

Angelopoulos, Konstantinos; Asimakopoulos, Stylianos; Malley, Jim

Working Paper

The Optimal Distribution of the Tax Burden over the Business Cycle

CESifo Working Paper, No. 4468

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Angelopoulos, Konstantinos; Asimakopoulos, Stylianos; Malley, Jim (2013) : The Optimal Distribution of the Tax Burden over the Business Cycle, CESifo Working Paper, No. 4468, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<http://hdl.handle.net/10419/89780>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



The Optimal Distribution of the Tax Burden over the Business Cycle

Konstantinos Angelopoulos
Stylianos Asimakopoulos
James Malley

CESIFO WORKING PAPER NO. 4468

CATEGORY 6: FISCAL POLICY, MACROECONOMICS AND GROWTH

NOVEMBER 2013

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

The Optimal Distribution of the Tax Burden over the Business Cycle

Abstract

This paper analyses optimal income taxes over the business cycle under a balanced-budget restriction, for low, middle and high income households. A model incorporating capital-skill complementarity in production and differential access to capital and labour markets is developed to capture the cyclical characteristics of the US economy, as well as the empirical observations on wage (skill premium) and wealth inequality. We find that the tax rate for high income agents is optimally the least volatile and the tax rate for low income agents the least counter-cyclical. In contrast, the path of optimal taxes for the middle income group is found to be very volatile and counter-cyclical. We further find that the optimal response to output-enhancing capital equipment technology and spending cuts is to increase the progressivity of income taxes. Finally, in response to positive TFP shocks, taxation becomes more progressive after about two years.

JEL-Code: E240, E320, E620.

Keywords: optimal taxation, business cycle, skill premium, income distribution.

Konstantinos Angelopoulos
University of Glasgow
United Kingdom - G12 8RT Glasgow
konstantinos.angelopoulos@glasgow.ac.uk

Stylianos Asimakopoulos
University of Glasgow
United Kingdom - G12 8RT Glasgow
s.asimakopoulos@research.gla.ac.uk

James Malley
University of Glasgow
United Kingdom - G12 8RT Glasgow
j.malley@lbss.gla.ac.uk

November 1, 2013

We would like to thank Andrew Clausen, Michael Hatcher, Wei Jiang, Ioana Moldovan, Apostolis Philippopoulos for helpful comments and suggestions. We are also grateful for financial support from the ESRC, Grant No. RES-062-23-2292, but the views expressed here are entirely our own.

1 Introduction

There is a considerable literature that aims to characterize the properties of optimal tax policy over the business cycle (see e.g. Chari *et al.* (1994), Stockman (2001) and Arseneau and Chugh (2012)). This research has acknowledged the importance of market imperfections and of restrictions to the policy menu for optimal taxation. For instance, while in a frictionless labour market the labour income tax should optimally not vary much over the business cycle and remain a-cyclical, Arseneau and Chugh (2012) show that under search frictions in the labour market, the optimal labour income tax becomes very volatile and counter-cyclical. Moreover, Stockman (2001) shows that a balanced-budget restriction leads to an increase of the optimal volatility of the labour relative to capital taxes.

The literature, however, has not yet examined optimal income taxes over the business cycle under imperfections that limit the participation of households in markets for skilled labour and capital. 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 economic policy in response to economic fluctuations.²

Income taxation has naturally been a focal point for the research on economic policy under income inequality (see e.g. the work reviewed and analysed in Kocherlakota (2010)). This is because, on one hand, progressive income taxation can be used to reduce income inequality and promote a fairer distribution of income. On the other hand, the disincentives associated with taxation and, in particular, with progressive taxation, need to be taken into account. In light of this, the normative properties relating to the progressivity of the tax system have been extensively analysed (see e.g. Mankiw *et al.* (2009) for an assessment of this literature). However, the response of optimal income taxes in business cycle frequencies to exogenous productivity and government spending shocks, under a balanced budget and both wage and asset inequalities, has not been examined. This is particularly relevant given the presence of these inequalities and the current economic reality that severely limits the use of debt to respond to economic fluctuations in most advanced economies. In such an environment, the revenue requirements for governments that are faced with exogenous aggregate shocks need to be financed by unpleasant taxes, so that a pertinent question for policymaking becomes how to distribute the tax burden over the business cycle to minimise the

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, for example, the papers by Campbell and Mankiw (1989), Mankiw (2000) and Galí *et al.* (2007).

negative effect of distorting taxes.

In light of the above, we aim to analyse optimal income taxes under a balanced budget over the business cycle in a model that captures the cyclical characteristics of the economy and the empirical observations on wage (skill premium) and wealth inequality. To this end, we develop a model economy characterised by capital-skill complementarity in the production process, a labour market that is fragmented with respect to skill and capital market imperfections that lead to the exclusion of a subset of the population from holding assets.

Our model thus consists of three types of households, representing high, middle and low income groups, as well as two labour markets, for skilled and unskilled labour. We assume imperfections which preclude. These prevent all households from participating in both markets. In particular, we assume that the first type of household provides skilled labour services, capturing the supply of college-educated workers in the labour market. The other two types of households, in contrast, provide unskilled labour, and they represent the part of the labour force that does not have a college qualification. The production structure implies that there are two wage rates in the model, leading to a skill premium. Following the contributions of Katz and Murphy (1992) and Krusell *et al.* (2000), we assume that the skill premium is driven by skill-biased technical change and capital-skill complementarities. In particular, we assume that the production process follows the technology specified in Krusell *et al.* (2000) which has received empirical support and has been shown to match the behavior of the skill premium in the data.

Capital market imperfections imply that households in our model differ with respect to their participation in the asset markets. In particular, following the contributions of Campbell and Mankiw (1989), Mankiw (2000) and Galí *et al.* (2007), we assume that a subset of the households does not have any savings and is thus earning only labour income, which it totally consumes. 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, middle income unskilled agents who also own assets and low income unskilled agents who do not have access to the capital market.

Compared to the representative agent optimal taxation literature of e.g. Chari *et al.* (1994), Stockman (2001) and Arseneau and Chugh (2012), our modeling instead emphasises the importance of imperfections that lead to wage and wealth inequality as well as the balanced-budget constraint imposed on optimal income taxation. In contrast to the heterogeneous agents literature of optimal taxation (see e.g. the work reviewed and analysed in

Kocherlakota (2010))³, our modeling emphasises wage inequalities that are driven by capital-skill complementarities, in conjunction with asset market participation inequalities and focuses on the business cycle properties of income taxes.

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

We find that the cyclical properties of the income tax rates differ significantly with each other and with those observed in the data. As expected, given the balanced budget restriction and the instruments available to the government, the tax rates are generally more volatile and more counter-cyclical than in the data. However, there are also important differences between the tax rates. These result from the trade-off that the government faces when deciding how to distribute the distortions reflected by the higher volatility and countercyclicity of the three tax rates over the business cycle. On one hand, such distortions have a large impact on hand-to-mouth households, since they are less able to smooth shocks. 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 this acts to amplify external shocks. Optimal policy resolves this trade-off by keeping the lowest volatility for the tax rate for skilled and the lowest countercyclicity for the hand-to-mouth. In contrast, the middle income group, made up by unskilled households with savings, receives very volatile, very counter-cyclical taxes.

We further analyse the optimal distribution of the tax burden in the short- and medium-run in response to temporary output-enhancing exogenous shocks. The government finds it optimal to respond to an increase in the productivity of capital equipment and to public spending cuts by in-

³This literature pays particular attention to the unobservability of idiosyncratic labour productivity, which drives wage inequality. However, here we emphasise observable University education and employment in skilled jobs which drive the college-premium wage inequality.

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

creasing the progressivity of income taxes. In the case of capital equipment technology shocks, in particular, the government finds it optimal to redistribute some of the gains to skilled workers, who are the main beneficiaries of such changes, to the more constrained households in the labour market. This is achieved by increasing the high-income tax and reducing the other two taxes. Public spending cuts allow the government to reduce all income taxes, but the reduction is higher the lower the income level of the household. Finally, the response to positive total factor productivity (TFP) shocks implies that the progressivity of the tax system increases after about two years. The sensitivity of the income of hand-to-mouth households with respect to TFP shocks implies that the government needs to use a pro-cyclical tax on impact in this case to help smooth consumption. As a result, income taxes to low income agents increase immediately after positive TFP shocks.

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

Our model is developed to capture the key business cycle features of an economy characterised by imperfections that limit participation in labour and capital markets. We first consider a fragmented labour market, so that there exist separate markets for "skilled" and "unskilled" labour, defined as workers with and without college education, and assume that there exist 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. In particular, using the data in Acemoglu and Autor (2011), we find that the standard deviation of the cyclical component of the skilled population share, relative

⁵When looking at longer horizons, it is natural to allow for mobility from unskilled to skilled labour, associated with human capital investment and university education (see e.g. He (2012) and Angelopoulos *et al.* (2013b) 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.

to that of output, is 0.27, while its correlation with output is -0.18.⁶ These findings suggest that the imperfections (in the form of, e.g. socio-economic barriers relating to access to education) which determine participation in labour markets are indeed more restrictive in shorter, business cycle horizons.

This environment leads to wage inequality. Following the literature on the skill premium driven wage inequality (see e.g. Hornstein *et al.* (2005) and Acemoglu and Autor (2011) for reviews), we assume that the production process involves two types of labour inputs, i.e. skilled and unskilled which have different degrees of complementarity with capital. In particular, the production technology involves two types of capital and it is assumed that skilled labour services complement equipment capital in production relatively more than unskilled labour services, such that there is an increase in the skill premium in the labour market when technological innovations are augmenting equipment capital.

We also allow for transaction costs for participating in capital markets (see e.g. Schmitt-Grohé and Uribe (2003) and Benigno (2009)). Given inequalities in asset ownership, and in particular, evidence that suggests higher wealth for skilled relative to unskilled workers⁷, we distinguish these costs between skilled and unskilled households. This leads to different asset holdings across workers, and in particular, implies that a subset of the population is excluded from the asset markets (see e.g. Aghion and Howitt (2009, ch. 6) for capital market imperfections and agent heterogeneity). Excluded agents are thus not permitted to accumulate capital stock to smooth consumption, as they consume all their (labour) income (see e.g. Campbell and Mankiw (1989), Mankiw (2000) and Galí *et al.* (2007) for hand-to-mouth consumers). We assume that hand-to-mouth households offer unskilled labour services.

The above assumptions lead to an economy where households differ in their participation in both the labour and the asset markets.⁸ More specifically, there are 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. Given the previous discussion, the composition of the population is assumed to be

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.

⁷Data 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.

⁸A similar population decomposition is considered in the analysis of U.K. policy reforms in Angelopoulos *et al.* (2013a).

constant and exogenous. For simplicity, we also assume that the total size of the population, N , is constant. The above implies that $N = N_s + N_u + N_h$, where we define $n_s = N_s/N$, $n_u = N_u/N$, and $n_h = 1 - n_s - n_u$. There are also N identical firms and a government.

In each period, households act as price takers and make decisions regarding how much to consume, work and save. Firms act competitively and employ two types of capital stock together with the two types of labour to produce a homogeneous product. The government runs a balanced budget and imposes different tax rates on each income level. It uses the revenue from these taxes to finance public spending.

2.1 Households

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

$$U_j = E_t \sum_{t=0}^{\infty} \beta^t u(C_{j,t}, \bar{C}_{j,t-1}, l_{j,t}) \quad (1)$$

where E_t is the conditional expectations operator at period t ; $0 < \beta < 1$ is a constant discount factor; $C_{j,t}$ and $l_{j,t}$ are private consumption and leisure respectively at period t ; $\bar{C}_{j,t-1}$ is the average consumption of the j -type households in period $t - 1$, which is taken as given at the household level, and captures external habits in consumption (see e.g. Campbell and Cochrane (1999) and Ljungqvist and Uhlig (2000));⁹ and $u(\cdot)$ is the utility function which satisfies the Inada conditions. As discussed below, the presence of habits allows the model to match the empirical cyclical properties of consumption in the data.

The specific form for utility is given by:

$$u(C_{j,t}, \bar{C}_{j,t-1}, l_{j,t}) = \frac{[(C_{j,t} - \omega \bar{C}_{j,t-1})^\gamma l_{j,t}^{1-\gamma}]^{(1-\sigma)}}{1 - \sigma} \quad (2)$$

where ω measures the weight attached to external consumption habits within each type of household; $\sigma > 1$ is coefficient of relative risk aversion; and $0 < \gamma < 1$ is the weight of effective consumption in utility.

A household of type j faces the following time constraint:

$$1 = l_{j,t} + h_{j,t} \quad (3)$$

⁹Hence, we assume that there is "catching-up with the Joneses in the neighborhood", since each household compares its consumption level to that of its socio-economic class. See Ljungqvist and Uhlig (2000) for a discussion on various forms of catching-up and keeping-up with the Joneses and internal versus external habits.

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

$$C_{j,t} + I_{j,t}^i = (1 - \tau_{j,t})w_{j,t}h_{j,t} + (1 - \tau_{j,t}) (r_t^q K_{j,t}^q + r_t^e K_{j,t}^e) - T_t - \psi_j \left((K_{j,t}^q)^2 + (K_{j,t}^e)^2 \right) \quad (4)$$

while hand-to-mouth households face the constraint:

$$C_{h,t} = (1 - \tau_{h,t})w_{u,t}h_{h,t} - T_t \quad (5)$$

where the superscript $i = q, e$ refers to structures, q , and equipment, e ; $I_{j,t}^i$ is investment; $K_{j,t}^i$ is the capital stock; $w_{j,t}$ is the wage rate; ψ_j is the capital transaction cost for $j = s, u$; and T_t is a lump-sum tax. The above budget constraints capture several key features of the model. First, the 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 transaction costs. In particular, the hand-to-mouth households implicitly face transaction costs that are infinite, so that they are excluded from the capital markets. The remaining households face finite transaction costs, modelled here as quadratic functions of the capital stock, following e.g. Persson and Tabellini (1992) and Benigno (2009). These may differ so that the households can be differentiated with respect to their steady-state holdings of wealth. Third, there are two types of capital holdings, in structures and equipment, which pay different rates of return. The importance of allowing for the two types of capital is explained below in the discussion of skill-biased technology in production. Fourth, for each level of income, as reflected by the household type, there is a different income tax rate.

Finally the motion of the capital stock for $j = s, u$ is:

$$K_{j,t+1}^i = (1 - \delta^i)K_{j,t}^i + I_{j,t}^i \quad (6)$$

where, $0 \leq \delta^i \leq 1$ is the depreciation rate.

Each household $j = s, u$ chooses $\{C_{j,t}, h_{j,t}, K_{j,t+1}^q, K_{j,t+1}^e, I_{j,t}^q, I_{j,t}^e\}_{t=0}^{\infty}$, to maximise (1) subject to (2), (3), (4) and (6), by taking policy variables, prices, and aggregate quantities (i.e. $\bar{C}_{j,t-1}$) as given. Similarly, hand-to-mouth households, $j = h$, choose $\{C_{h,t}, h_{h,t}\}_{t=0}^{\infty}$, to maximise (1) subject to (2), (3) and (5), by taking policy variables, prices, and aggregate quantities (i.e. $\bar{C}_{h,t-1}$) as given. The optimality conditions for the households are given in Appendix A.

2.2 Production and firms

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

$$Y_t = A_t \left(K_t^{f,q} \right)^\alpha \times \left[\lambda \left(\nu \left\{ (A_t^e)^\rho \left(K_t^{f,e} \right)^\rho \right\} + (1 - \nu) \left(h_{s,t}^f \right)^\rho \right)^{\varphi/\rho} + (1 - \lambda) \left(h_{u,t}^f \right)^\varphi \right]^{\frac{1-a}{\varphi}} \quad (7)$$

where,

$$0 < a, \lambda, \nu < 1; \quad -\infty < \varphi, \rho < 1;$$

A_t is total factor productivity; A_t^e is the efficiency level of capital equipment; φ , and ρ are the parameters determining the factor elasticities, i.e. $1/(1-\varphi)$ is the elasticity of substitution between equipment capital and unskilled labour and between skilled and unskilled labour, whereas $1/(1-\rho)$ is the elasticity of substitution between equipment capital and skilled labour; and a, λ, ν are the factor share parameters. In this specification, capital-skill complementarity is obtained if $1/(1-\rho) < 1/(1-\varphi)$. Appendix B analytically confirms that the skill premium, defined as $\frac{w_s}{w_u}$, is increasing in equipment capital, $K_t^{f,e}$, and decreasing in the relative supply of skilled labour, $\frac{h_{s,t}^f}{h_{u,t}^f}$, for the parameter restrictions considered.

Following the literature, A_t and A_t^e are assumed to follow stochastic exogenous AR(1) processes:

$$A_{t+1} = (1 - \rho_A) A + \rho_A A_t + \varepsilon_t^A \quad (8)$$

$$A_{t+1}^e = (1 - \rho_{A^e}) A^e + \rho_{A^e} A_t^e + \varepsilon_t^{A^e} \quad (9)$$

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

where ε_t^A and $\varepsilon_t^{A^e}$ are independently and identically distributed Gaussian random variables with zero means and standard deviations given respectively by σ_A and σ_{A^e} .¹¹

Under this production technology, an increase in the efficiency level of capital equipment, A_t^e , favours the productivity of skilled workers more than the productivity of unskilled workers and is thus skill-biased. Hence, the model is consistent with the empirical evidence that points to rising productivity for equipment capital 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).

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

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

subject to the technology constraint in (7). In equilibrium, profits are zero.

2.3 The government

The government runs a balanced budget in every period which is given by:

$$\begin{aligned} G_t^c = & n_s \tau_{s,t} w_{s,t} h_{s,t} + n_u \tau_{u,t} w_{u,t} h_{u,t} + n_h \tau_{h,t} w_{u,t} h_{h,t} + \\ & + \tau_{s,t} n_s (r_t^q K_{s,t}^q + r_t^e K_{s,t}^e) + \tau_{u,t} n_u (r_t^q K_{u,t}^q + r_t^e K_{u,t}^e) + T_t \end{aligned} \quad (10)$$

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

$$G_{t+1}^c = (1 - \rho_{G^c}) G_t^c + \rho_{G^c} G_t^c + \varepsilon_t^{G^c} \quad (11)$$

where $\varepsilon_t^{G^c} \sim iidN(0, \sigma_{G^c}^2)$. Regarding the tax rates, we consider below policy regimes where they are exogenously set or they are optimally chosen by the government. Following Arseneau and Chugh (2012), when we consider how the model economy behaves in response to exogenous fiscal policy, we use lump-sum taxes as the residual variable in the government budget constraint,

¹¹We consider the total factor productivity (TFP) and equipment capital augmenting exogenous processes due to the predominant role attached to them in the literature on economic fluctuations and the skill premium. For example, see Lindquist (2004) and Pourpourides (2011) who examine the skill premium in business cycle frequencies under these two exogenous processes.

since for this experiment we are not studying government financing issues. However, for the optimal policy analysis, again as in Arseneau and Chugh (2012), lump-sum taxes are fixed to zero.

In the literature that examines the optimality or not of tax smoothing (see e.g. Chari *et al.* (1994) and Arseneau and Chugh (2012)), the government budget constraint includes debt. In contrast, here we focus on the optimal allocation of the tax burden over the business cycle given the revenue requirements of the government. Hence we do not allow the government to issue debt to balance the budget (see also Stockman (2001), who considers optimal capital and labour taxes with and without access to debt, albeit in a different setup).

2.4 Market clearing conditions

The labour and capital market clearing conditions are given by:

$$h_s^f = n_s h_s \quad (12)$$

$$h_u^f = n_u h_u + n_h h_h \quad (13)$$

$$K_t^i = n_s K_{s,t}^i + n_u K_{u,t}^i. \quad (14)$$

The aggregate resource constraint is:

$$Y_t = G_t^c + n_s C_{s,t} + n_u C_{u,t} + n_h C_{h,t} + n_s (I_{s,t}^q + I_{s,t}^e) + n_u (I_{u,t}^q + I_{u,t}^e) + n_s \psi_s \left[(K_{s,t}^q)^2 + (K_{s,t}^e)^2 \right] + n_u \psi_u \left[(K_{u,t}^q)^2 + (K_{u,t}^e)^2 \right]. \quad (15)$$

3 Exogenous policy

Before studying the model's implications for optimal tax policy, we analyse its cyclical properties under an exogenous fiscal policy. In this section, we calibrate the model so that it generates empirically relevant business cycle fluctuations. We concentrate on the key labor market dimension that determines inequality, i.e. the skill premium, when driven by the empirically relevant government spending and income tax rate processes. Hence, we assume that the income tax rates for $j = s, u, h$ also follow AR(1) processes:

$$\tau_{j,t+1} = (1 - \rho_j) \tau_j + \rho_j \tau_{j,t} + \varepsilon_t^{\tau_j} \quad (16)$$

where $\varepsilon_t^{\tau_j} \sim Niid(0, \sigma_{\tau_j})$.

3.1 Decentralized competitive equilibrium

Given initial levels of capital stock for structures, K_0^q , and equipment, K_0^e , the four policy instruments $(\tau_{s,t}, \tau_{u,t}, \tau_{h,t}, G_t^e)$ and the stationary stochastic processes $\{A_t, A_t^e\}_{t=0}^\infty$, the DCE system of equations is characterized by a sequence of allocations $\{C_{s,t}, C_{u,t}, C_{h,t}, h_{s,t}, h_{u,t}, h_{h,t}, K_{s,t+1}^q, K_{s,t+1}^e, K_{u,t+1}^q, K_{u,t+1}^e, I_{s,t}^q, I_{s,t}^e, I_{u,t}^q, I_{u,t}^e\}_{t=0}^\infty$, prices $\{w_{s,t}, w_{u,t}, r_t^q, r_t^e\}_{t=0}^\infty$, and the residual policy instrument $\{T_t\}_{t=0}^\infty$ such that: (i) households maximize their welfare and firms their profits, taking policy, prices and aggregate variables as given; (ii) the government budget constraint is satisfied in each time period; (iii) all markets clear and (iv) $\bar{C}_{j,t-1} = C_{j,t-1}$. The full decentralized competitive equilibrium (DCE) is set out in Appendix A.

3.2 Data analysis and targets

We aim for the exogenous-policy model to replicate the long-run great ratios and key labour market averages as well as explaining the cyclical volatilities and correlations with output of key variables in the economy. We use quarterly data for U.S. economy, which are obtained from datasets constructed by Lindquist (2004), Piketty and Saez (2007), Castro and Coen-Pirani (2008), Pourpourides (2011), Arsenau and Chugh (2012) as well as data series from the Bureau of Economic Analysis (BEA).¹²

Table 1: Business cycle statistics of main endogenous variables

Variable	Correlation with output	Standard deviation
Y	1	0.013-0.014
C	0.83	0.011-0.012
I	0.76-0.91	0.037-0.063
$\frac{w_s}{w_u}$	0.09-0.19	0.006-0.013
h_s	0.42-0.69	0.008-0.010
h_u	0.59-0.73	0.006-0.012

Sources: The data ranges are constructed using the results reported in Lindquist (2004), Castro and Coen-Pirani (2008) and Pourpourides (2011).

In Table 1 we report the data volatilities and correlations with output from existing studies for variables which correspond with key endogenous variables in our model. These are quarterly data for the period 1979-2002 (Lindquist (2004)) and 1979-2003 (Castro and Coen-Pirani (2008) and Pourpourides (2011)). Their cyclical component has been obtained by taking the

¹²We are particularly grateful to Matthew Lindquist and Daniele Coen-Pirani for providing their datasets.

logarithms of the series and then using an HP-filter with a smoothing parameter of 1600.¹³ As can be seen in Table 1, these studies document some interesting results regarding the labour market statistics. In particular, they point out that the skill premium is effectively uncorrelated with output and smoother than output in business cycle frequencies. Moreover, the cyclical properties of the labour supply of skilled and unskilled workers do not differ qualitatively, both having a positive correlation with output, while being less volatile than output. The statistics regarding consumption, investment and capital are similar to those commonly obtained in macroeconomic research.

Table 2: Data averages and business cycle statistics
of policy variables

	G^c	τ_{60-100}	τ_{20-60}	τ_{0-20}
Averages	-	0.247	0.180	0.144
Autocorrelations	0.770	0.950	0.920	0.890
Correlations with Y	-0.066	0.587	0.654	0.198
Standard deviation	0.014	0.006	0.005	0.004

Sources: The statistics are obtained using the data from Piketty and Saez (2007), Arsenau and Chugh (2012) and BEA.

We further examine the statistical properties of the fiscal policy variables in Table 2 where we present the means, as well as the first-order autocorrelations, standard deviations and correlations with output for the tax rates and government spending series. The government spending series is obtained using quarterly data from the BEA for the period 1979 to 2002.¹⁴ The income tax data are obtained using the Piketty and Saez (2007) dataset, which reports annual data on income tax rates per income group (in quantiles) for

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 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. We use the quarterly data for our business cycle analysis. However, note that the average skill premium as well as its second moments using the annual data and the classification in Acemoglu and Autor (2011) gives very similar results. In particular, the skill premium on average is 1.60, while its cyclical relative volatility and correlation with output are respectively given by 0.49 and -0.13.

¹⁴This series refers to government consumption expenditures and gross investment as it is reported in NIPA Table 1.1.5. To calculate the statistical properties of the cyclical component of the series, we log it and apply the HP-filter with a smoothing parameter of 1600.

the period 1966-2001.¹⁵ As we explain below, we calibrate the share of hand-to-mouth agents to be 20%, the share of unskilled workers who also have savings to be 40% and the share of skilled workers to 40%. Since our model predicts that the income levels of these three groups increase in the order mentioned above, we use the Piketty and Saez (2007) dataset to obtain three income tax rates, the first for the lowest quantile, the second as the average for the two middle quantiles, and the third as the average for the two top quantiles. We use these series of tax rates as proxies for $\tau_{j,t}$, $j = h, u, s$, respectively, in our model.

As can be seen in Table 2, the average tax for the bottom quantile, τ_{0-20} , is equal to 14.4%, for the next two, τ_{20-60} , 18% and for top two, τ_{60-100} , 24.7%, suggesting that income taxation at this disaggregation is progressive.¹⁶ Regarding the business cycle statistics of the tax series, the results suggest that, as expected, these are highly persistent and have low volatility. The spending process is less persistent and more volatile. The correlations with output suggest that all the tax rates are pro-cyclical and the government spending is essentially uncorrelated with output.¹⁷ Finally, using data on the productivity of capital equipment from the BEA for the period 1988-2011, we estimate the autocorrelation of HP-filtered series to be 0.975 and its standard deviation to be 0.007.¹⁸

3.3 Calibration

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

¹⁵These tax rates refer to average tax rates by income groups. To obtain quarterly series from the annual data, so that the cyclical statistics from the tax series are comparable to the remaining data used in the analysis, we follow the interpolation method in Litterman (1983). We use as an indicator variable the quarterly time-series of labour income tax rates from Arsenau and Chugh (2012). To obtain the cyclical component of the series we again use the HP-filter with a smoothing parameter of 1600 (see also e.g. Arsenau and Chugh (2012)).

¹⁶While Piketty and Saez (2007) discuss many aspects of the progressivity of income taxation, we focus here on income taxes for the three income groups that best correspond to the household disaggregation in our model.

¹⁷Note that using the annual series for tax rates, the results regarding the statistical properties of the cyclical components of the series are similar qualitatively.

¹⁸The time-series on the productivity of equipment capital is obtained from the Bureau of Labor Statistics and refers to annual data. The respective quarterly series is obtained using the methodology in Litterman (1983), where the indicator variable is the quarterly time-series of total investment for the same period. The series is then logged and HP-filtered with a smoothing parameter of 1600.

with the data. The calibrated parameters are summarised in Table 3.

3.3.1 Population shares

We assume that the population breakdown in our model economy is given as $n_s = 0.4$, $n_u = 0.4$, $n_h = 0.2$. The share of skilled households is consistent with the data in Acemoglu and Autor (2011), which implies that the average share of the labour force with a college degree is about 45%. The split of unskilled households into hand-to-mouth and those who can access the asset market, coheres with empirical evidence from Traum and Yang (2010), who estimate the share of the hand-to-mouth population for the U.S. at 18% and Cogan *et al.* (2010), who estimate the share of the hand-to-mouth population at 26.5%. The above split is also consistent with data from the 2010 U.S. Census, which indicates that 43% of the populations has a college degree and that the percentage of households without any assets is 18.7%.¹⁹

3.3.2 Tax-spending policy

A particular advantage of the 40/40/20 percent split 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 quantile, as described above. Therefore, we set the constant terms in the AR processes described above, τ_j , $j = h, u, s$, to be equal to the data averages for the respective income quantiles, i.e. for τ_{0-20} , τ_{20-60} and τ_{60-100} . Moreover, we set ρ_j and σ_j equal to the autocorrelation parameters and the standard deviations of cyclical component of the respective tax series in the data (see also e.g. Arsenau and Chugh (2012)). Following the same procedure, we also set the autocorrelation and standard deviation parameters for the processes for government spending (ρ_G and σ_G) to be equal to the respective estimates of the cyclical component of the public spending series described above (see also e.g. Arsenau and Chugh (2012)). Finally, we calibrate the long-run value of government spending to obtain a public spending to output ratio of 19%, consistent with the data discussed above.

3.3.3 Production and capital and labour markets

The elasticities of substitution between skilled labour and capital equipment and between unskilled labour and capital equipment (or skilled labour) have been estimated by Krusell *et al.* (2000). We use their estimates, so that $\varphi =$

¹⁹This information is obtained from Table 4 of the Census Bureau, Survey of Income and Program Participation.

0.401 and $\rho = -0.495$. The remaining parameters in the production function are calibrated to make the steady-state predictions of the model in asset and labour markets 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 1.66 and a labour share of income of 69%, which are consistent with the U.S. data. In particular, the target value for the skill premium is obtained from U.S. Census data and is within the range of estimates in Table 1.²⁰ The share of labour income in GDP is obtained from BEA data on personal income for the period 1970-2011. The productivity of capital structures, α is set at the same value as Lindquist (2004) and helps to bring the model's capital to output ratios close to the data. The calibrated parameters in the production function are generally very similar to those estimated or calibrated in the literature.

The depreciation rates of capital structures and capital equipment are calibrated to obtain a quarterly capital to output ratio equal to 6.95 in the steady-state. This is within the range presented in Table 1 and is consistent with an annual capital to output ratio of 1.74, obtained using BEA annual data on capital stocks from 1970 to 2011. In particular, we set $\delta^e = 0.028$ to be within the range of Lindquist (2004) and Pourpourides (2011) and calibrate $\delta^q = 0.016$ residually.²¹

We set the transaction cost parameters as $\psi_s = 0.0002$ and $\psi_u = 0.0018$. There are two targets for these parameters. The first is that the total capital 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 2010 Census²², which 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. The second target is that in the steady-state the transaction costs cohere with a real return to capital (that excludes depreciation, taxes and transaction costs) of about 1% per quarter.²³

3.3.4 Utility function

The coefficient of relative risk aversion, σ , is set following previous studies (e.g. Schmitt-Grohé and Uribe, 2007) at $\sigma = 2$. The time discount fac-

²⁰The specific source is the Current Population Survey, 2011 Annual Social and Economic Supplement from the U.S. Census Bureau.

²¹For instance, Krusell *et al.* (2000) report $\delta^q = 0.0125$ and $\delta^e = 0.031$; Pourpourides $\delta^q = 0.014$ and $\delta^e = 0.027$; and Lindquist $\delta^q = 0.014$ and $\delta^e = 0.031$.

²²The specific information is obtained using Table 1 from the 2010 U.S. Census Bureau, Survey of Income and Program Participation.

²³The real rate of return to capital at an annual frequency is 4%, using data from the World Bank.

tor, $\beta = 0.99$, is calibrated to target the investment to output ratio in the data. The weight of consumption to utility, $\gamma = 0.225$, is set so that in the steady-state each household devotes about one third of its time to work, consistent with the long-run averages reported in Table 1. The consumption habit parameter, ω , is calibrated so that the model's predicted volatility of consumption is similar to the data. The value employed of 0.58 is also within the range (0.52 – 0.71) suggested by Christiano *et al.* (2005).

Table 3: Model parameters

Parameter	Value	Definition	Source
$0 \leq \delta^q \leq 1$	0.016	depreciation rate of capital structures	calibration
$0 \leq \delta^e \leq 1$	0.028	depreciation rate of capital equipment	calibration
$0 < \beta < 1$	0.990	time discount factor	calibration
$0 \leq \omega < 1$	0.580	habit persistence parameter	calibration
$0 < \gamma < 1$	0.225	weight attached to consumption in utility	calibration
$\sigma > 1$	2.000	coefficient of relative risk aversion	assumption
$0 \leq \alpha \leq 1$	0.130	income share of capital structures	calibration
$\frac{1}{1-\rho}$	0.669	capital equipment to skilled labour elasticity	assumption
$\frac{1}{1-\varphi}$	1.669	capital equipment to unskilled labour elasticity	assumption
$0 < \lambda < 1$	0.560	share of composite input to output	calibration
$0 < \nu < 1$	0.580	share of capital equipment to composite input	calibration
$0 < \frac{G^c}{Y} < 1$	0.190	government spending	calibration
$\psi_s > 0$	0.0002	transaction cost for skilled agents	calibration
$\psi_u > 0$	0.0018	transaction cost for unskilled agents	calibration
τ_s	0.247	average income tax rate, skilled	data
τ_u	0.180	average income tax rate, unskilled	data
τ_h	0.144	average income tax rate, hand-to-mouth	data
Stochastic processes			
σ_A	0.004	standard deviation of TFP	calibration
ρ_A	0.950	AR(1) coefficient of TFP	data
σ_{A^e}	0.007	standard deviation of cap. equipment	data
ρ_{A^e}	0.975	AR(1) coefficient of cap. equipment	data
σ_{τ_s}	0.006	standard deviation of income tax, skilled	data
ρ_{τ_s}	0.950	AR(1) coefficient of income tax, skilled	data
σ_{τ_u}	0.005	standard deviation of income tax, unskilled	data
ρ_{τ_u}	0.920	AR(1) coefficient of income tax, unskilled	data
σ_{τ_h}	0.004	standard deviation of income tax, hand-to-mouth	data
ρ_{τ_h}	0.890	AR(1) coefficient of income tax, hand-to-mouth	data
σ_{G^c}	0.014	standard deviation of public spending	data
ρ_{G^c}	0.770	AR(1) coefficient of public spending	data

3.3.5 Technology

The constant terms in the processes for TFP and capital equipment productivity are normalized to unity (i.e. $A = 1$ and $A^e = 1$ respectively). We set the autocorrelation and standard deviation parameters for the processes for capital equipment technology (ρ_{A^e} and σ_{A^e}), equal to the respective estimates of the cyclical component of the relevant data series described above (see also e.g. Lindquist (2004), Pourpourides (2011)). The autocorrelation parameter of TFP is set equal to 0.95, following Lindquist (2004) and Pourpourides (2011), while σ^A is calibrated to match the volatility of output observed in the data (see Table 1).

3.4 Solution and results

The steady-state solution of the DCE system for key variables is compared with their corresponding data averages in Table 4. To study dynamics, we compute a first-order approximation of the equilibrium conditions around the deterministic steady-state, by implementing the perturbation methods in Schmitt-Grohé and Uribe (2003).²⁴ We use the first-order accurate decision rules to simulate time paths of the equilibrium under shocks to total factor productivity, capital equipment augmenting technology, government spending, and income tax realizations, that are obtained from the distributions specified above. We conduct 1000 simulations, each 296 periods long. We drop the initial 200 periods so that the remaining series length of 96 periods corresponds with the number of observations in the data, i.e. 1979:1-2002:4. For each simulation, we then compute the required moments and report the means of these moments across the simulations in Table 5. This table also reports, for convenience, the predicted business cycle statistics from the studies of Lindquist (2004) and Pourpourides (2011).

Table 4: Steady-state of the exogenous policy model

Variable	Model	Data	Variable	Model	Data
$\frac{K}{Y}$	6.946	6.550	h_s	0.362	0.317
$\frac{I}{Y}$	0.151	0.159	h_u	0.364	0.348
$\frac{C}{Y}$	0.653	0.659	h_h	0.398	-
$\frac{G^c}{Y}$	0.189	0.195	$\frac{w_s}{w_p}$	1.649	1.659
r_{net}	0.010	0.010	$\frac{wh}{Y}$	0.699	0.686

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

²⁴We present results using a first-order approximation throughout the paper. Our findings do not change by using a second-order approximation.

well with the data.²⁵ In particular, Table 4 shows that all model predictions are quantitatively similar to the long-run averages in the data. It is also worth noting that the ratio of average hours worked by unskilled workers to the average hours worked by skilled labour in the model is 1.027, which is similar to the 1.099 obtained for the U.S. from the Bureau of Labor Statistics (BLS) data.²⁶ These work-time allocations imply Frisch (or λ -constant) labour supply elasticities of 1.08 for skilled, 1.07 for unskilled and 0.93 for hand-to-mouth workers, which are again generally consistent with the literature (see e.g. Browning *et al.* (1999), Chetty *et al.* (2011), and Keane and Rogerson (2012)).²⁷

Table 5: Business cycle statistics of the exogenous policy model

Variable	Correlation with Output			Standard deviation		
	Model	Lindquist (2004)	Pourpourides (2011)	Model	Lindquist (2004)	Pourpourides (2011)
Y	1	1	1	0.014	0.013	0.014
C	0.89	0.81	0.93	0.012	0.006	0.012
I	0.80	0.96	0.76	0.048	0.053	0.040
$\frac{w_s}{w_u}$	0.15	0.06	-0.09	0.005	0.003	0.005
h_s	0.68	0.81	0.95	0.005	0.006	0.007
h_u	0.28	0.95	0.54	0.004	0.006	0.012
h_h	-0.42	N/A	N/A	0.007	N/A	N/A

Turning to the business cycle statistics in Table 5, the overall fit is comparable to existing research on business cycle models with the skill premium (see e.g. Lindquist (2004) and Pourpourides (2011)). In particular, the model matches the key stylised facts about the skill premium in the data, i.e. that it is effectively not correlated with output and that its volatility is less than that of output (refer to Table 1). In addition, the model predictions regard-

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) plus Table 7.2B (line 32) for the capital stock; (iii) BLS, Current Employment Statistics survey for hours worked; (iv) World Bank for the real rate of return; (v) BEA, NIPA Table 2.1 for labour's share in income; and (vi) U.S. Census Bureau, Survey of Income and Program Participation for the skill premium. Comparable averages are obtained using the dataset in Lindquist (2004), for those variables that are similar in both studies.

²⁶The Castro and Coen-Pirani (2008) quarterly data U.S. from 1979-2003 gives an unskilled to skilled labour supply ratio in efficiency units equal to 1.030. Our model prediction of 1.027 for the weighted ratio of unskilled to skilled labour coheres well with this figure.

²⁷Table 4 also shows that the labour's share of income in the model, $\frac{w_s n_s h_s + w_u n_u h_u + w_h n_h h_h}{Y} = 0.699$ is close to the value (i.e. 0.686) obtained from the BEA Table 2.1 for 1979-2002.

ing the second moments of the hours worked are generally consistent with the data both qualitatively and quantitatively. However, the model quantitatively under-predicts the correlation of unskilled hours with output. The model also matches the second moments of consumption and investment. Overall, the model's predictions regarding the key endogenous variables are empirically relevant.

4 Optimal tax policy over the business cycle

Having established the empirical relevance of the calibration, we now discard the exogenous processes for the income tax rates in (16) and instead assume that the paths of these tax rates are optimally chosen by a government that seeks to maximise a utilitarian objective function under commitment, taking the revenue requirements as given.

4.1 The problem of the government

The government chooses the paths of the three income tax rates to maximise aggregate welfare subject to the DCE under the following assumptions:²⁸ (i) It takes the spending side of the government budget as given. In particular, government consumption spending continues to follow the exogenous process in (11); (ii) it does not have a complete tax system at its disposal and can only tax each agent's total income;²⁹ and (iii) it cannot issue debt. Treating the spending side of the budget as given is a common assumption in the analysis of optimal taxation and allows us to focus on the revenue side of the budget. The second and third assumptions are motivated by current economic reality. In particular, the politico-economic framework does not allow governments to tax all sources of income differently and existing levels of debt imply that it cannot be easily used to smooth economic fluctuations.

In light of the labour and asset market imperfections and the restrictions placed on government policy we ask - what is the optimal distribution of the income tax burden over the business cycle? The requirements imposed

²⁸We also keep lump-sum taxes fixed to zero, as is common for optimal taxation analysis. We note, however, that our results do not change qualitatively if we keep the lump-sum instrument fixed to its steady-state value obtained from the model under exogenous policy.

²⁹We examine an optimal taxation problem with an incomplete set of tax instruments, since tax rates are not available for all pairs of goods in the economy (see e.g. Chari and Kehoe (1999), for the definition of a complete tax system). Motivated by constraints imposed on economic policy in practice, we consider a government that is restricted to tax capital and labour income at the same rate. However it is allowed to choose a different tax rate for each income group in the economy.

on the tax system certainly dictate some properties that optimal income taxation must satisfy over the business cycle. In particular, we would expect income taxes on average to be more volatile compared to the data since tax smoothing via public debt or expenditure management is not an option. We would also expect them to be generally counter-cyclical, given that negative shocks that reduce the tax bases necessitate a rise in the tax rates to make up for the loss in the tax revenue. However, allowing the government to choose different tax rates for each income group, implies that the government still has to decide on whether these tax rates should have the same volatility and co-movement with output over the business cycle and, if not, how to set these cyclical properties for each tax rate.

We examine the problem of a government that has Utilitarian preferences, so that its objective function is given by the expected lifetime utility of the weighted average of the welfare of the three types of households, where the weights attached to each type are equal to the population share of that type, n_j . In this case, the government chooses $\{\tau_{s,t}, \tau_{u,t}, \tau_{h,t}, C_{s,t}, C_{u,t}, C_{h,t}, h_{s,t}, h_{u,t}, h_{h,t}, K_{s,t+1}^q, K_{s,t+1}^e, K_{u,t+1}^q, K_{u,t+1}^e, I_{s,t}^q, I_{s,t}^e, I_{u,t}^q, I_{u,t}^e, w_{s,t}, w_{u,t}, r_t^q, r_t^e\}_{t=0}^\infty$, to maximise:

$$U^g = E_t \sum_{j=s,u,h} \left[n_j \sum_{t=0}^{\infty} \beta^t u(C_{j,t}, C_{j,t-1}, l_{j,t}) \right] \quad (17)$$

subject to the DCE equations (18) – (47) in Appendix A, where we set $\{T_t \equiv 0\}_{t=0}^\infty$. Note that the government internalises the externalities in consumption when making its optimal choices. We also assume that the government can commit to the optimal paths.

As in the exogenous-policy baseline, we first compute the deterministic steady-state equilibrium under optimal policy and next approximate the dynamic equilibrium paths using the first-order approximation of the equilibrium conditions under optimal policy for time $t > 0$ around the deterministic steady-state of these conditions. As is common in the literature (see e.g. Arsenau and Chugh (2012)), when characterizing asymptotic policy dynamics, we also make the auxiliary assumption that the initial state of the economy at $t = 0$ is the steady-state under optimal policy. We use the first-order accurate decision rules to simulate the optimal-policy equilibrium under shocks to TFP, equipment capital and government consumption spending that are obtained as in the exogenous-policy experiments in the previous section.³⁰

³⁰To calculate the business cycle statistics of the model under optimal policy we work as in the exogenous policy case. In particular, we conduct simulations under shocks to the exogenous processes, obtain the required statistics for each simulation and then calculate their mean value across the simulations. We conduct 1000 simulations, each 296 periods long and drop the initial 200 periods.

4.2 Optimal taxes over the business cycle

Table 6 presents the optimal properties of the tax system under shocks to all stochastic processes. As can be seen by the steady-state income taxes, optimal tax policy is progressive and, in fact, relatively more progressive compared with the data averages. In particular, $\tau_s > \tau_{60-100}$, $\tau_u < \tau_{20-60}$, $\tau_h < \tau_{0-20}$.³¹ This is noteworthy since the progressivity of the tax system as captured by the three tax rates considered here has indeed increased since the mid-1960s (see e.g. the data in Piketty and Saez (2007)).³² Therefore, the assumed imperfections and inequalities in our model justify progressive income taxation.

Table 6: Optimal tax policy

	τ_s	τ_u	τ_h
Steady-state	0.285	0.143	-0.041
Autocorrelations	0.754	0.832	0.891
Correlations with Y	-0.461	-0.732	-0.110
Standard deviation	0.012	0.070	0.173

These results in Table 6 further suggest that the volatility and co-movement of optimal taxes with output differ significantly with each other and with the data reported in Table 2. As discussed above, the policy problem we consider implies that the tax rates need to be generally more volatile and more counter-cyclical, compared with the data. However, Table 6 also reveals important differences between the tax rates. These result from the trade-off that the government faces when deciding how to distribute the distortions reflected by the higher volatility and countercyclicity of the three tax rates over the business cycle. On one hand, such distortions have a larger impact on hand-to-mouth households, since they are less able to smooth shocks. 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

³¹The bigger difference between the optimal and the actual level of taxes is observed in the taxation of the lowest income group. However, in practice, targeted subsidies may reduce the effective tax burden for this group as well, bringing the optimal effective tax burden for the low income, relative to the higher income groups predicted here closer to reality. Since, as explained above, we exclude lump-sum instruments from the policy menu, we do not consider further the role of targeted transfers.

³²For instance, using the Piketty and Saez (2007) dataset, τ_{60-100} has increased from 20% to 25%, τ_{20-60} has decreased from 18% to 14% and τ_{0-40} has decreased from 16% to 11%, between 1966 and 2001.

incentive to minimise distortions to the choices of skilled households, since this acts to amplify external shocks.

To analyse further the cyclical properties of optimal taxes in this framework, we start with the standard deviations. Recall that in the data, the volatilities of the tax rates are increasing with the level of the tax rate, but these differences are small, as the standard deviations of the three rates are quantitatively similar. In contrast, given the balanced budget restriction and the instruments available to the government, the optimal volatility of taxes is, as expected, higher than that found in the data. However, the increase in volatility is not the same across the three taxes. It is the highest for the low-income group and the lowest for the high-income group.³³

A smoother income tax creates fewer distortions in the household's optimisation problem. Since income taxes are the only choices for the policy-maker in this framework, implying that smoothing is not possible for all tax rates, the government finds it optimal to keep the tax rate which distorts incentives the most, the smoothest. Given the higher complementarities of skilled hours with equipment capital, implying that tax-induced fluctuations in skill supply propagate more in the economy via the equipment capital channel, the government finds it optimal to keep the tax to skilled households the least volatile. In contrast, since hand-to-mouth households do not own capital stock, their choices affect the endogenous propagation mechanism in the economy the least, so that their taxes are optimally the most volatile.

Recall from Table 2 that all income taxes are found to be pro-cyclical in the data, with output correlations in the range of 0.2 (for τ_{0-20}) to 0.65 (for τ_{20-60}). As expected, the results in Table 6 confirm that the correlations become negative when the tax instruments are the only optimally chosen instruments in the government's budget constraint. However, this Table also shows that the optimal correlations of the tax rates are symmetrically opposite to the data correlations. For example, the strongest pro-cyclical tax in the data becomes optimally the strongest counter-cyclical one, while the least cyclical in the data becomes the least counter-cyclical. Hence, the requirement of the government to make the tax system generally counter-cyclical, is not translated into a proportional reduction of the correlation coefficients of all tax rates.³⁴

³³Our results regarding the magnitude and range in volatilities of taxes are within the range of the optimal volatilities for tax instruments considered in the model with search frictions and government debt in Arsenau and Chugh (2012) and the neoclassical model with a balanced budget restriction in Stockman (2001).

³⁴While the literature has not considered optimal income taxation for different levels of income as we do in this paper, it has nevertheless demonstrated that market frictions

Counter-cyclical taxes intensify fluctuations in income, as they amplify the effects of exogenous productivity shocks. Therefore, a government that needs to make use of counter-cyclical taxes over the business cycle, does so with a view to minimise the distortions that they cause. In this setup, it is optimal to minimise such policy distortions to the income of the hand-to-mouth households, by making their tax rate to be the least counter-cyclical. This is because these agents are the most exposed to economic fluctuations. Comparing skilled and unskilled households, it is optimal to least distort the choices of skilled, given the higher complementarities of skilled hours with equipment capital. As a result, the unskilled workers face the most counter-cyclical income tax.

Finally, there are not significant differences regarding the persistence of tax rates. Recall that in the data, all taxes are highly persistent. When chosen optimally, they are somewhat less persistent (particularly τ_s), but they are qualitatively of similar magnitudes.

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

We are now in a position to analyse the optimal reaction of tax policy to different exogenous shocks by examining the impulse responses (IRs) of the key economic variables after a temporary standard deviation shock to each of the exogenous processes. These are plotted in Figures 1 and 2. This allows us to evaluate how the government optimally distributes the tax burden in the short- and medium-run in response to output-enhancing exogenous shocks. Figure 1 concentrates on the optimal taxes and the distribution of the tax burden post shock whereas Figure 2 documents the responses of the remaining endogenous variables.³⁵

Starting with the impulse responses to a positive TFP shock, we see that τ_h is the only tax that increases on impact (see Figure 1). This is because the government, in the short-run, finds it optimal to make this tax pro-cyclical to facilitate a smoother consumption response from hand-to-mouth households.

do matter for the optimal level of cyclicality of income taxes. For instance, while in the neoclassical model the labour income tax is generally positively related with the TFP shocks and thus tends to be pro-cyclical, in the labour market frictions model of Arsenau and Chugh (2012), the labour income tax is optimally counter-cyclical.

³⁵Figure 1 plots the optimal response of the tax rates in: (i) levels (row 1); (ii) percent deviations from the steady-state (row 2); and (iii) differences between the relevant percent deviations (row 3). Thus, row 3 shows how the relative tax burden is distributed in response to the shocks considered. Figure 2, as in row 2 of Figure 1 plots the standard impulse responses.

As can also be seen, in the short-run both τ_s and τ_u fall, as the government can reduce these tax rates and maintain the same tax revenue, given the rise in tax bases. The reduction is bigger for the tax rate on unskilled labour. Therefore, the response of optimal income taxes with respect to the progressivity of the tax system is mixed after a positive TFP shock, as, on one hand, the tax rate on the low income is increased, while, on the other, the middle-income households benefit from the biggest reduction.

Moreover, the dynamic paths for the taxes show that the fall in τ_s is temporary, as in the medium-run the government finds it optimal to increase τ_s above its steady-state. As the economy returns to the steady-state, it is optimal for the government to generate the tax revenue required by increasing faster the tax rate for the largest income source, so that it can maintain a low tax rate for unskilled and hand-to-mouth for longer. Figure 2 shows that the increase in equipment capital and decrease in relative skill supply induced by the changes in TFP and income taxation increase the skill premium. As a result, the skill premium is procyclical.

[Figures 1 and 2 here]

The response of optimal taxes is very different under a positive capital equipment technology shock. In particular, this shock creates an increase in wage inequality (see Figure 2), reflecting the rise in the productivity of skilled labour via capital-skill complementarity, to which the government responds, by increasing the tax for skilled workers and decreasing the tax for unskilled workers.³⁶ Hence, in this case, the government finds it optimal, given the exogenous productivity gains for skilled households, to redistribute the tax burden in favour of the unskilled households. As a result, income taxation becomes more progressive.

After an output-enhancing temporary reduction to government spending all taxes can fall, since the government needs to generate less tax revenue. However, driven by the incentive to support the households who are most constrained in the asset market, the reduction is bigger for τ_h , followed by that for τ_u . Therefore, a temporary spending cut is followed by an increase in the progressivity of income taxes. The reductions in the income tax rates lead to a proportionately larger increase for the supply of unskilled labour, relative to skilled, which in turn increases the skill premium (see Figure 2).

³⁶Note that in this case the amount of equipment capital is reduced after a positive capital equipment technological shock (see Figure 2). This result is due the important income gains accruing to skilled workers by the increase in the returns to this asset. This allows them to reduce investment in this asset and increase consumption. A similar response of equipment capital is obtained in Lindquist (2004).

4.3.1 Robustness of results

We further examine the optimal reaction of tax policy to different exogenous shocks under different assumptions regarding the quantitative importance of transaction costs and capital-skill complementarity. In Figure 3, we show, for each case, the optimal path of the three income tax rates starting from the non-stochastic steady state. For comparison, we also present the optimal paths under the benchmark case analysed above.

[Figure 3 here]

In particular, we consider three alternative calibrations and present the optimal taxes in each case, in response to exogenous output-enhancing shocks. In Model 1, we consider a case where the transaction costs for skilled and unskilled workers are five times greater than their benchmark value in Table 3. In Model 2, we examine a case where the equipment capital-skill substitutability rises, as ρ is reduced from $\rho = -0.495$ to $\rho = -0.40$.³⁷ In Model 3, we examine the case where the equipment capital-unskilled substitutability falls, as φ is reduced from $\varphi = 0.401$ to $\varphi = 0.30$.³⁸

The general message from Figure 3 is that the responses of optimal policy are quantitatively very similar across the different model variants. Hence, our results regarding the optimal reaction to exogenous shocks are robust to changes in the magnitude of transaction costs and production function elasticities.

5 Conclusions

In this paper, we studied how the progressivity of the income tax system optimally changed over the business cycle in an economy characterised by capital-skill complementarity, a wage premium to skilled workers and market imperfections that create unequal opportunities to households regarding their participation in the labour and capital markets. In particular, we assumed that a subset of households provided skill labour services, whereas the rest worked as unskilled. Moreover, participation premia in the capital 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

³⁷This requires that we re-calibrate $\lambda = 0.5676$ and $\nu = 0.5243$, so that the factor shares remain the same and the production function is uniquely identified (see e.g. Cantore and Levine, 2012).

³⁸In this case we re-calibrate $\lambda = 0.5727$ and $\nu = 0.5808$.

was consistent with key features of the labour markets and wealth ownership that we considered.

Our analysis considered the problem of a government that chose the paths of income tax rates to maximise aggregate welfare. In doing so we constrained the policy menu of the government to focus on income taxes that applied to total household income only. We then characterised the business cycle properties of optimal income taxes under these policy constraints and examined the optimal reaction of income taxation under different exogenous shocks.

With respect to business cycle properties, we found that the volatility and co-movement of optimal taxes with output differed significantly with each other and with the data. In particular, optimal policy kept the tax rate for skilled agents the least volatile and the tax rate for the hand-to-mouth agents the least countercyclical. In contrast, the path of optimal taxes for the middle income group, made up by unskilled households with savings, was very volatile and counter-cyclical.

Regarding exogenous shocks, the government found it optimal to respond to output-enhancing capital equipment technology and spending cuts by increasing the progressivity of income taxes. On the other hand, the optimal response to positive TFP shocks implied that the progressivity of the tax system increases after about two years.

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., Fernandez, B. and J. Malley (2013a). ‘The Distributional Consequences of Tax Reforms Under Capital-Skill Complementarity’, *Economica*, (in press).
- [4] Angelopoulos, K., Malley, J. and A. Philippopoulos (2013b). ‘Human capital, social mobility and the skill premium’, *CEPrifo Working Paper 4388*.
- [5] Arseneau, D. and S. Chugh (2012). ‘Tax smoothing in frictional labor markets’, *Journal of Political Economy*, 120, 926-985.

- [6] Benigno, P. (2009). ‘Price stability with imperfect financial integration’, *Journal of Money, Credit and Banking*, 41, 121-49.
- [7] Browning, M., Hansen, L. and J. Heckman (1999). ‘Micro data and general equilibrium models’, in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, 1, 543-633, Amsterdam: Elsevier.
- [8] 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.
- [9] Campbell, J. and J. Cochrane (1999). ‘By force of habit: A consumption-based explanation of aggregate stock market behavior’, *Journal of Political Economy*, 107, 205–251.
- [10] Cantore, C. and P. Levine (2012). ‘Getting normalization right: Dealing with ‘dimensional constants’ in macroeconomics’, *Journal of Economic Dynamics and Control*, 36, 1931-1949.
- [11] Castro R. and D. Coen-Pirani (2008). ‘Why have aggregate skilled hours become so cyclical since the mid-1980s?’, *International Economic Review*, 49 (1), 135-184.
- [12] Chari, V., Lawrence, C. and P. Kehoe (1994). ‘Optimal fiscal policy in a business cycle model’, *The Journal of Political Economy*, 102 (4), 617-652.
- [13] Chari, V. and P. Kehoe (1999). ‘Optimal fiscal and monetary policy’, in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, 1, 1671-1745, Amsterdam: Elsevier.
- [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] Christiano, C., Eichenbaum, M. and C. Evans (2005). ‘Nominal rigidities and the dynamic effects of a shock to monetary policy’, *Journal of Political Economy*, 113 (1), 1-45.
- [16] 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.

- [17] Galí, J., Lopez-Salido, D. and J. Valles (2007). ‘Understanding the effects of government spending on consumption’, *Journal of the European Economic Association*, 5 (1), 227-270.
- [18] He, H. (2012). ‘What Drives the Skill Premium: Technological change or demographic variation?’, *European Economic Review*, 56, 1546-1572.
- [19] He, H and Z. Liu (2008). ‘Investment-specific technological change, skill accumulation, and wage inequality’, *Review of Economic Dynamics*, 11, 314-334.
- [20] 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.
- [21] Katz, L. and K. Murphy (1992). ‘Changes in relative wages, 1963-1987: supply and demand factors’, *The Quarterly Journal of Economics*, 107, 35-78.
- [22] Keane, M. and R. Rogerson (2012). ‘Micro and macro labor supply elasticities: A reassessment of conventional wisdom’, *Journal of Economic Literature*, 50 (2), 464-476.
- [23] Kocherlakota, N. (2010). ‘The new dynamic public finance’, Princeton University Press, Princeton, New Jersey.
- [24] Krusell, P., Ohanian, L., Ríos-Rull J. and G. Violante (2000). ‘Capital–skill complementarity and inequality: a macroeconomic analysis’, *Econometrica*, 68, 1029–1053.
- [25] Ljungqvist, L. and H. Uhlig (2000). ‘Tax policy and aggregate demand management under catching up with the Joneses’, *American Economic Review*, 90 (3), 356-366.
- [26] Lindquist, M. (2004). ‘Capital–skill complementarity and inequality over the business cycle’, *Review of Economic Dynamics*, 7, 519-540.
- [27] Litterman, R. (1983). ‘A random walk, Markov model for the distribution of time-series’, *Journal of Business & Economic Statistics*, 1 (2), 169-173.
- [28] Mankiw, G. (2000). ‘The savers-spenders theory of fiscal policy’, *American Economic Review*, 90 (2), 120-125.

- [29] Mankiw, G., Weinzierl, M. and D. Yagan (2009). ‘Optimal taxation in theory and practice’, *Journal of Economic Perspectives*, 23, 147-174.
- [30] Matsuyama K. (2006). ‘The 2005 Lawrence R. Klein Lecture: emergent class structure’, *International Economic Review*, 47, 327-360.
- [31] Persson, T. and G. Tabellini (1992). ‘The politics of 1992: fiscal policy and European integration’, *Review of Economic Studies*, 59, 689-701.
- [32] 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 (1), 3-24.
- [33] Pourpourides, P. (2011). ‘Implicit contracts and the cyclicity of the skill-premium’, *Journal of Economic Dynamics and Control*, 35, 963-979.
- [34] Schmitt-Grohé, S. and M. Uribe (2003). ‘Closing small open economy models’, *Journal of International Economics*, 61, 163-185.
- [35] Schmitt-Grohé, S. and M. Uribe (2007). ‘Optimal simple and implementable monetary and fiscal rules’, *Journal of Monetary Economics*, 54, 1702–1725.
- [36] Stockman, D. (2001). ‘Balanced-budget rules: Welfare loss and optimal policies’, *Review of Economic Dynamics*, 4, 438-459.
- [37] Traum, N. and S. Yang (2010). ‘Does government debt crowd out investment? A Bayesian DSGE approach’, Congressional Budget Office Working Paper 2010-02.

6 Appendix A: DCE system of equations

6.1 Skilled households

FOC for consumption:

$$0 = \frac{[(C_{s,t} - \omega C_{s,t-1})^\gamma (1 - h_{s,t})^{1-\gamma}]^{(1-\sigma)\gamma}}{C_{s,t} - \omega C_{s,t-1}} + \lambda_t^1 \quad (18)$$

FOC for hours worked:

$$0 = -\frac{[(C_{s,t} - \omega C_{s,t-1})^\gamma (1 - h_{s,t})^{1-\gamma}]^{(1-\sigma)(1-\gamma)}}{C_{s,t} - \omega C_{s,t-1}} - \lambda_t^1 (1 - \tau_{s,t}) w_{s,t} \quad (19)$$

FOC for capital structures:

$$0 = \lambda_t^2 + \beta \{ E_t \lambda_{t+1}^1 [(\tau_{s,t+1} - 1) r_{t+1}^q + 2\psi_s K_{s,t+1}^q] + E_t \lambda_{t+1}^2 (\delta^q - 1) \} \quad (20)$$

FOC for capital equipment:

$$0 = \lambda_t^3 + \beta \{ E_t \lambda_{t+1}^1 [(\tau_{s,t+1} - 1) r_{t+1}^e + 2\psi_s K_{s,t+1}^e] + E_t \lambda_{t+1}^3 (\delta^e - 1) \} \quad (21)$$

FOC for investment in capital structures:

$$0 = \lambda_t^1 - \lambda_t^2 \quad (22)$$

FOC for investment in capital equipment:

$$0 = \lambda_t^1 - \lambda_t^3 \quad (23)$$

Capital structures evolution equation:

$$0 = K_{s,t+1}^q - (1 - \delta^q) K_{s,t}^q - I_{s,t}^q \quad (24)$$

Capital equipment evolution equation:

$$0 = K_{s,t+1}^e - (1 - \delta^e) K_{s,t}^e - I_{s,t}^e \quad (25)$$

where λ_t^1 , λ_t^2 and λ_t^3 are respectively, the Lagrange multipliers with respect to the budget constraint as well as the capital structures and capital equipment evolution equations of the skilled agent. Based on Walras' law only $N - 1$ constraints are required in the DCE, hence we drop the budget constraint of the skilled worker.

6.2 Unskilled households

FOC for consumption:

$$0 = \frac{[(C_{u,t} - \omega C_{u,t-1})^\gamma (1 - h_{u,t})^{1-\gamma}]^{(1-\sigma)\gamma}}{C_{u,t} - \omega C_{u,t-1}} + \lambda_t^4 \quad (26)$$

FOC for hours worked:

$$0 = -\frac{[(C_{u,t} - \omega C_{u,t-1})^\gamma (1 - h_{u,t})^{1-\gamma}]^{(1-\sigma)(1-\gamma)}}{C_{u,t} - \omega C_{u,t-1}} - \lambda_t^4 (1 - \tau_{u,t}) w_{u,t} \quad (27)$$

FOC for capital structures:

$$0 = \lambda_t^5 + \beta \{ E_t \lambda_{t+1}^4 [(\tau_{u,t+1} - 1) r_{t+1}^q + 2\psi_u K_{u,t+1}^q] + E_t \lambda_{t+1}^5 (\delta^q - 1) \} \quad (28)$$

FOC for capital equipment:

$$0 = \lambda_t^6 + \beta \{ E_t \lambda_{t+1}^1 [(\tau_{u,t+1} - 1) r_{t+1}^e + 2\psi_u K_{u,t+1}^e] + E_t \lambda_{t+1}^6 (\delta^e - 1) \} \quad (29)$$

FOC for investment in capital structures:

$$0 = \lambda_t^4 - \lambda_t^5 \quad (30)$$

FOC for investment in capital equipment:

$$0 = \lambda_t^4 - \lambda_t^6 \quad (31)$$

Budget constraint:

$$0 = C_{u,t} + I_{u,t}^q + I_{u,t}^e - (1 - \tau_{u,t}) r_t^q K_{u,t}^q - (1 - \tau_{u,t}) r_t^e K_{u,t}^e - (1 - \tau_{u,t}) w_{u,t} h_{u,t} + T_t + \psi_u \left[(K_{u,t}^q)^2 + (K_{u,t}^e)^2 \right] \quad (32)$$

Capital structures evolution equation:

$$0 = K_{u,t+1}^q - (1 - \delta^q) K_{u,t}^q - I_{u,t}^q \quad (33)$$

Capital equipment evolution equation:

$$0 = K_{u,t+1}^e - (1 - \delta^e) K_{u,t}^e - I_{u,t}^e \quad (34)$$

where λ_t^4 , λ_t^5 and λ_t^6 are respectively, the Lagrange multipliers with respect to the budget constraint as well as the capital structures and capital equipment evolution equations of the unskilled agent.

6.3 Hand-to-mouth households

FOC for consumption:

$$0 = \frac{[(C_{h,t} - \omega C_{h,t-1})^\gamma (1 - h_{h,t})^{1-\gamma}]^{(1-\sigma)\gamma}}{C_{h,t} - \omega C_{h,t-1}} + \lambda_t^7 \quad (35)$$

FOC for labour:

$$0 = - \frac{[(C_{h,t} - \omega C_{h,t-1})^\gamma (1 - h_{h,t})^{1-\gamma}]^{(1-\sigma)(1-\gamma)}}{C_{h,t} - \omega C_{h,t-1}} - \lambda_t^7 (1 - \tau_{h,t}) w_{u,t} \quad (36)$$

Budget constraint:

$$0 = C_{h,t} - (1 - \tau_{h,t}) w_{u,t} h_{h,t} + T_t \quad (37)$$

where λ_t^7 is the Langrange multiplier with respect to the budget constraint of the hand-to-mouth agent.

6.4 Firms

FOC for capital structures:

$$0 = \frac{\alpha Y_t}{K_t^{f,q}} - r_t^q \quad (38)$$

FOC for capital equipment:

$$0 = - \frac{K_t^{f,e} \lambda \nu Y_t}{K_t^{f,e} \Omega_t^1 (\Omega_t^1)^{\frac{\alpha-1}{\varphi}} \left[(A_t^e)^\rho (K_t^{f,e})^\rho + (1 - \nu) (h_{s,t}^f)^\rho \right]} - r_t^e \quad (39)$$

where $\Omega_t^1 \equiv \lambda \left\{ \nu \left[(A_t^e)^\rho (K_t^{f,e})^\rho \right] + (1 - \nu) (h_{s,t}^f)^\rho \right\}^{\varphi/\rho} + (1 - \lambda) (h_{u,t}^f)^\varphi$.

FOC for skilled labour hours:

$$0 = \frac{A_t (K_t^{f,q})^\alpha \lambda (1 - \nu) (1 - \alpha) (h_{s,t}^f)^\rho \Omega_t^2}{h_{s,t}^f \Omega_t^1 (\Omega_t^1)^{\frac{\alpha-1}{\varphi}} \left[\nu (A_t^e)^\rho (K_t^{f,e})^\rho + (1 - \nu) (h_{s,t}^f)^\rho \right]} - w_{s,t} \quad (40)$$

where $\Omega_t^2 \equiv \left\{ \nu \left[(A_t^e)^\rho (K_t^{f,e})^\rho \right] + (1 - \nu) (h_{s,t}^f)^\rho \right\}^{\varphi/\rho}$.

FOC for unskilled labour hours:

$$0 = \frac{A_t (K_t^{f,q})^\alpha (1 - \alpha) (1 - \lambda) (n_u h_{u,t} + n_h h_{h,t})^\varphi}{\Omega_t^1 (\Omega_t^1)^{\frac{\alpha-1}{\varphi}} (n_u h_{u,t} + n_h h_{h,t})} - w_{u,t} \quad (41)$$

6.5 Government budget constraint

$$0 = G_t^c - T_t - n_h \tau_{h,t} w_{u,t} h_{h,t} - n_s \tau_{s,t} (w_{s,t} h_{s,t} + r_t^q K_{s,t}^q + r_t^e K_{s,t}^e) - n_u \tau_{u,t} (w_{u,t} h_{u,t} + r_t^q K_{u,t}^q + r_t^e K_{u,t}^e) \quad (42)$$

6.6 Aggregate Resource Constraint

$$0 = Y_t - G_t^c - n_s C_{s,t} - n_u C_{u,t} - n_h C_{h,t} - n_s (I_{s,t}^q + I_{s,t}^e) - n_u (I_{u,t}^q + I_{u,t}^e) - \psi_s \left((K_{s,t}^q)^2 + (K_{s,t}^e)^2 \right) - \psi_u \left((K_{u,t}^q)^2 + (K_{u,t}^e)^2 \right) \quad (43)$$

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

6.7 Market clearing conditions

$$h_{s,t}^f = n_s h_{s,t} \quad (44)$$

$$h_{u,t}^f = n_u h_{u,t} + n_h h_{h,t} \quad (45)$$

$$K_t^{f,e} = n_s K_{s,t}^e + n_u K_{u,t}^e \quad (46)$$

$$K_t^{f,q} = n_s K_{s,t}^q + n_u K_{u,t}^q \quad (47)$$

7 Appendix B: The skill premium

Using equations (40) and (41) implies the following expression for the skill premium:

$$\frac{w_{s,t}}{w_{u,t}} = \frac{\lambda (1 - \nu) \left(h_{s,t}^f \right)^{\rho-1}}{(1 - \lambda) \left(h_{u,t}^f \right)^{\varphi-1}} \left(\Xi_t^1 \right)^{\varphi/\rho-1}. \quad (48)$$

where $\Xi_t^1 \equiv \nu \left[(A_t^e)^\rho \left(K_t^{f,e} \right)^\rho \right] + (1 - \nu) \left(h_{s,t}^f \right)^\rho$.

The skill premium is increasing with respect to capital equipment as long as the equipment-skill complementarity is present, i.e. $\rho < 0$, and the unskilled agents are substitutes to both of them, i.e. $0 < \varphi < 1$. Also,

for the $\frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(K_t^{f,e}\right)} > 0$ condition to hold it is necessary that $0 < \lambda, \nu < 1$ which is satisfied through our calibration as in Krusell *et al.* (2000):

$$\frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(K_t^{f,e}\right)} = \frac{\lambda(1-\nu)}{(1-\lambda)} \frac{\left(h_{s,t}^f\right)^{\rho-1}}{\left(h_{u,t}^f\right)^{\varphi-1}} (\varphi - \rho) \nu (A_t^e)^\rho \left(K_t^{f,e}\right)^{\rho-1} \left(\Xi_t^1\right)^{\varphi/\rho-2}. \quad (49)$$

Moreover the skill premium is decreasing to skilled labour supply, $\frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(h_{s,t}^f\right)} < 0$:

$$\begin{aligned} \frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(h_{s,t}^f\right)} &= \frac{\lambda(1-\nu)}{(1-\lambda)} \frac{\left(h_{s,t}^f\right)^{\rho-1}}{\left(h_{u,t}^f\right)^{\varphi-1}} \left(\Xi_t^1\right)^{\varphi/\rho-1} \times \\ &\times \left[\rho + (\rho - \varphi)(1 - \nu) \left(h_{s,t}^f\right)^{\rho-1} \right] \end{aligned} \quad (50)$$

where the terms that define the sign are: $\frac{\lambda(1-\nu)}{(1-\lambda)} > 0$ and the term inside the squared brackets $\rho + (\rho - \varphi)(1 - \nu) \left(h_{s,t}^f\right)^{\rho-1} < 0$ (due to the fact that $\rho < 0$ and $0 < \varphi < 1$).

Also, the skill premium is increasing to unskilled labour supply, $\frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(h_{u,t}^f\right)} > 0$:

$$\begin{aligned} \frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(h_{u,t}^f\right)} &= \frac{\lambda(1-\nu)}{(1-\lambda)} (1 - \varphi) \frac{\left(h_{s,t}^f\right)^{\rho-1}}{\left(h_{u,t}^f\right)^\varphi} \times \\ &\times \left\{ \nu \left[(A_t^e)^\rho \left(K_t^{f,e}\right)^\rho \right] + (1 - \nu) \left(h_{s,t}^f\right)^\rho \right\}^{\varphi/\rho-1} \end{aligned} \quad (51)$$

now the crucial terms are: $\frac{\lambda(1-\nu)}{(1-\lambda)} (1 - \varphi)$, where $\frac{\lambda(1-\nu)}{(1-\lambda)} > 0$ and also $(1 - \varphi)$ since $0 < \varphi < 1$.

The last two derivatives, (50 – 51), imply that the skill premium is decreasing with respect to the relative labour supply of skilled over unskilled agents, i.e. $\frac{\partial\left(\frac{w_{s,t}}{w_{u,t}}\right)}{\partial\left(\frac{h_{s,t}^f}{h_{u,t}^f}\right)} < 0$.

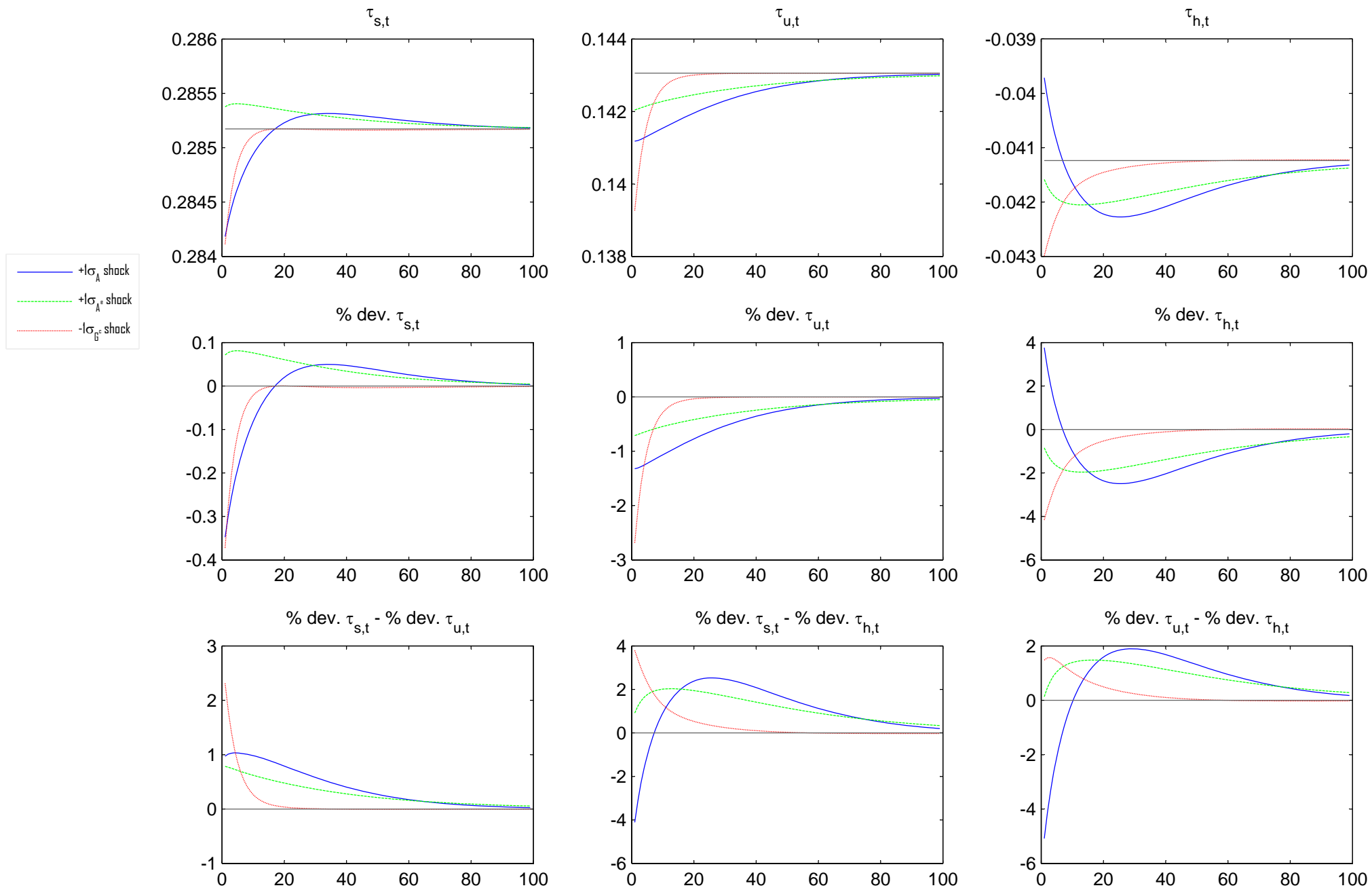


Figure 1: Optimal distribution of the tax burden

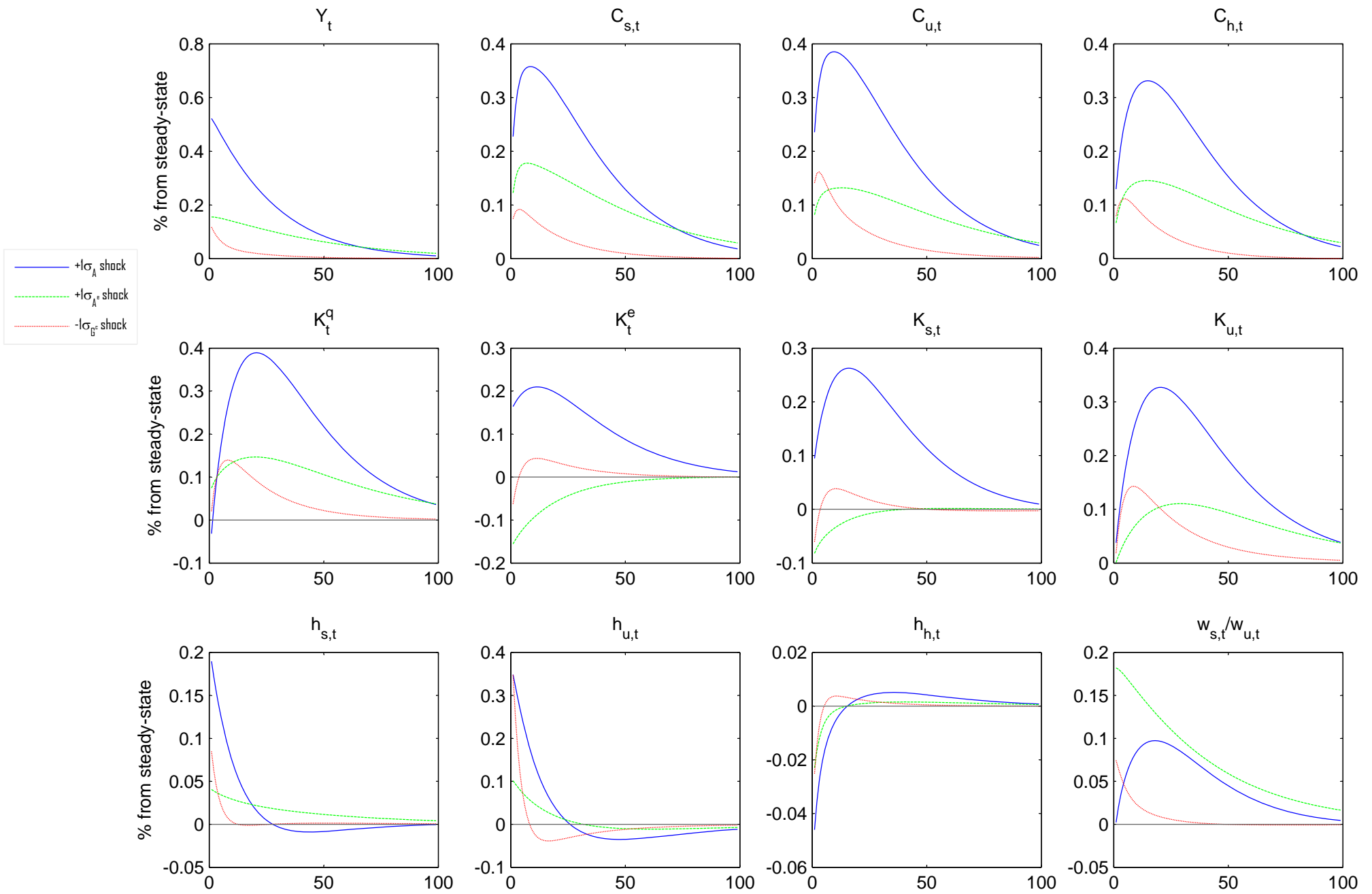


Figure 2: Impulse responses (optimal policy)

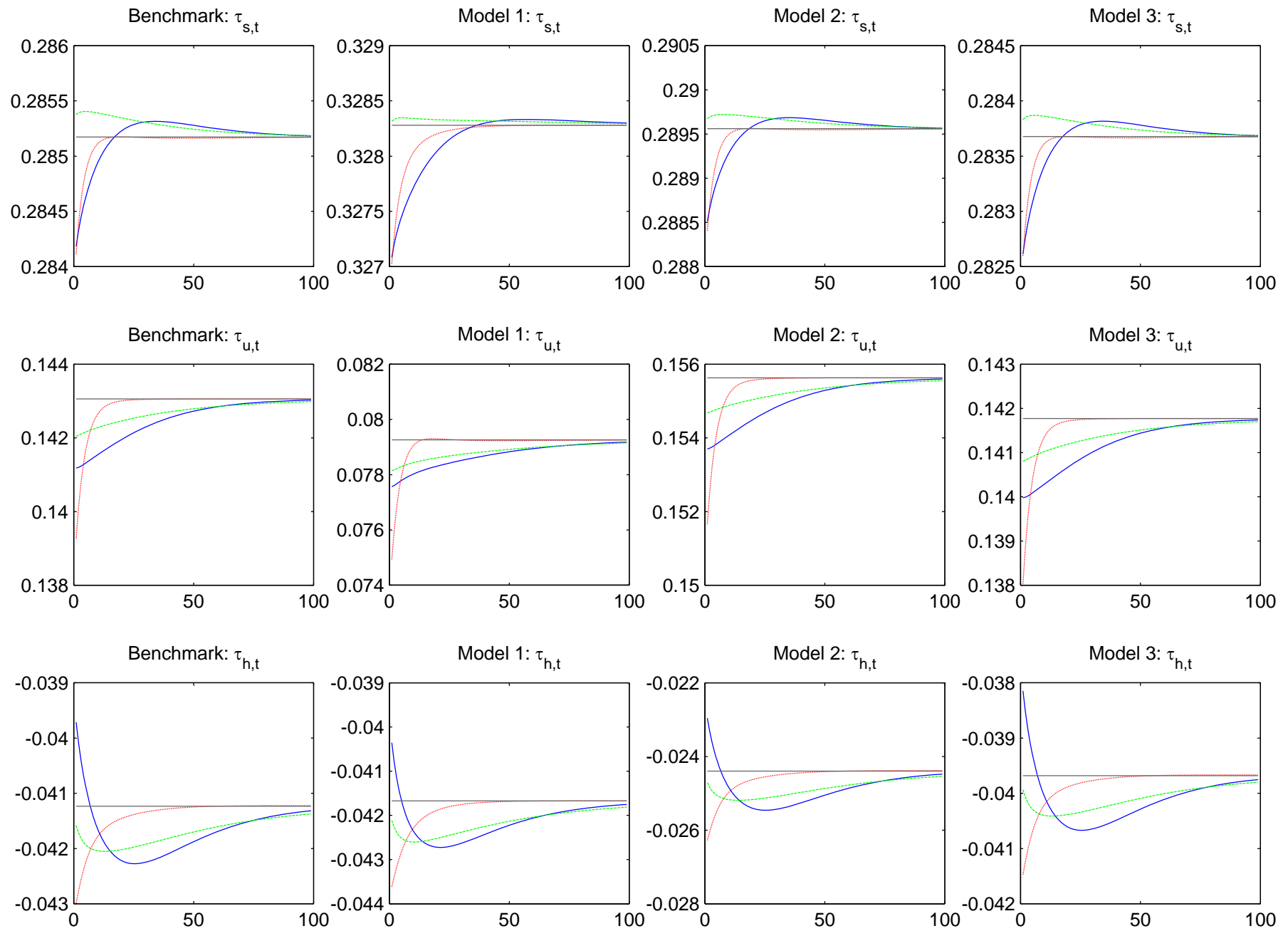


Figure 3: Optimal tax rates for benchmark model and alternative calibrations