u(C_{j,t}, l_{j,t}) \right] \quad (19)$$

²⁶In 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 heterogeneous households and differences in the effective income taxation between 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). We assume that the government has access to a commitment technology and chooses processes for $\{\tau_{s,t}^l, \tau_{u,t}^l, \tau_{h,t}^l, C_{s,t}, C_{u,t}, C_{h,t}, h_{s,t}, h_{u,t}, h_{h,t}, K_{s,t+1}, K_{u,t+1}, I_{s,t}, I_{u,t}, w_{s,t}, w_{u,t}, r_t\}_{t=0}^\infty$ and $\{\tau_{s,t}^k, \tau_{u,t}^k\}_{t=1}^\infty$ ²⁷ that satisfy the budget constraints (7), (8), (11), market clearing (13)-(15), and optimality conditions (A1)-(A8) in the Appendix.²⁸ As is common in optimal policy problems, we assume that the first-order conditions from this problem are necessary and sufficient, and all allocations are interior (see e.g. Arseneau and Chugh (2012)). To preserve space, we only discuss the first-order conditions for the policy instruments (income taxes). The equilibrium conditions for the policy instruments encapsulate key incentives that the Utilitarian government needs to consider when setting taxes 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 conditions for $\tau_{s,t}^l$ and $\tau_{s,t}^k$ are given by:

$$\frac{\partial U^g}{\partial \tau_{s,t}^l} = \lambda_t^{gbc} n_s h_{s,t} w_{s,t} - \lambda_t^{bcs} h_{s,t} w_{s,t} - \lambda_t^{lss} U_{c_s}(t) w_{s,t} = 0 \quad (20)$$

$$\frac{\partial U^g}{\partial \tau_{s,t}^k} = \lambda_t^{gbc} n_s K_{s,t} r_t - \lambda_t^{bcs} K_{s,t} r_t - \lambda_{t-1}^{kss} U_{c_s}(t) r_t = 0 \quad (21)$$

where λ_t^{lss} , λ_{t-1}^{kss} , λ_t^{bcs} and λ_t^{gbc} refer to the Lagrange multipliers attached to the skilled household's FOCs for hours worked and capital, the household's budget constraint and the government budget constraint respectively.²⁹ The

²⁷Note that, as is typically the case in problems with optimal capital taxation, the government is not allowed to choose the capital tax for period 0.

²⁸Note that the constraints do not include a present value implementability constraint, which is the typical approach in Ramsey problems with debt; instead, the government budget constraint is imposed in each period.

²⁹Note 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 these conditions.

first-order condition for the choice of taxes dictate that these should be chosen so that the net marginal social benefits equals the marginal social costs. For instance, regarding $\tau_{s,t}^l$, the net marginal benefits, MB^s , arising from an increase in this tax are given by: $\lambda_t^{gbc} n_s h_{s,t} w_{s,t} - \lambda_t^{bcs} h_{s,t} w_{s,t}$. 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: $\lambda_t^{lss} U_{c_s}(t) w_{s,t}$ and capture the distortions in labour supply, associated with the fall in the marginal net returns to work time. 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} . Similar arguments apply to the costs and benefits associated with $\tau_{s,t}^k$, where now the costs are associated with the distortion of the investment decision of the skilled household.

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

$$\frac{\partial U^g}{\partial \tau_{u,t}^l} = \lambda_t^{gbc} n_u h_{u,t} w_{u,t} - \lambda_t^{bcu} h_{u,t} w_{u,t} - \lambda_t^{lsu} U_{c_u}(t) w_{u,t} = 0 \quad (22)$$

$$\frac{\partial U^g}{\partial \tau_{u,t}^k} = \lambda_t^{gbc} n_u K_{u,t} r_t - \lambda_t^{bcu} K_{u,t} r_t - \lambda_{t-1}^{ku} U_{c_u}(t) r_t = 0 \quad (23)$$

$$\frac{\partial U^g}{\partial \tau_{h,t}^l} = \lambda_t^{gbc} n_h h_{h,t} w_{u,t} - \lambda_t^{bch} h_{h,t} w_{u,t} - \lambda_t^{lsh} U_{c_h}(t) w_{u,t} = 0 \quad (24)$$

where λ_t^{lsu} , λ_{t-1}^{ku} , and λ_t^{bcu} refer to the Lagrange multipliers attached to the unskilled household's FOCs for hours worked and capital 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 first-order conditions follows the same reasoning as for the conditions for the labour and capital income taxes applied to the skilled household.

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, as well as between the taxes, given heterogeneity in opportunities, so that a quantitative evaluation is required to determine optimal taxes. In particular, the net marginal benefit described above 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 for aggregate productivity are stronger for taxes on skilled households, followed by unskilled and hand-to-mouth.

4.1.2 Trade-off for optimal policy

The literature on optimal taxation suggests that capital taxes should be smaller on average and more volatile over the business cycle, compared with labour income taxes (see e.g. Chari *et al.* (1994) and Ljungqvist and Sargent (2012)). Optimal policy in our framework requires the government to decide not only how to set capital versus labour income taxes over the business cycle, but also how to distinguish capital and labour income taxes for each factor supply and (implicitly) income level. In particular, the government needs to decide if all labour income taxes should have similar stochastic properties, in terms of volatility and cyclical, or, if not, how they should differ across households; and similarly for the two capital income taxes. The barriers and transaction costs 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 taxes over the business cycle, the government must evaluate a key trade-off, which follows from the previous sub-section. On one hand, unfavourable income taxes, in the form of higher level, volatility and counter-cyclical, 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 size of the tax base to which they apply. Therefore, the government has an incentive, for both reasons, to keep the taxes that apply to lower income, lower wealth households 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 due to the complementarity between capital and skill and because the high income skilled households supply more capital than either of the unskilled households. Thus, taxes applied to skilled households imply higher distortions at the aggregate level than the other taxes. This creates an incentive for the government to smooth taxes for the skilled household and to set their taxes to be the least counter-cyclical. By extension, 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.

The analysis of the results below elaborates on this trade-off. In subsections 4.4 - 4.7, we further investigate the strength of the channels embedded in the model and relate these to this trade-off work by considering appropriate model variants and changes in related parameters.

4.1.3 Solution

To numerically solve for the outcomes under optimal policy, we follow e.g. Arseneau and Chugh (2012) in that we first compute the deterministic steady-state under optimal policy and then approximate the equilibrium conditions under optimal policy around the deterministic steady-state of these conditions.³⁰ 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 to employ a second-order approximation. This is because the effects of the stochastic environment on households depend on their ability to smooth consumption, thus affecting differently households with differential access to asset markets.

Moreover, in our setup, since there is no government debt, to solve for the deterministic steady-state we shut down shocks and drop time subscripts in the first-order equations arising from the government's problem (19), and then solve the non-linear system of equations using numerical methods. The asymptotic non-stochastic equilibrium is independent of initial conditions and can thus be solved for independently of the dynamic paths.³¹ For this

³⁰Hence, as also discussed in e.g. Arseneau and Chugh (2012), this implies the auxiliary assumption that when calculating asymptotic policy dynamics the initial state is the asymptotic steady state associated with optimal policy.

³¹This property of the model is particularly useful in our setup, because the existence of two types of households with access to the asset market would imply two present value implementability constraints if the government could issue debt as in the unrestricted Ramsey problem. In turn, this implies that solving for the optimal dynamic paths from period 0 to the steady state is a complex problem.

property of the model, the transaction costs in capital ownership at the household level are also important. In particular, the presence of transaction costs means that the two Euler equations remain different in the non-stochastic steady-state even if $\tau_s^k = \tau_u^k$, in turn implying that the asymptotic equilibrium can be solved without requiring additional restrictions (e.g. in the form of initial conditions). Instead, if the transaction cost parameters are set to zero, the two Euler equations in Appendix A collapse to one in the deterministic steady-state if $\tau_s^k = \tau_u^k$ and thus the latter cannot be obtained without additional restrictions. Thus, the setup does not rule out $\tau_s^k = \tau_u^k$ as a potential optimal policy outcome given the computational strategy adopted.³²

4.2 Optimal policy in the long-run

The optimal tax rates in the non-stochastic steady-state are presented in Table 6.

Table 6: Optimal policy in the steady-state

	τ_s^l	τ_u^l	τ_h^l	τ_s^k	τ_u^k	τ_s^e	τ_u^e
Steady-state	0.510	0.116	0.042	0.005	0.001	0.371	0.074

As can be seen, optimal capital income taxes are near zero, consistent with the literature since Chamley (1986) and Judd (1985). They are not exactly zero, as in the frictionless neoclassical model, because of the intermediation costs in holding capital, which create a resource implication for the aggregate economy that the government needs to take into account when choosing taxes optimally.³³ The capital income tax for skilled households is higher than that for unskilled, however, quantitatively, optimal long-run capital taxes are small.

The small capital income taxes imply that to generate the additional tax revenue required, the government needs to employ non-zero labour income taxes, which distort the labour supply decisions of the households. Optimal labour income taxation is progressive in the long-run and ranges from about 50% for skilled households to about 5% for hand-to-mouth households. The effective income taxes per income group, calculated as total tax revenue collected per household in the group over the total income of the household, suggest that the skilled (higher income group) would optimally pay about 37% of their total income in taxes, whereas the unskilled households 7.4% and the hand-to-mouth 4.2%.

³²In fact, note that $\tau_s^k = \tau_u^k$ was assumed in order to solve the model under exogenous policy.

³³See, for example, Guo and Lansing (1999) and Judd (1997) for previous examples of non-zero capital taxes in the long-run that arise because of frictions in the capital market.

The results in Table 6 imply that the government increases aggregate welfare in this framework, when using distorting taxes, by redistributing 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.

Total income taxation is progressive in the steady-state under the market exclusions assumed in this model. It should be noted that if these market exclusions were relaxed, for instance by allowing unskilled workers to educate and become skilled, optimal policy would also need to take the skill-creation incentives into account, which could have important implications for optimal tax progressivity. However, focusing on business cycle frequencies, where the proportions of skilled and unskilled workers do not vary significantly, as noted earlier, we assume away such movements between these skill groups.

4.3 Optimal taxes over the business cycle

We next compute fluctuations around the deterministic optimal steady-state by simulating the optimal-policy equilibrium under the exogenous processes to TFP, investment-specific technological change and government spending described in Section 3.³⁴ We present 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 cyclicity of taxes: base results

Consistent with the results in e.g. Chari *et al.* (1994) and Stockman (2001), the capital income taxes are more volatile over the business cycle, compared with labour income taxes. This is because capital taxes apply to a tax base that is pre-determined when the exogenous shocks are realised in each period, which creates efficiency incentives for a higher response of the capital income

³⁴To 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.

taxes.³⁵ The relative magnitude between capital and labour tax standard deviations in Table 7 is comparable to the results reported in Stockman (2001) under a balanced budget restriction.³⁶ Moreover, and again consistent with the literature, capital taxes fluctuate around zero over time, as the confidence intervals in Table 7 indicate. The confidence intervals for the labour income taxes are also non-trivial and imply that 95% of the time labour income taxes fluctuate between about 2 for (τ_s^l) and 4 (for τ_u^l) percentage points.³⁷

Table 7: Optimal tax policy for all shocks

	$\mu(X_i)$	$CI(X_i)$	$\frac{\sigma(X_i)}{\sigma(Y)}$	$\rho(X_{i,t}, X_{i,t-1})$	$\rho(X_i, Y)$	$CI(\rho(X_i, Y))$
τ_s^l	0.510	0.501, 0.520	0.358	0.903	0.631	0.493, 0.738
τ_u^l	0.116	0.096, 0.135	0.723	0.938	0.149	-0.053, 0.340
τ_h^l	0.042	0.024, 0.060	0.664	0.950	-0.197	-0.382, 0.004
τ_s^k	0.005	-0.061, 0.071	2.459	0.780	-0.776	-0.845, -0.682
τ_u^k	0.001	-0.075, 0.077	2.825	0.630	-0.677	-0.773, -0.551
τ_s^e	0.371	0.359, 0.383	0.453	0.659	-0.698	-0.788, -0.579
τ_u^e	0.074	0.050, 0.099	0.905	0.587	-0.686	-0.779, -0.563
τ_h^e	0.042	0.024, 0.060	0.664	0.950	-0.197	-0.382, 0.004

Note: the confidence intervals, CI, are at the 95% level here and in the tables below.

The main results in Table 7 pertain to the differences in the volatility between the capital taxes and between the labour income taxes as well as between the effective income taxes that apply to different households and income levels. Regarding capital taxes, the rate levied on the unskilled households is more volatile than the rate on the skilled households. This relative difference is even greater when comparing the volatility of labour income taxes for unskilled (middle-income) and skilled households, as τ_u^l is twice more volatile than τ_s^l . The labour income tax rate to hand-to-mouth households is more volatile than the labour income tax to skilled households

³⁵See also Fahri (2010) and Gervais and Mennumi (2015) on optimal capital taxation.

³⁶This is also confirmed when we calculate results under higher risk aversion results in later sections. For instance, for a very similar high degree of risk aversion (see our results for $\sigma = 8$ below) both our model and Stockman (2001) under a balanced budget constraint predict that optimally the ratio of standard deviations of capital to labour income taxes is about 10. In contrast, note that without the balanced budget restriction, Stockman (2001) finds that such ratios exceed 100.

³⁷The comparison with previous results in the literature here focused on qualitative properties and relative magnitudes. Quantitatively, the results in Table 7 imply different levels of standard deviations for capital and labour taxes compared with e.g. Stockman (2001). As discussed earlier, the models differ in many dimensions. Moreover, the stochastic environment captured by the simulation exercise is also different. Therefore, it is difficult to compare the effect of exogenous stochasticity on the endogenous quantities and thus on the magnitude of the response required by the government.

but less volatile than the labour income tax to unskilled households that can smooth consumption. Moreover, the government chooses counter-cyclical capital income taxation and sets effectively a-cyclical labour income taxes for hand-to-mouth and unskilled households, and pro-cyclical labour income taxes for skilled workers.³⁸

The properties of the effective income taxes are important because they encapsulate the overall effect of taxation on household income. The outcome of the effects on capital and labour income taxes is such that the effective income tax rate applying to the middle-income group is twice more volatile than the effective income tax rate applying to high income households and about 50% more volatile compared to the effective income tax rate applying to hand-to-mouth (note that this effective rate is equal to τ_h^l). Moreover, effective income taxes are counter-cyclical for the skilled and unskilled groups and a-cyclical for the hand-to-mouth households. Hence, although in terms of labour income the tax to the skilled, high-income households is pro-cyclical, the counter-cyclicity of the capital income taxes implies that, overall, these households face counter-cyclical effective income taxes.

The quantitative resolution of the trade-off faced by the policy maker when deciding how to distribute the distortions reflected by the higher volatility and counter-cyclicity of the tax rates over the business cycle leads to the optimal income taxes reported in Table 7. On one hand, tax distortions have a larger impact on hand-to-mouth households, since they have lower income and are not able to smooth shocks over time. There is thus an incentive to minimise the impact of policy for this type of household. On the other hand, tax-induced distortions to skilled households have the strongest propagation effects in the economy, given the complementarity of skilled hours with equipment capital. Therefore, there is also an incentive to minimise distortions to the choices of skilled households, since such distortions act to amplify external shocks.

As a result, the government finds it optimal to respond to the public finance implications of technology and expenditure shocks by setting taxes to the middle income group of unskilled households (with savings) to be the least smooth. For the high income group, it is the tax to skilled labour supply is the least volatile of all taxes in the economy, highlighting the key role played by skilled labour in this model. On the other hand, counter-cyclicity of effective income taxation is monotonic with income, since the lower income households face the higher marginal utility losses from an increase in income taxation in bad times. For the higher income households, however, the gov-

³⁸Note that the confidence intervals for the correlation coefficients were calculated using the Fisher (1915) transformation.

ernment chooses to generate this counter-cyclical with respect to effective income taxation by making their capital taxation strongly counter-cyclical. In turn, the strong counter-cyclical of capital income taxation implies that labour income taxation for the households who own the capital stock needs to be set in a less counter-cyclical fashion, to compensate them for their capital income losses.

The trade-off that the government faces in setting taxes for the different households will be further analysed in the next sub-section, when we look at the effects of each shock in isolation. This also leads us to address directly the question of how the government optimally redistributes the tax burden between the income groups in bad times. Following this analysis, to further elaborate on the workings of the key channels underpinning the trade-off discussed above, we will also investigate the importance of skill heterogeneity, capital-skill complementarity and severity of limited consumption smoothing.

4.3.2 Optimal responses to productivity and spending shocks

We next evaluate how the government optimally changes the tax rates in the short- and medium-run by examining the impulse responses (IRs) of the optimal tax rates after a temporary, output-reducing standard deviation shock to each of the exogenous processes. These are plotted in Figure 1 as deviations (in levels) from the steady-state. Moreover, in Figure 2, we plot the responses of the effective income tax rates implied by the capital and labour income tax movement in Figure 1. To complement the impulse response analysis, Table 8 presents the optimal properties of the tax system under separate sources of economic fluctuations.

Productivity shocks Figure 1 shows that in response to negative shocks to A and A^e , the labour taxes are optimally reduced in the short-run, to stimulate labour supply and support the income of the households, while capital taxes are increased to make up for the loss in tax revenue. As in Stockman (2001), the dynamic response of the capital taxes displays a more gradual adjustment to the steady-state, instead of a very short-lived spike after the shock, which is typically associated with optimal capital taxation with debt. Given that the government cannot use debt to absorb the revenue implications of quick reductions to capital taxes, the latter return to their steady-state values more slowly.

[Figures 1 and 2 here]

The labour taxes increase above their pre-shock levels in the medium-run before returning to the steady-state. The dynamic patterns are generally

symmetric for the respective taxes for all households. However, as seen in Table 8, the relative standard deviation of τ_s^l is lower and its correlation with the technology shocks is the most strongly positive. Both characteristics capture the incentives of the government to use this tax in the least distorting way over the business cycle. As pointed out earlier, skilled labour supply works more as an amplification mechanism in response to exogenous shocks, given the stronger complementarity it has with capital. For efficiency reasons, the capital taxes are more volatile than labour taxes and skilled labour supply is supported by the least volatile and most pro-cyclical of the taxes.

Table 8: Optimal tax policy under individual shocks

	τ_s^l	τ_u^l	τ_h^l	τ_s^k	τ_u^k	τ_s^e	τ_u^e	τ_h^e
<i>A</i> shock								
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.326	0.558	0.519	2.147	2.186	0.357	0.636	0.519
$\rho(X_i, Y)$	0.749	0.293	-0.143*	-0.937	-0.926	-0.959	-0.991	-0.143*
$\rho(X_i, A)$	0.894	0.523	0.110	-0.993	-0.990	-0.996	-0.933	0.110
<i>A^e</i> shock								
$\frac{\sigma(X_i)}{\sigma(Y)}$	0.532	1.191	1.426	3.404	3.447	0.659	1.383	1.426
$\rho(X_i, Y)$	0.206	-0.324	-0.402	-0.426	-0.481	-0.253	-0.635	-0.402
$\rho(X_i, A^e)$	0.789	0.363	0.299	-0.805	-0.530	-0.588	-0.286	0.299
<i>G</i> shock								
$\frac{\sigma(X_i)}{\sigma(Y)}$	1.000	1.500	2.500	26.25	51.75	6.500	18.00	2.500
$\rho(X_i, Y)$	-0.645	-0.922	-0.174*	0.228	0.577	0.201	0.549	-0.174*
$\rho(X_i, G)$	-0.981	-0.775	0.721	0.747	0.892	0.689	0.887	0.721

* indicates that the 95% confidence interval for the correlation coefficient includes zero.

Note that the special treatment of skilled labour supply is especially pronounced under investment-specific shocks, which affect capital directly. Thus, the need to use skilled labour to counteract these negative productivity effects is stronger. As can be seen in Table 8, the relative volatility of τ_s^l is significantly smaller than the remaining taxes in this case. Note that this is the only tax that is pro-cyclical and also is the most strongly correlated with the investment specific process.

The overall income implications of the responses of capital and labour taxes are summarised by the responses of the effective income taxes, plotted in Figure 2. The optimal responses of the effective income taxes after exogenous negative productivity shocks imply that there is a short-run reduction in the rate for hand-to-mouth households, whereas the effective tax rate on the remaining households is increased to meet the financing requirements of the government. These dynamic patterns capture the equity incentives of the government and, in particular, the social benefits by supporting the income

of the poorest households, since a marginal decrease of income for the hand-to-mouth households carries a higher social cost. As seen in Figure 1, the increase in τ_s^e and τ_u^e are driven by increases in the capital taxes, as labour income taxes for households with asset holdings decrease in the short-run, following negative productivity shocks.

However, note that under investment-specific shocks the hand-to-mouth tax is more volatile than the remaining effective taxes and has comparable counter-cyclicalities. Under more neutral productivity shocks, there are significant differences in counter cyclicalities in favour of hand-to-mouth, even though the volatility of τ_h^l is comparable to that of τ_s^e and τ_u^e . Therefore, although some protection is afforded to the poorer households under negative productivity shocks, this cover is limited. The reason is that such shocks have direct negative aggregate effects and the tax system needs to respond to the aggregate productivity losses.

Spending shocks The responses of the taxes are not the same under fiscal expenditure shocks. In particular, as can be seen in the third row of Figure 1, the government increases capital taxes as well as the labour tax rate to hand-to-mouth households to increase tax revenue in response to increased fiscal spending. Moreover, it decreases the labour income taxes to the remaining households in the short-run but τ_s^l and τ_u^l are increased above the pre-shock levels in later periods, and in fact the increase is faster compared with the increase in these taxes in the medium-run following the negative productivity shocks. The effective tax rates in Figure 2 display a more symmetric response across households in this case, as these tax rates follow more similar dynamic patterns for all households. These patterns are consistent with the correlations between the various income taxes and the fiscal spending shock in Table 8.

Compared with negative technology shocks, which imply a drop in productivity and thus to returns to labour and capital, positive fiscal spending shocks do not directly affect productivity and thus do not directly change the income of the households. Hence, in choosing the taxes, the government does not need to address a productivity shock that has aggregate as well as redistributive implications on its own, but rather seeks to find a way to minimise the efficiency and equity implications that are created directly by the requirement to finance spending policy shocks. Efficiency considerations imply a stronger reaction of capital relative to labour taxes (as explained earlier) and this stronger reaction is pronounced here, relative to the responses after negative productivity shocks. This is because the negative productivity shocks act as an adverse shock to capital accumulation in the first place.

Thus, creating an incentive for the government to support capital accumulation, which works in the opposite direction from the incentive to tax capital in the short-run for efficiency purposes. Accordingly, under government spending shocks, the difference in the volatility of labour and capital taxes is more pronounced. In addition, efficiency requires that the increase in capital taxes is higher proportionately for the investors who face the higher transaction costs, since it is less efficient at the aggregate level for them to hold assets. Finally, efficiency considerations imply that the skilled labour supply needs to be supported. Hence, creating an incentive to treat more favorably labour income taxation for skilled workers.

At the same time, equity considerations imply an incentive to keep differences between the changes in effective taxation across households relatively small. This implies that given the big increases in capital income taxes, the labour income taxes for the households with assets are reduced in the short-run, and the labour income tax for hand-to-mouth households is increased. As capital income taxes return to their steady-state value, the labour income taxes for households with assets are increased. However, note that again the increase in the effective tax to hand-to-mouth households is smaller proportionately compared with the increase in effective taxes to the remaining households. Indeed, τ_h^l is the least volatile of the effective taxes, although its correlation with the shock is comparable in magnitude.

4.3.3 The re-distribution of the relative tax burden in bad times

The previous analysis first shows that, as the economy fluctuates over time, a Utilitarian government would choose tax rates that are more volatile and less pro-cyclical for the middle income households, relative to the other household types. This implies that, in business cycle frequencies, the second moments of the tax rates are set least favorably for this income group. We examine next the redistribution of the relative tax burden between the three income groups in response to negative shocks associated with lower productivity and increased expenditure requirements. We can identify the direction of this redistribution and quantify it by calculating the change in the share of the total tax revenue paid by each of the three income groups. This is shown in Figure 3 and Table 9 for each of the three shocks examined in the previous sub-section.

[Figure 3 here]

As can be seen in Figure 3, following negative shocks, the contribution of tax revenue to total tax revenue is reduced for the group of high-income, skilled households and increased for that of middle-income, unskilled households, for all types of shocks. Regarding the group of hand-to-mouth, low

income households, their relative contribution is initially reduced and then increased following negative productivity shocks, while it is increased following increased government spending shocks.

Table 9: Distribution of the tax burden

	skilled	unskilled	hand-to-mouth
shock	sum of the change in shares		
A	-0.2485	0.1890	0.0595
A^e	-0.2675	0.1953	0.0723
G	-0.0556	0.0517	0.0039
	discounted sum of the change in shares		
A	-0.1672	0.1321	0.0351
A^e	-0.1500	0.1155	0.0345
G	-0.0329	0.0298	0.0031

In Table 9 above, we calculate the sum of the deviations of the relative tax revenue contributions from the steady-state ratio over time. The results confirm that the share of total tax revenue paid by the middle-income households is increased by most, followed by a rise in the share for hand-to-mouth that is about half of that for the unskilled, while the relative tax burden falls for the skilled households. The results are similar across the three types of shocks. As Table 9 shows, these results are robust to calculating discounted sums (second part of the Table).

4.4 Importance of skill heterogeneity and capital-skill complementarity

We next further investigate the importance of labour market heterogeneity and capital-skill complementarity in relation to the main results in Table 7. We first consider a version of the model where the production function is a standard Cobb-Douglas form with capital and two types of labour input, identified by exogenous productivity parameters that apply to skilled and unskilled workers.³⁹ These parameters are calibrated so that the model generates the same steady-state skill premium as in the base model. The remaining parameters are calibrated following the relevant, where appropriate, steps described in Section 3, to provide a steady-state that is effectively the same for the comparable aggregate quantities. We then obtain optimal taxes

³⁹Hence labour input is a weighted average of the labour hours of skilled and unskilled labour, where skilled and unskilled are differentiated by exogenous labour productivity parameters.

and their statistics over the business cycle, working as above, and report the results in the upper half of Table 10.

Changing the production function reduces the strength of the link between skilled labour and capital, since the relationship between different types of labour and capital is symmetric in the production function. However, this change does not eliminate the relative importance of skilled hours for production. This is because the higher productivity of the labour input of the skilled worker implies that this input has a stronger effect on the marginal product of capital in the economy. Therefore, the new production function reduces but does not eliminate the strength of the propagation effects associated with skilled labour and the importance of this for optimal taxation is seen in the upper half of Table 10. In particular, the gap between the volatility of the taxes applying to the skilled worker and the unskilled workers is clearly reduced, compared with the results in Table 7, although it is not totally eliminated. In terms of effective tax rates, the middle-income unskilled households still face the most volatile tax, but the skilled households now face a clearly more counter-cyclical tax than the other households, and the second most volatile effective tax. Hence, certain elements of the protection afforded to skilled workers have been reduced.

To further investigate the importance of the privileged position of skilled workers in production, we take the previous experiment one step further and next eliminate differences in labour productivity. In particular, we examine optimal policy results in a version of the base model without labour market heterogeneity and without capital-skill complementarity, by assuming that all types of households offer identical labour services so that there is a single wage rate. We also employ a Cobb-Douglas production function with one type of capital and one type of labour as inputs. Therefore, the only dimension in which the households differ is in terms of their participation in the capital market. The results for the optimal taxes and their statistics over the business cycle, working as above, are reported in the lower half of Table 10.

This experiment allows us to confirm the importance of labour market exclusions for the results obtained earlier under capital-skill complementarity. As can be seen, τ_s^l is now the most volatile of the labour income taxes, since the high income households do not enjoy a privileged position in the labour market. Their labour input is comparable to that of the remaining households in terms of its implications for aggregate productivity and the propagation of productivity shocks. Moreover, the volatility of income taxation is now increasing with the wealth of the households, similar to the counter-cyclicity results. In particular, note that while in terms of effective taxation, counter-cyclicity is increasing with income both in Table 7 and in Table 10, it is only when the labour supply of the three households has the same productivity

that the volatility of effective income taxation increases with income.⁴⁰ In terms of the trade-off that the government faces, only one side of this is relevant here. In particular, the need to protect the income of the less well-off households in the face of exogenous fluctuations which gives rise to the monotonic results summarised in the lower half of Table 10.

Table 10: Optimal tax policy without capital-skill complementarity

	$\mu(X_i)$	$CI(X_i)$	$\frac{\sigma(X_i)}{\sigma(Y)}$	$\rho(X_{i,t}, X_{i,t-1})$	$\rho(X_i, Y)$	$CI(\rho(X_i, Y))$
Cobb-Douglas case with wage inequality						
τ_s^l	0.493	0.483, 0.503	0.185	0.894	0.623	0.481, 0.731
τ_u^l	0.194	0.181, 0.207	0.241	0.922	0.246	0.048, 0.425
τ_h^l	0.118	0.103, 0.133	0.274	0.915	0.141	-0.061, 0.332
τ_s^k	0.010	-0.094, 0.114	1.970	0.297	-0.500	-0.637, -0.333
τ_u^k	0.003	-0.122, 0.128	2.356	0.089	-0.237	-0.417, -0.038
τ_s^e	0.327	0.295, 0.359	0.604	0.203	-0.460	-0.605, -0.286
τ_u^e	0.149	0.106, 0.192	0.807	0.080	-0.186	-0.373, 0.015
τ_h^e	0.118	0.103, 0.133	0.274	0.915	0.141	-0.061, 0.332
Cobb-Douglas case without wage inequality						
τ_s^l	0.366	0.355, 0.378	0.331	0.900	0.527	0.365, 0.658
τ_u^l	0.300	0.293, 0.308	0.213	0.923	0.250	0.052, 0.429
τ_h^l	0.252	0.241, 0.263	0.309	0.912	0.159	-0.042, 0.349
τ_s^k	0.005	-0.051, 0.062	1.612	0.669	-0.619	-0.729, -0.477
τ_u^k	0.004	-0.087, 0.095	2.607	0.139	-0.398	-0.555, -0.215
τ_s^e	0.201	0.180, 0.223	0.618	0.558	-0.577	-0.697, -0.426
τ_u^e	0.229	0.208, 0.250	0.601	0.100	-0.371	-0.532, -0.184
τ_h^e	0.252	0.241, 0.263	0.309	0.912	0.159	-0.042, 0.349

4.5 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.⁴¹ This can be relevant in our setup, since hand-to-mouth households are particularly exposed to fluctuations in income. We

⁴⁰As with the results in Table 7, given the strong counter-cyclicality of capital income taxes, labour income taxes for households who own assets need to be set less counter-cyclically, to compensate them for the income loss sustained by the increase in capital taxation in bad times.

⁴¹Analyses 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)).

therefore present in Table 11 results for optimal taxes when $\sigma = 8$.⁴²

Table 11: Optimal tax policy with higher RRA (all shocks)

	$\mu(X_i)$	$CI(X_i)$	$\frac{\sigma(X_i)}{\sigma(Y)}$	$\rho(X_{i,t}, X_{i,t-1})$	$\rho(X_i, Y)$	$CI(\rho(X_i, Y))$
τ_s^l	0.607	0.596, 0.617	0.466	0.916	0.721	0.608, 0.805
τ_u^l	0.052	0.042, 0.063	0.474	0.950	0.160	-0.042, 0.350
τ_h^l	-0.097	-0.111, -0.084	0.603	0.939	0.055	-0.147, 0.253
τ_s^k	0.005	-0.106, 0.116	4.888	0.238	-0.398	-0.554, -0.215
τ_u^k	-0.0003	-0.140, 0.140	6.164	0.163	-0.444	-0.592, -0.267
τ_s^e	0.458	0.433, 0.483	1.095	0.012	-0.130	-0.322, 0.072
τ_u^e	0.029	-0.002, 0.061	1.388	0.115	-0.407	-0.561, -0.224
τ_h^e	-0.097	-0.111, -0.084	0.603	0.939	0.055	-0.147, 0.253

Given the higher costs attached to income fluctuations particularly for hand-to-mouth households, this calibration implies that these households experience optimally the least volatile effective income taxes, which are also a-cyclical. Moreover, the difference between volatilities of the labour income tax rates for skilled and unskilled households that have access to savings is reduced. However, the difference in the volatilities of capital taxes is increased, so that the overall difference in the relative standard deviation of τ_s^e and τ_u^e does not change significantly.

4.6 Importance of capital holding costs

The costs associated with holding assets in this setup distinguish the two types of households in the asset markets and allow the model to be calibrated to replicate observed asset market inequality. As explained in previous sections, if these costs are equal, then the Euler equations can be distinguished in the steady-state only if the capital taxes differ. To quantitatively evaluate the importance of reducing the distance in asset market inequality between skilled and unskilled workers, we next consider an experiment where we reduce the asset holding cost parameter for the unskilled households by 50%. The results are summarised in Table 12. As can be seen, this change affects the volatility attached to the two capital taxes in the expected way, i.e. it increases the standard deviation of τ_s^k relative to τ_u^k , compared with the base

⁴²Note that the higher risk aversion parameter implies that the model under exogenous policy predicts lower volatilities for consumption and labour supplies, compared with the data, while the correlation of consumption and the labour supplies with output are much higher compared to the data. In other words, it worsens the overall fit of the model relative to the benchmark case under $\sigma = 2$ presented in Section 3.4. Hence, although not calibrated to hit a specific target, the value of $\sigma = 2$ employed for the benchmark results is consistent with empirically relevant cyclical properties for the model.

results in Table 7, because the efficiency advantages of allocating the more distorting tax to unskilled capital income, as opposed to skilled, are significantly reduced. Other than this, however, the main results in Table 7 are not affected.

Table 12: Optimal tax policy, with lower ψ_u

	$\mu(X_i)$	$CI(X_i)$	$\frac{\sigma(X_i)}{\sigma(Y)}$	$\rho(X_{i,t}, X_{i,t-1})$	$\rho(X_i, Y)$	$CI(\rho(X_i, Y))$
τ_s^l	0.484	0.477, 0.492	0.283	0.908	0.574	0.424, 0.695
τ_u^l	0.146	0.127, 0.165	0.710	0.923	0.321	0.128, 0.490
τ_h^l	0.044	0.027, 0.061	0.645	0.947	-0.163	-0.352, 0.039
τ_s^k	0.004	-0.076, 0.084	2.964	0.398	-0.641	-0.746, -0.506
τ_u^k	0.001	-0.072, 0.073	2.688	0.776	-0.690	-0.782, -0.568
τ_s^e	0.379	0.365, 0.393	0.514	0.218	-0.506	-0.641, -0.340
τ_u^e	0.081	0.053, 0.108	1.022	0.699	-0.669	-0.766, -0.541
τ_h^e	0.044	0.027, 0.061	0.645	0.947	-0.163	-0.352, 0.039

4.7 Importance of government preferences

In our analysis we focused on the choices of a Utilitarian government that cares equally about all households in the economy. As we have seen, such a government in this setup would choose a tax system with higher progressivity in the steady-state compared to the actual data. In this section we consider instead the choices of a government that weighs the households in the social welfare function in such a way that the implied optimal effective taxes for the three income groups are the same with those in the exogenous policy equilibrium. We calibrated the weights in the social welfare function following this approach and obtained a weight attached to the skilled households of 0.568, a weight to the unskilled households with assets of 0.262 and a weight to hand-to-mouth households of 0.1. These weights imply that the policymaker is biased in favour of high-income, skilled households, and against the lower income households. The results for the choices of this policymaker over the business cycle are summarised in Table 13.

Given that the government choices now incorporate a preference bias in favour of the skilled workers that was pinned down by targeting the effective tax rates, it is not surprising that the volatility of τ_u^e relative to τ_s^e is even higher compared with the Utilitarian case in Table 7. Consistent with this, τ_u^e is also more counter-cyclical relative to τ_s^e . On the other hand, the properties of τ_h^l are not as much affected. Given that the importance of this group in generating tax revenue is small, it seems that there is no need for big changes

when this group receives a lower weight in the social welfare function.

Table 13: Non-Utilitarian tax policy

	$\mu(X_i)$	$CI(X_i)$	$\frac{\sigma(X_i)}{\sigma(Y)}$	$\rho(X_{i,t}, X_{i,t-1})$	$\rho(X_i, Y)$	$CI(\rho(X_i, Y))$
τ_s^l	0.404	0.391, 0.417	0.471	0.907	0.514	0.349, 0.648
τ_u^l	0.343	0.330, 0.356	0.478	0.915	0.392	0.207, 0.549
τ_h^l	0.170	0.154, 0.186	0.610	0.944	-0.154	-0.344, 0.048
τ_s^k	0.003	-0.093, 0.099	3.603	0.325	-0.427	-0.578, -0.248
τ_u^k	0.004	-0.022, 0.232	8.559	0.090	-0.359	-0.522, -0.171
τ_s^e	0.278	0.250, 0.306	1.051	0.153	-0.262	-0.439, -0.064
τ_u^e	0.240	0.174, 0.305	2.441	0.043	-0.352	-0.516, -0.163
τ_h^e	0.170	0.154, 0.186	0.610	0.944	-0.154	-0.344, 0.048

We can finally see in Table 13 that the difference in the volatility of τ_u^e relative to τ_s^e is now due exclusively to the capital income taxes, as the labour income taxes for these two groups have more similar properties. For the Utilitarian case in Table 7, the government cares more about the skilled workers relative to the unskilled indirectly, because of the key role that skilled labour has in production and thus is propagation of shocks and aggregate outcomes. This role works via skilled labour supply, which is why, in Table 7, it is the τ_u^l relative to τ_s^l relationship where most of the action in tax heterogeneity takes place. On the other hand, under the preference bias in Table 13, the government cares more about skilled versus unskilled workers directly, via the social welfare function. Hence the emphasis on supporting skilled labour supply for its benefits to the aggregate economy is reduced, and instead it is the total income effects of taxation (as captured by changes in τ_u^e and τ_s^e) that are targeted directly by the government.

5 Conclusions

This paper analysed optimal taxes 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.

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 income tax rates to

maximise aggregate welfare. We found that optimal taxes for middle-income households were more volatile than the remaining taxes. In addition, in bad times, the share of total tax revenue paid by middle-income households rose, followed by a significantly smaller increase for low-income households, while the relative tax share for skilled households fell.

The exclusions in asset and labour markets under capital-skill complementarity implied that the importance that skilled workers enjoyed in production and the vulnerability of hand-to-mouth consumers to income fluctuations were such that unfavourable changes in taxation to the middle income group were generally the least harmful to aggregate welfare over the business cycle.

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] Angelopoulos, K., Asimakopoulos, S., and J. Malley (2015). ‘Tax smoothing in a business cycle model with capital-skill complementarity’, *Journal of Economic Dynamics and Control*, 51, 420-444.
- [4] Angelopoulos, K., Fernandez, B. and J. Malley (2014). ‘The distributional consequences of tax reforms under capital-skill complementarity’, *Economica*, 81, 747-767.
- [5] Angelopoulos, K., Malley, J. and A. Philippopoulos (2017). ‘Human capital accumulation and transition to skilled employment’, *Journal of Human Capital*, 11, 72-105.
- [6] Arseneau, D. and S. Chugh (2012). ‘Tax smoothing in frictional labor markets’, *Journal of Political Economy*, 120, 926-985.
- [7] 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.
- [8] Benigno, P. (2009). ‘Price stability with imperfect financial integration’, *Journal of Money, Credit and Banking*, 41, 121-49.

- [9] Blau, F. and L. Kahn (2007). ‘Changes in the labor supply behavior of married women: 1980–2000’, *Journal of Labor Economics*, 25, 393-438.
- [10] 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.
- [11] 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.
- [12] Chamley, C. (1986). ‘Optimal taxation of capital income in general equilibrium with infinite lives’, *Econometrica*, 54 (3), 607-622.
- [13] Chari, V., Lawrence, C. and P. Kehoe (1994). ‘Optimal fiscal policy in a business cycle model’, *The Journal of Political Economy*, 102, 617-652.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] Farhi, E. (2010). ‘Capital taxation and ownership when markets are incomplete’, *Journal of Political Economy*, 118 (5), 908-948.
- [18] Fisher, R. (1915). ‘Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population’, *Biometrika*, 10, 507–521.
- [19] Gervais, M. and A. Mennuni (2015). ‘Optimal fiscal policy in the neo-classical growth model revisited’, *European Economic Review*, 73, 1-17.
- [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] Guo, J.T and K.J. Lansing (1999). ‘Optimal taxation of capital income with imperfectly competitive product markets’, *Journal of Economics Dynamics and Control*, 23, 967-995.
- [25] He, H. (2012). ‘What Drives the Skill Premium: Technological change or demographic variation?’, *European Economic Review*, 56, 1546-1572.
- [26] He, H. and Z. Liu (2008). ‘Investment-specific technological change, skill accumulation, and wage inequality’, *Review of Economic Dynamics*, 11, 314-334.
- [27] Hindriks, J. and G. Myles (2013). *Intermediate Public Economics*, MIT Press.
- [28] 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.
- [29] Judd, K. (1985). ‘Redistributive taxation in a simple perfect foresight model’, *Journal of Public Economics*, 28 (1), 59-83.
- [30] Judd, K. (1997). ‘The optimal tax on capital income is negative’, National Bureau of Economic Research, Working Paper 6004.
- [31] Katz, L. and K. Murphy (1992). ‘Changes in relative wages, 1963-1987: supply and demand factors’, *The Quarterly Journal of Economics*, 107, 35-78.
- [32] Keane, M. and R. Rogerson (2012). ‘Micro and macro labor supply elasticities: A reassessment of conventional wisdom’, *Journal of Economic Literature*, 50, 464-476.

- [33] Kimball, M. and M. Shapiro, (2008).. ‘Labor supply: Are the income and substitution effects both large or both small?’, NBER Working Papers 14208, National Bureau of Economic Research.
- [34] Kocherlakota, N. (2010). ‘The new dynamic public finance’, Princeton University Press, Princeton, New Jersey.
- [35] Krusell, P., Ohanian, L., Ríos-Rull J. and G. Violante (2000). ‘Capital–skill complementarity and inequality: a macroeconomic analysis’, *Econometrica*, 68, 1029–1053.
- [36] Lindquist, M. (2004). ‘Capital–skill complementarity and inequality over the business cycle’, *Review of Economic Dynamics*, 7, 519-540.
- [37] Litterman, R. (1983). ‘A random walk, Markov model for the distribution of time-series’, *Journal of Business & Economic Statistics*, 1, 169-173.
- [38] Ljungqvist L. and T. Sargent (2012). ‘Recursive macroeconomic theory’, 3rd edition, The MIT Press, Cambridge, Massachusetts.
- [39] Mankiw, G. (2000). ‘The savers-spenders theory of fiscal policy’, *American Economic Review*, 90, 120-125.
- [40] Mankiw, G., Weinzierl, M. and D. Yagan (2009). ‘Optimal taxation in theory and practice’, *Journal of Economic Perspectives*, 23, 147-174.
- [41] Martinez-Mongay, C. (2000). ‘ECFIN’s effective tax rates. Properties and comparisons with other tax indicators’, Economic Paper No. 146, (October), Brussels, European Commission.
- [42] Matsuyama K. (2006). ‘The 2005 Lawrence R. Klein Lecture: emergent class structure’, *International Economic Review*, 47, 327-360.
- [43] Persson, T. and G. Tabellini (1992). ‘The politics of 1992: fiscal policy and European integration’, *Review of Economic Studies*, 59, 689-701.
- [44] Philippon, T. (2014). ‘Has the American finance industry become less efficient?’, *The American Economic Review* (forthcoming).
- [45] 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.

- [46] Pourpourides, P. (2011). ‘Implicit contracts and the cyclicity of the skill-premium’, *Journal of Economic Dynamics and Control*, 35, 963-979.
- [47] Schmitt-Grohé, S. and M. Uribe (2003). ‘Closing small open economy models’, *Journal of International Economics*, 61, 163-185.
- [48] Stockman, D. (2001). ‘Balanced-budget rules: Welfare loss and optimal policies’, *Review of Economic Dynamics*, 4, 438-459.
- [49] Traum, N. and S. Yang (2010). ‘Does government debt crowd out investment? A Bayesian DSGE approach’, Congressional Budget Office Working Paper 2010-02.
- [50] Werning, I. (2007). ‘Optimal fiscal policy with redistribution’, *The Quarterly Journal of Economics*, 122, 925-967.

Appendix: FOCs for households and firms

Skilled households

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

Hours worked

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

Capital

$$\beta \left\{ E_t U_{c_s}(t+1) \left[(1 - \tau_{s,t+1}^k) r_{t+1} + \frac{(1-\delta)}{A_{t+1}^e} - 2\psi_s K_{s,t+1} \right] \right\} - \frac{U_{c_s}(t)}{A_t^e} = 0 \quad (\text{A2})$$

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

Unskilled households

Hours worked

$$U_{c_u}(t) (1 - \tau_{u,t}^l) w_{u,t} + U_{h_u}(t) = 0 \quad (\text{A3})$$

Capital

$$\beta \left\{ E_t U_{c_u}(t+1) \left[(1 - \tau_{u,t+1}^k) r_{t+1} + \frac{(1-\delta)}{A_{t+1}^e} - 2\psi_u K_{u,t+1} \right] \right\} - \frac{U_{c_u}(t)}{A_t^e} = 0 \quad (\text{A4})$$

These first-order conditions equate the marginal benefits from labour hours and investment in capital and to their respective marginal costs. Note that the presence of the asset holding costs, ψ_s and ψ_u ensures that (A2) and (A4) differ in the steady-state, which is required so that the latter is uniquely determined even if $\tau_s^k = \tau_u^k$.

Hand-to-mouth households

Hours worked

$$U_{c_h}(t) (1 - \tau_{h,t}^l) w_{u,t} + U_{h_h}(t) = 0 \quad (\text{A5})$$

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:

$$MPK = r_t \quad (\text{A6})$$

where

$$\begin{aligned} MPK = & A_t \{ \lambda \left[\nu \left(K_t^f \right)^\rho + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right]^{\varphi/\rho} + (1 - \lambda) \times \\ & \times \left(h_{u,t}^f \right)^\varphi \}^{\frac{1-\varphi}{\varphi}} \left[\nu \left(K_t^f \right)^\rho + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right]^{\frac{\varphi-\rho}{\rho}} \times \\ & \times \lambda \nu \left(K_t^f \right)^{\rho-1} \end{aligned} \quad (\text{A7})$$

where:

$$\begin{aligned}
MPL^s &= A_t \{ \lambda \left[\nu \left(K_t^f \right)^\rho + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right]^{\varphi/\rho} + (1 - \lambda) \times \\
&\quad \times \left(h_{u,t}^f \right)^\varphi \}^{\frac{1-\varphi}{\varphi}} \left[\nu \left(K_t^f \right)^\rho + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right]^{\frac{\varphi-\rho}{\rho}} \times \\
&\quad \times \lambda (1 - \nu) \left(h_{s,t}^f \right)^{\rho-1} \\
MPL^u &= w_{u,t}
\end{aligned} \tag{A8}$$

where:

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

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

Figure 1: Response of Tax Rates by Shock

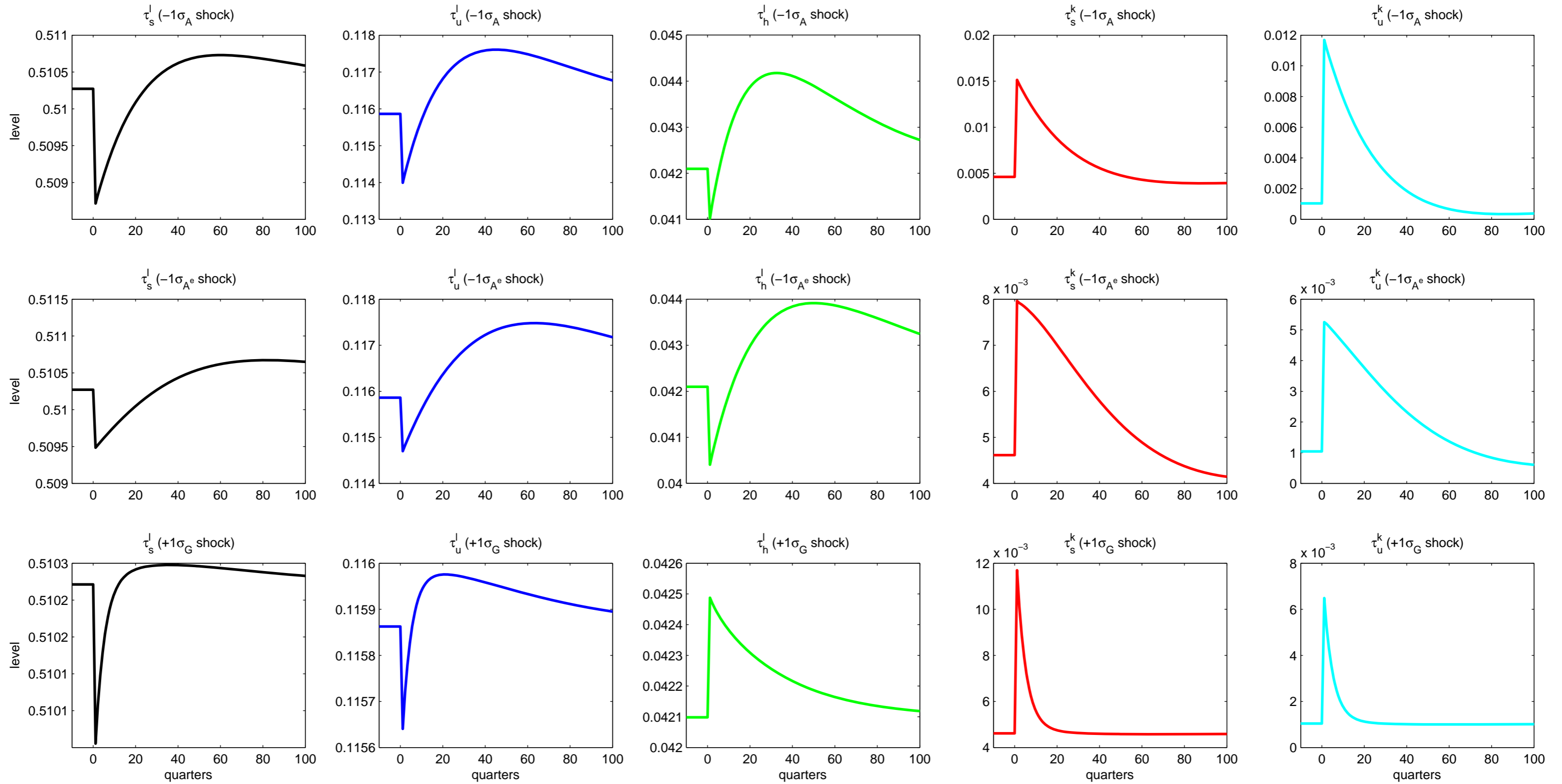


Figure 2: Response of Effective Tax Rates by Shock

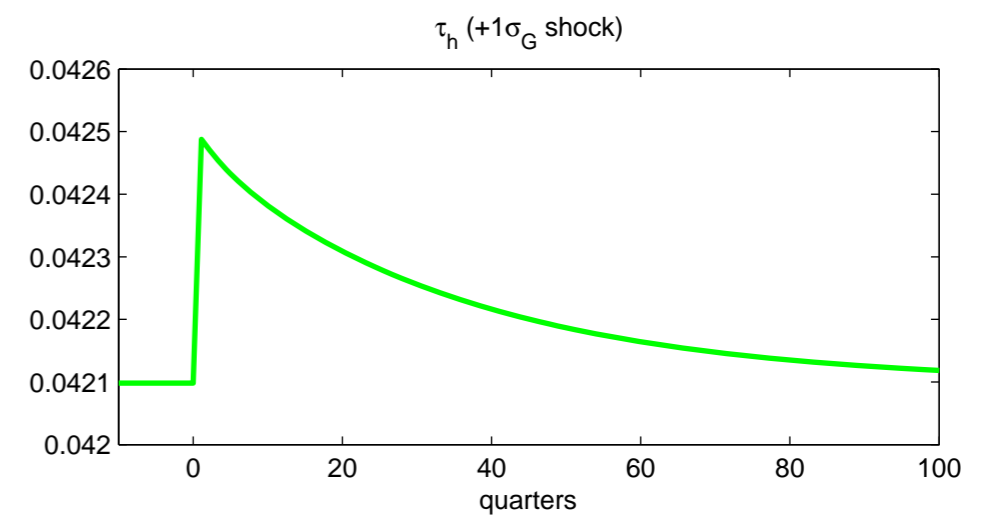
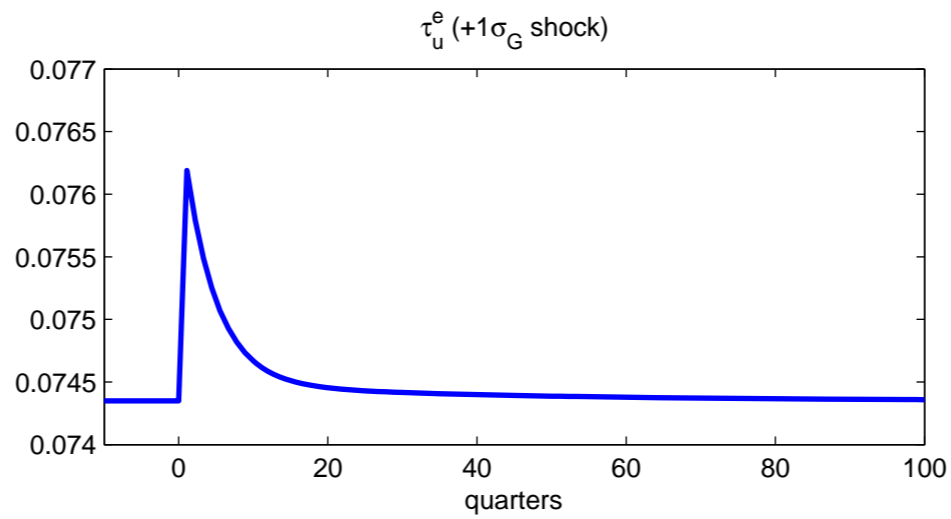
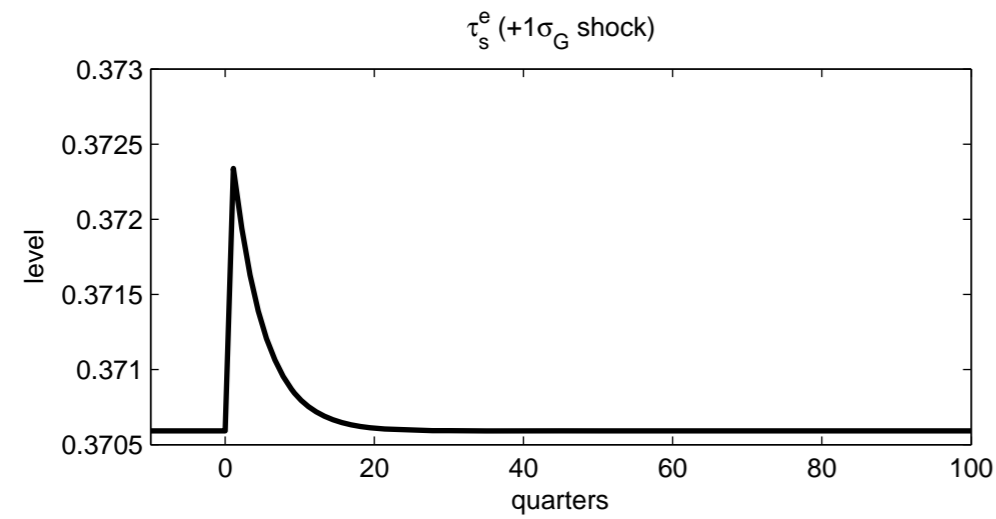
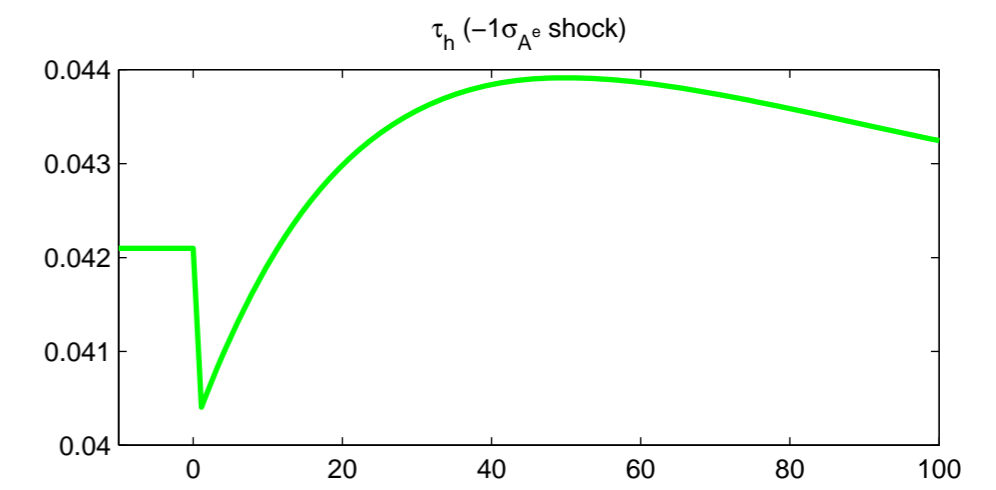
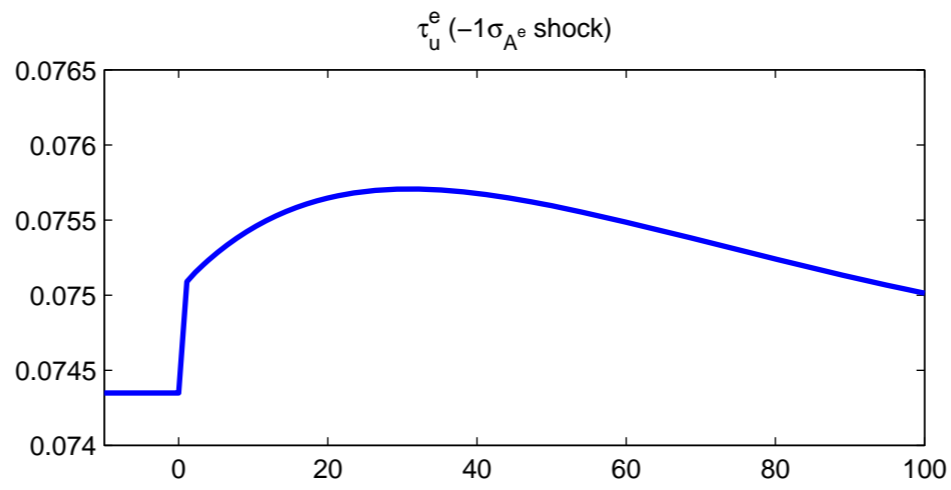
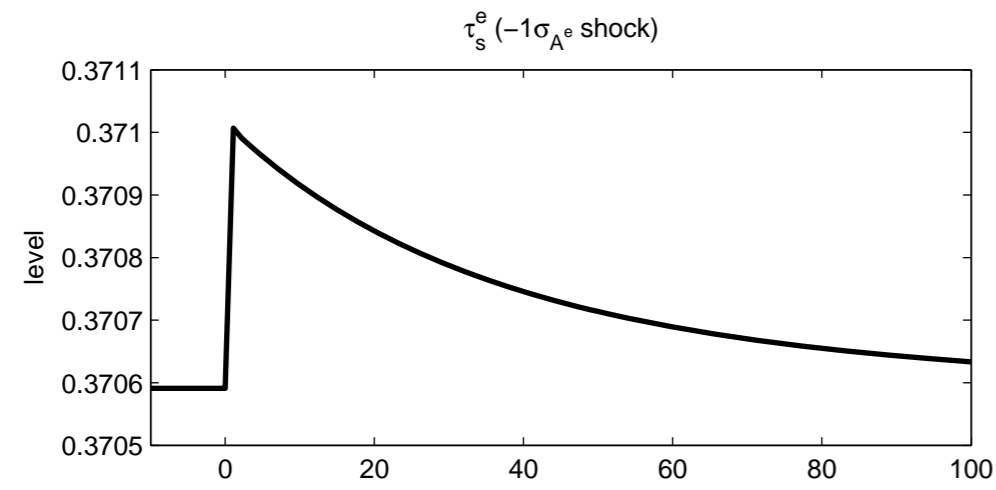
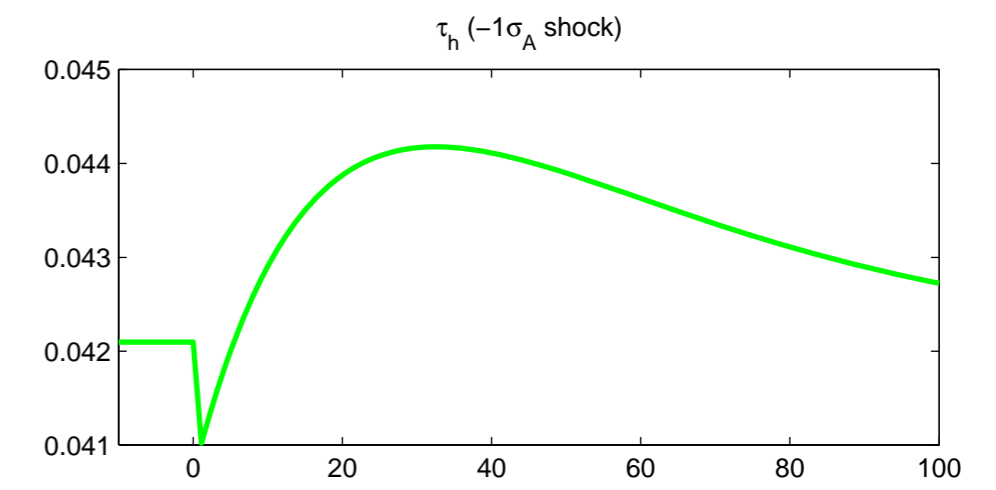
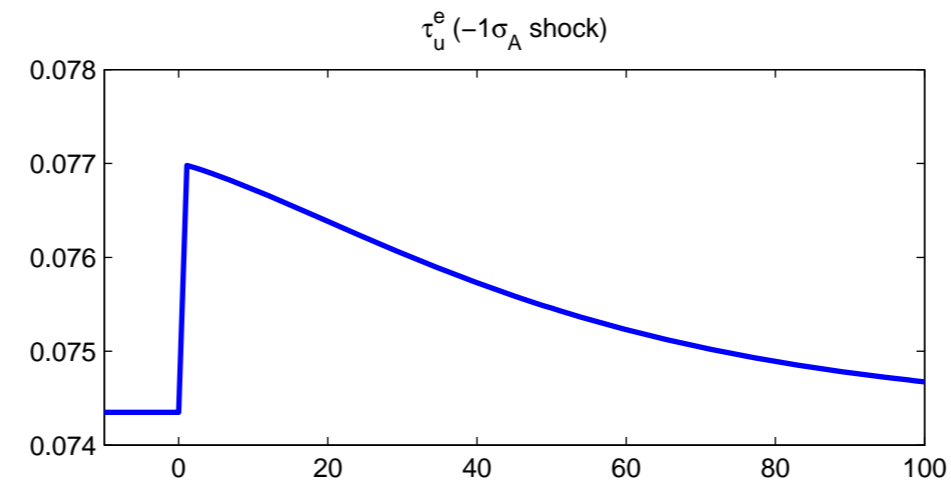
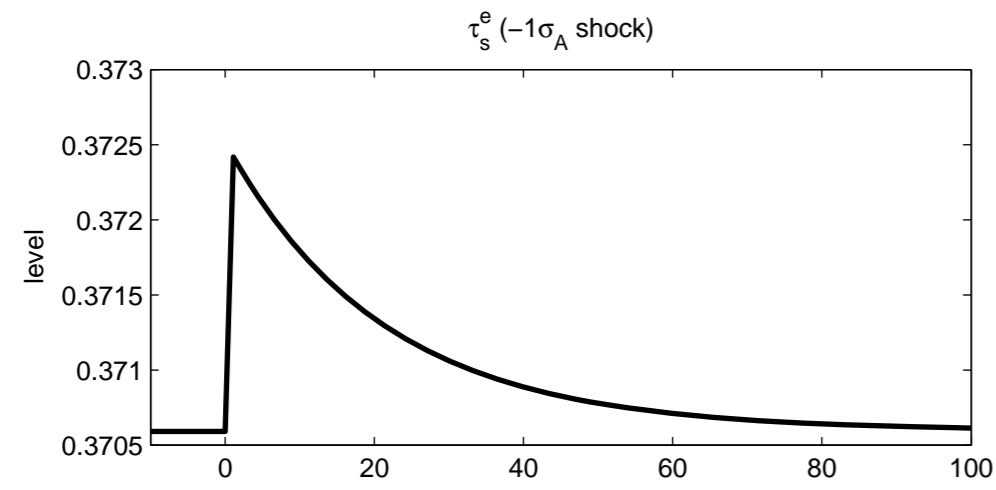


Figure 3: Tax Revenue Shares Paid by Group by Shock

