CFCM

CENTRE FOR FINANCE, CREDIT AND MACROECONOMICS

Working Paper 14/12

Optimal progressive taxation in a model with endogenous skill supply

Konstantinos Angelopoulos, Stylianos Asimakopoulos, James Malley

Produced By:

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

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

Optimal progressive taxation in a model with endogenous skill supply*

Konstantinos Angelopoulos
University of Glasgow

Stylianos Asimakopoulos
University of Nottingham

James Malley
University of Glasgow and CESifo

September 29, 2014

Abstract

This paper examines quantitatively the extent of progressivity or regressivity of optimal labour income taxation in a model with skill heterogeneity, endogenous skill acquisition and a production sector with capital-skill complementarity. We find that wage inequality driven by the resource requirements of skill-creation implies progressive labour income taxation in the steady-state as well as along the transition path from the exogenous to optimal policy steady-state. In particular, in the steady state, skilled labour income is taxed about 40% more than unskilled labour income. We further find that these results are explained by a lower work time elasticity for skilled versus unskilled labour which results from the introduction of the skill acquisition technology.

Keywords: optimal progressive taxation, skill premium, allocative efficiency
JEL Classification: E24, E32, E62
Corresponding author: jim.malley@glasgow.ac.uk

*We would like to thank Fabrice Collard, Guido Cozzi, Pedro Gomes, Charles Nolan, Apostolis Philippopoulos, Peter Rosenkranz, Peter Birch Sørenson, Ulrich Woitek, participants at the Crestasee Workshop, University of Zurich, July 2014 and the CESifo Workshop on Reforming the Public Sector, Venice, July 2014. for helpful comments and suggestions.

1 Introduction

The literature on optimal taxation has examined extensively the question of the optimal progressivity of the tax system in environments with heterogeneous agents and income inequality (see e.g. Mirrlees (1971), Diamond (1998), Saez (2001) and Kocherlakota (2010)). This framework is mainly chosen to capture the key trade-off underpinning the choice of optimal progressive taxation, namely equity versus efficiency. On one hand, equity ambitions typically prescribe progressivity of the tax system, while, on the other hand, efficiency goals are generally associated with regressive tax structures. The literature has also found that in some circumstances progressive taxation may improve resource allocation by correcting for an underlying market failure. For instance, when lower income is related to market exclusion, redistributive taxation may increase economic growth by increasing participation and economic performance (see e.g. Drazen (2000)).

In contrast to the studies referred to above, this paper examines quantitatively the optimal progressivity or regressivity of labour income taxation in a model without redistributive motives. To this end we employ a representative agent setup without market failures, incorporating skill heterogeneity, capital-skill complementarity, endogenous skill acquisition and wage inequality. Our interest in this question is motivated by the empirical relevance of both wage inequality and the proposed model structure. For example, following reductions in earnings inequality in the U.S. for most of the 20th century, this trend has reversed since 1980 such that the wage premium for skilled workers is at its highest level since 1910 (see e.g. Goldin and Katz (2008)).[1] Additionally, Goldin and Katz (2008) provide historical evidence for the 20th century demonstrating that wage inequality has developed within a production sector characterised by capital-skill complementarity.

As discussed above, optimal progressive taxation generally follows from equity considerations and may lead to increased efficiency in the presence of market failures. However, while we generally expect some form of regressive taxation for efficiency reasons, the implications of income taxation for resource allocation ultimately depend on the structure of the underlying economy and, in particular, on the effects of taxes on the optimal reactions of the private economic agents. Thus, in the context of the empirically relevant analytical framework sketched out above, this paper concentrates on the efficiency incentives of optimal taxation by employing perfect capital and labour markets to derive the Ramsey plan that minimises tax distortions. As

[1] Given the importance of these developments, an extensive literature has studied wage differentials between college and high school graduates (see e.g. Acemoglu and Autor (2011) Goldin and Katz (2008) and Hornstein *et al.* (2005) for reviews).

is common in the public finance literature of Ramsey optimal taxation, the requirement to tax will be exogenously imposed on the government, which is assumed to have access to a commitment technology. Additionally, since we do not allow policy makers to have access to lump-sum policy instruments, we focus on the *second-best* Ramsey problem.

In light of the above, we calculate optimal factor income taxation in an environment embodying two types of labour services (skilled and unskilled), two types of capital (structures and equipment) and endogenous acquisition of skill. We employ the production technology in Krusell *et al.* (2000) and also used in e.g. Lindquist (2004), He and Liu (2008) and Pourpourides (2011), since this has been shown to provide a good match to the data. This technology specifies that equipment capital complements skilled labour more than unskilled so that changes in its accumulation are skill biased.

Our analysis of skill acquisition and capital-skill complementarity builds on and extends the model in He and Liu (2008).[2] In particular, we assume that a representative household decides how to allocate its expenditure into investment in the two types of capital stock and into goods for creating skilled labour. Moreover, it decides how to allocate its time endowment into leisure, work time in skill and unskilled jobs, and in education or training for creating skilled labour. The technology assumed for the creation of skilled hours follows a standard Cobb-Douglas form, which allows the model to capture the goods and time opportunity costs of creating skilled labour services. The resource requirements associated with skill acquisition in turn imply that there is a wage premium accruing to skilled labour to compensate for these costs.

In other recent work, Angelopoulos *et al.* (2014), we analyse optimal tax smoothing under skill heterogeneity and capital-skill complementarity, when the government has access to state-contingent debt and a complete set of state-contingent tax instruments. This is carried out in a stochastic environment with endogenous and exogenous skill supply by different workers, under externalities in skill creation. In contrast, our interest here is in optimal factor return taxation in a deterministic environment with a representative worker, without market failures, both in the long-run as well along the transition to the Ramsey steady-state. In particular, by first focusing on the long-run under Ramsey policy, we examine the degree of optimal labour income tax progressivity. Second, by calibrating the model under exogenous policy to data averages for the U.S., we calculate the optimal transition paths

[2] The model in He and Liu (2008) provides a useful framework in which they study policy reforms in the presence of wage inequality (see also Angelopoulos *et al.* (2013)). Since our aim here is to analyse optimal policy, we modify their model to allow for an endogenous labour-leisure choice, which is necessary when examining optimal labour taxes.

for policy and allocations from the exogenous policy economy to the Ramsey steady-state.

In contrast to general expectations, when only efficiency motives maintain, we find that wage inequality driven by the resource requirements of skill-creation implies progressive labour income taxation. In particular, in the steady state, skilled labour income is taxed about 40% more than unskilled labour income. We further find that this is explained by the lower elasticity for skilled labour relative to that of unskilled. The intuition for this result is that the resources employed for creating skilled labour generates additional opportunity costs for the provision of skilled labour, which act to reduce the responsiveness of skilled work time when the tax on skilled labour income changes. When the model is calibrated to U.S. data, these effects, on balance, lower the skilled labour elasticity relative to that for the unskilled. Thus they create an incentive for the Ramsey planner to tax skilled labour income more than unskilled for efficiency purposes.

We next find that the Ramsey plan requires capital taxes to be set very high in the first period and then rapidly decrease towards zero, as is common in the literature on optimal capital taxation (see e.g. Chamley (1986) and Chari and Kehoe (1999)). By contrast, both labour income taxes turn into subsidies in the first period, before converging to their steady-state levels. Notably, the tax system becomes progressive from the first period. As is also common in the optimal taxation literature, the government runs big surpluses in the first period. This allows it to create a stock of assets, which is in turn used to finance primary deficits in the future. Finally, it is worth noting the Ramsey plan implies a sharp increase in wage inequality in the first periods, before the skill premium returns effectively to its initial level. However, the increase in tax progressivity implies that net wage inequality is reduced under Ramsey policy.

The rest of the paper is organised as follows. Section 2 presents the theoretical model. Section 3 first specifies the functional forms used for production, utility and skill acquisition, followed by the model calibration and the steady-state solution under exogenous fiscal policy. Section 4 solves the Ramsey model and discusses the steady-state results for optimal policy together with the transition paths of the optimal policy instruments and allocations. Finally, section 5 contains the conclusions.

2 The model

The economy is populated by a representative household which supplies skilled and unskilled labour services. Following He and Liu (2008) skilled labour requires the creation of skill, which is determined by time and goods. There is also a representative firm that uses two types of capital and two types of labour for the production of a homogeneous product. Following Krusell *et al.* (2000), skilled labour is assumed to be more complementary to capital equipment than unskilled labour. Thus, capital equipment accumulation is skill biased and tends to increase the skill premium, defined as the ratio of the skilled to unskilled wage. In contrast, increases in the relative supply of skilled labour tend to reduce the skill premium. Since provision of skilled labour comes at a cost to the household, a wage premium for skilled labour is required in equilibrium to maintain net wage parity. Finally, the government finances exogenous public spending by issuing debt, taxing all sources of income and subsidising investment in skill acquisition.

2.1 The representative household

2.1.1 Utility

The lifetime utility of the representative household is given by:

$$U = \sum_{t=0}^{\infty} \beta^t u(C_t, l_t) \tag{1}$$

where $0 < \beta < 1$ is a constant discount factor and denotes the time preference of the individual; C_t and l_t are consumption and leisure respectively at period t; and $u(\cdot)$ is increasing in its arguments, strictly concave and three times continuously differentiable.

2.1.2 Constraints

The representative household faces the following time constraint:

$$1 = l_t + h_t^s + h_t^u + e_t \tag{2}$$

where h_t^s and h_t^u denote skilled and unskilled labour work time respectively in period t and e_t is time devoted to education or other training for skills acquisition in period t.

The skill acquisition function is given by:

$$h_t^s = g(E_t^g, e_t) \tag{3}$$

where E_t^g is expenditure on creating skills, and $g(.)$ is increasing in its arguments, strictly concave and three times continuously differentiable.

The law of motion for the two types of capital stock , $i = p, e$, where p and e denote plant and equipment capital respectively, is given by:

$$K_{t+1}^{i} = (1 - \delta^{i})K_{t}^{i} + I_{t}^{i}. \tag{4}$$

The depreciation rate is denoted $0 \leq \delta_i \leq 1$ and $I_{i,t}$ is the investment in period t.

Finally, the household has the following budget constraint equating total expenditure with total income in period t:

$$\begin{aligned} &C_t + I_t^p + I_t^e + (1 - s_t^g)\, E_t^g + \tfrac{b_{t+1}}{R_t^b} = (1 - \tau_t^s) w_t^s h_t^s + \\ &+ (1 - \tau_t^u) w_t^u h_t^u + (1 - \tau_t^p) r_t^p K_t^p + (1 - \tau_t^e) r_t^e K_t^e + b_t \end{aligned} \tag{5}$$

where, $\frac{b_{t+1}}{R_t^b}$ is the discounted value of bonds bought by the household at start of period t; $R_t^b \equiv \left(1 + r_t^b\right)$ is the gross return to bonds; b_t is the payout value of bonds bought at period $t-1$; s_t^g is a subsidy for spending on goods for skill acquisition; τ_t^s, τ_t^u, τ_t^p and τ_t^e are the tax rates on skilled and unskilled labour income as well as plant and equipment capital income in period t respectively.

2.1.3 First-order conditions

Using equation (4) for $i = p, e$ to substitute out plant and equipment investment, the Lagrangian for the household problem is:

$$\begin{aligned} \mathcal{L} = \max \sum_{j=0}^{\infty} &\beta^j u(C_{t+j}, l_{t+j}) + \Lambda_{t+j}\{C_{t+j} + K_{t+j+1}^p - \\ &- (1 - \delta^p) K_{t+j}^p + K_{t+j+1}^e - (1 - \delta^e) K_{t+j}^e + (1 - s_t^g)\, E_t^g + \\ &+ \frac{b_{t+j+1}}{R_{t+j}^b} - (1 - \tau_{t+j}^s) w_{t+j}^s h_{t+j}^s - (1 - \tau_{t+j}^u) w_{t+ij}^u h_{t+j}^u - \\ &- \left[(1 - \tau_{t+j}^p) r_{t+j}^p\right] K_{t+j}^p - [(1 - \tau_{t+j}^e) r_{t+j}^e] K_{t+j}^e - b_{t+j}\} + \\ &+ M_{t+j}[h_t^s - g(E_{t+j}^g, e_{t+j})] \end{aligned} \tag{6}$$

where from the time constraint, $l_{t+j} = 1 - h_{t+j}^s - h_{t+j}^u - e_{t+j}$. The representative household chooses $\{C_t, h_t^s, h_t^u, e_t, E_t^g, K_{t+1}^p, K_{t+1}^e, b_{t+1}\}_{t=0}^{\infty}$ given prices and taxes to maximize equation (6) which gives respectively the following first-order conditions (FOCs):

$$U_{C_t} = -\Lambda_t \tag{7}$$

$$U_{h_t^s} = \Lambda_t (1 - \tau_t^s) w_t^s - M_t \tag{8}$$

$$U_{h_t^u} = \Lambda_t(1-\tau_t^u)w_t^u \tag{9}$$

$$U_{e_t} = M_t g_{e_t} \tag{10}$$

$$g_{E_t^g} = \frac{\Lambda_t}{M_t}(1-s_t^g) \tag{11}$$

$$\Lambda_t = \beta\left\{\Lambda_{t+1}\left[r_{t+1}^p(1-\tau_{t+1}^p)+(1-\delta^p)\right]\right\} \tag{12}$$

$$\Lambda_t = \beta\left\{\Lambda_{t+1}\left[(r_{t+1}^e(1-\tau_{t+1}^e)+(1-\delta^e)\right]\right\} \tag{13}$$

$$\Lambda_t = \beta\Lambda_{t+1}R_t^b \tag{14}$$

where U_x and g_x are the derivatives of the utility and skill accumulation functions with respect to the relevant variable denoted generically as x; and Λ_t and M_t are the Langrange multipliers associated with the budget constraint and the skill acquisition equation respectively.

These equilibrium conditions imply first, that the marginal utility of consumption, U_{C_t}, is equal to shadow price of the budget constraint, Λ_t which measures the change in maximised utility, when the constraint is relaxed by a unit. Second, the marginal disutility of skilled/unskilled work time, $U_{h_t^s}$ and $U_{h_t^u}$, are equal to the net of tax return to skilled/unskilled work, $(1-\tau_t^s)w_t^s$ and $(1-\tau^u)w_t^u$ respectively valued by the shadow price, Λ_t. Additionally, the return to skilled work is also net of the shadow price, M_t, of the skilled employment constraint, capturing the valuation, in utility terms, of the cost of creating skill. Third, the marginal disutility of education, U_{e_t}, is equal to the marginal increase in skilled employment due to a unit increase in education time, g_{e_t}, valued by the shadow price, M_t. Fourth, the marginal increase in skilled employment for a unit increase in goods expenditure, $g_{E_t^g}$, is equal to the ratio of shadow prices, $\frac{\Lambda_t}{M_t}$, net of the subsidy to goods invested to create skill, $(1-s_t^g)$. Finally the last three conditions, equate the marginal utility of consumption in period t, Λ_t, with discounted marginal utility of consumption in period $t+1$, $\beta\Lambda_{t+1}$, which includes the consumption due to saving in plant/equipment capital net of taxes and depreciation, and bonds.

By combining (8) with (9), and noting that $U_{h_t^s} = U_{h_t^u}$, we see that the return to skilled labour net of tax and the cost for skill acquisition, must equal in equilibrium, the net of tax return to unskilled labour:

$$(1-\tau_t^s)w_t^s - \frac{M_t}{\Lambda_t} = (1-\tau_t^u)w_t^u. \tag{15}$$

In other words, wage parity requires that the net returns to an hour in either skilled or unskilled labour are equalised. Therefore, in an economy without market failures, the skill premium is the compensation to skilled labour for the opportunity cost of acquiring skills.

We next substitute the condition relating to the return to bonds (14) into (12) and (13) to obtain:

$$R_t^b = r_{t+1}^p(1-\tau_{t+1}^p) + (1-\delta^p) \tag{16}$$

$$R_t^b = r_{t+1}^e(1-\tau_{t+1}^e) + (1-\delta^e). \tag{17}$$

These define the no-arbitrage conditions for capital and bonds ensuring that the three assets have the same rate of return in equilibrium. Finally, the following transversality conditions for $i = p, e$ must hold for the economy to reach a stationary equilibrium:

$$\lim_{t\to\infty} \beta^t U_{C_t} \frac{b_{t+1}}{R_t^b} = 0 \tag{18}$$

$$\lim_{t\to\infty} \beta^t U_{C_t} K_{t+1}^i = 0. \tag{19}$$

2.2 The representative firm

The representative firm produces a homogeneous consumption good, Y_t, using skilled, $\widetilde{h}_t^s$, and unskilled, $\widetilde{h}_t^u$, labour as well as plant, $\widetilde{K}_t^p$, and equipment, $\widetilde{K}_t^e$, capital. Acting in perfectly competitive factor markets, taking prices, policy and exogenous variables as given, the firm maximises its profits, Π_t:

$$\max_{\widetilde{h}_t^s, \widetilde{h}_t^u, \widetilde{K}_t^p, \widetilde{K}_t^e} \Pi_t = Y_t - w_t^s \widetilde{h}_t^s - w_t^u \widetilde{h}_t^u - r_t^p \widetilde{K}_t^p - r_t^e \widetilde{K}_t^e \tag{20}$$

subject to a Krusell *et al.* (2000) type production function:

$$Y_t = f\left(\widetilde{h}_t^s, \widetilde{h}_t^u, \widetilde{K}_t^p, \widetilde{K}_t^e\right) \tag{21}$$

where $f(\cdot)$ is homogenous of degree one; a $\tilde{}$ over a variable denotes firm quantities; w_t^s and w_t^u are the returns to skilled and unskilled labour respectively; and r_t^e and r_t^p are the returns to capital holdings in equipment and structures respectively.

Choosing the optimal amount of hours of skilled and unskilled labour to hire and the optimal quantity of plant and equipment capital to rent yields the following first-order conditions:

$$w_t^s - f_{\widetilde{h}_t^s} = 0 \tag{22}$$

$$w_t^u - f_{\widetilde{h}_t^u} = 0 \tag{23}$$

$$r_t^p - f_{\widetilde{K}_t^p} = 0 \tag{24}$$

$$r_t^e - f_{\widetilde{K}_t^e} = 0 \tag{25}$$

equating the returns to each factor to their respective marginal products. Given the structure employed, profits are zero in equilibrium.

2.3 Government budget constraint

The government's budget constraint in each period is:

$$G^c + s_t^g E_t^g + b_t = \tau_t^s w_t^s h_t^s + \tau_t^u w_t^u h_t^u + \tau_t^p r_t^p K_t^p + \tau_t^e r_t^e K_t^e + \frac{b_{t+1}}{R_t^b} \tag{26}$$

and states that expenditure on public consumption, G^c, the subsidy to spending on education and repayments on existing debt (issued at the start of period $t-1$) must be equal to the revenues from taxing labour and capital income plus the discounted value of new debt issued at the start of period t.

2.4 Market clearing conditions

Output can be used for public and private consumption, plant and equipment investment as well as goods spending to acquire skills, implying the following aggregate resource constraint:

$$Y_t = G^c + C_t + I_t^p + I_t^e + E_t^g. \tag{27}$$

Additionally, the market clearing conditions in the capital and labour markets are given by:

$$\widetilde{K}_t^p = K_t^p \tag{28}$$

$$\widetilde{K}_t^e = K_t^e \tag{29}$$

$$\widetilde{h}_t^s = h_t^s \tag{30}$$

$$\widetilde{h}_t^u = h_t^u. \tag{31}$$

2.5 Decentralised competitive equilibrium

The decentralised competitive equilibrium (DCE) with exogenous policy is summarized by a sequence of allocations $\{C_t, h_t^s, h_t^u, e_t, E_t^g, K_{t+1}^p, K_{t+1}^e, \widetilde{h}_t^s, \widetilde{h}_t^u, \widetilde{K}_t^p, \widetilde{K}_t^e\}_{t=0}^{\infty}$, the residual policy instrument $\{b_{t+1}\}_{t=0}^{\infty}$ and prices $\{w_t^s, w_t^u, r_t^p, r_t^e\}_{t=0}^{\infty}$ such that the representative household solves its optimisation problem and the firm maximizes profits, taking prices, tax rates, initial conditions for capital and debt, and fixed G^c as given; the government budget constraint is satisfied and all markets clear. The DCE system is presented in the Appendix, see equations (58)-(69).

3 Quantitative analysis exogenous policy

In this section we first specify the functional forms for production, utility and skill acquisition. We then calibrate the exogenous policy model using annual U.S. data for the period 1970-2011 and solve for the steady-state.

3.1 Functional forms

The production function follows the CES form as in Krusell *et al.* (2000):

$$\begin{aligned} Y_t = {} & A\left(\widetilde{K}_t^p\right)^{\alpha}\{\lambda\left[\nu\left(A^e\widetilde{K}_t^e\right)^{\rho}+(1-\nu)\left(\widetilde{h}_t^s\right)^{\rho}\right]^{\varphi/\rho}+ \\ & +(1-\lambda)\left(\widetilde{h}_t^u\right)^{\varphi}\}^{\frac{1-a}{\varphi}} \end{aligned} \tag{32}$$

where, $-\infty < \varphi, \rho < 1$; $0 < a, \lambda, \nu < 1$. The parameters φ and ρ determine 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. The elasticity of substitution between equipment capital and skilled labour is given by $1/(1-\rho)$.[3] The parameters a, λ, ν denote the factor shares and finally, $A > 0$ and $A^e > 0$ are the total factor productivity and capital equipment augmenting technology parameters respectively.

The utility function follows the CRRA form in Chari *et al.* (1994):

$$u(C_t, l_t) = \frac{\left(C_t^{\gamma} l_t^{1-\gamma}\right)^{1-\sigma}}{1-\sigma} \tag{33}$$

where $(\sigma, \gamma) > 0$ represent the preference parameters of the representative household. Specifically, γ determines the weight given to consumption, and σ is the relative risk aversion coefficient.

Finally, the skill acquisition equation is a variant of the form used in He and Liu (2008):

$$h_t^s = g(E_t^g, e_t) = S\left[\left(E_t^g\right)^{\phi}\left(e_t\right)^{1-\phi}\right]^{\xi} \tag{34}$$

where the shares of goods and time in the creation of skills are given by ϕ and $1-\phi$ respectively, with $0 < \phi < 1$. The parameter $S > 0$ determines the efficiency of the skill-creation process. Finally, $0 < \xi < 1$ is a measure of the returns to scale and is positive but less than unity to ensure that the model has a unique solution (see, e.g. He and Liu (2008)).

[3] Note that capital-skill complementarity maintains if $1/(1-\rho) < 1/(1-\varphi)$.

3.2 Calibration and steady-state

We calibrate the model under exogenous fiscal policy to target the key great-ratios using U.S. annual data for the period 1970-2011. Table 2.1 below reports the model's quantitative parameters. Starting with the share of leisure in utility, $(1-\gamma)$, we calibrate it to 0.65 so that, in the steady-state, the household devotes about one third of its time to labour and education. The relative risk aversion parameter, $\sigma = 2$, is set to the value commonly employed in the literature.

The elasticities of substitution between skilled labour and capital and between unskilled labour and capital (or skilled labour) have been estimated by Krusell *et al.* (2000). Following the literature (see e.g. Lindquist (2004), and Pourpourides (2011)), we also use these estimates, to set $\varphi = 0.401$ and $\rho = -0.495$. Moreover, the income share of capital structures, a, is set to 0.12, as in Lindquist (2004). 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. More specifically, the unskilled labour weight in composite input share, $(1-\lambda) = 0.3$, is calibrated to obtain a skilled to unskilled labour ratio of about 79% and the capital equipment weight share in composite input, $\nu = 0.47$, is set to obtain a skill premium of approximately 1.64.[4] We also normalize the steady-state values of TFP and capital equipment efficiency to unity (i.e. $A = A^e = 1$).

The depreciation rates of capital structures and capital equipment, $\delta^p = 0.08$ and $\delta^e = 0.1$, are calibrated to obtain an annual capital to output ratio of about 1.94, which is consistent with the annual data reported by the BEA on capital stocks.[5] These values are also in line with the works of Greenwood *et al.* (1997) and Krusell *et al.* (2000). The time discount factor, $\beta = 0.96$, is set to obtain a post-tax post-depreciation annual real rate of return on capital of roughly 4.17%, which coheres with the value obtained in the data from the World Bank.[6]

The returns to scale parameter, ξ, in the skill acquisition equation is calibrated to be equal to 0.425, to obtain an investment in education to output ratio of about 1.8% which is similar to the average private expenditure on education in the U.S.[7] The weight on education time, $1-\phi$, is set equal

[4] The target value for the skill premium is obtained from U.S. Census data and the skilled to unskilled labour data is from the Acemoglu and Autor (2011) dataset for the past 20 years.

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

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

[7] Using annual data from U.S. National Center for Education Statistics, Digest of Ed-

to 0.45 to target average time in education as a share of total non-leisure time of about 5%.[8] The efficiency of skills transformation, B, is normalised to unity.[9]

Table 1: Calibration

	Definitions	Values
δ^p	depreciation rate of capital structures	0.080
δ^e	depreciation rate of capital equipment	0.100
β	time discount factor	0.960
γ	weight attached to consumption in utility	0.350
σ	coefficient of relative risk aversion	2.000
ϕ	weight on goods investment for skill acquisition	0.550
S	efficiency of skills production	1.000
ξ	returns to scale in skill creation	0.424
α	income share of capital structures	0.120
$\frac{1}{1-\rho}$	capital equipment to skilled labour elasticity	0.670
$\frac{1}{1-\varphi}$	capital equipment to unskilled labour elasticity	1.670
λ	share of composite input to output	0.700
ν	share of capital equipment to composite input	0.470
τ^s	skilled labour income tax rate	0.250
τ^u	unskilled labour income tax rate	0.200
τ^p	tax rate on capital structures income	0.310
τ^e	tax rate on capital equipment income	0.310
s^g	subsidy for goods investment in skill acquisition	0.000
A	total factor productivity	1.000
A^e	capital equipment productivity	1.000

Finally, we use the data from the ECFIN to construct series on effective capital and labour tax rates, following Martinez-Mongay (2000), to obtain an average tax rate for capital and labour.[10] Therefore, we set the tax rate for both capital income $\tau^p = \tau^e = 0.31$ and the two labour income tax rates

ucation Statistics, for the period 1970-2011, the relevant share to output is about 2%.

[8] To obtain this value we assume that the total time spent in higher education is on average 4 years. Note that the average years of working is 35. Therefore, the percentage of time spent in education is $\frac{4}{35} = 0.1143$. Taking into account that on average, 40-45% of the overall population in the U.S. are college educated (see Table 4 of the Census Bureau, Survey of Income and Program Participation), we obtain: $\frac{4}{35} \times 0.45 = 0.0514$.

[9] These parameters are within the range suggested in the related literature (i.e. Heckman (1976) and Stokey (1996)).

[10] In particular, we use data for 1970-2011, to construct the LITR and KITN rates for effective average labour and capital taxes respectively (see Martinez-Mongay (2000)), as they treat self-employed income as capital income.

$\tau^u = 0.20$ and $\tau^s = 0.25$.[11] Given that it is difficult to obtain data for the education investment subsidy, s^g, we set it to zero under the exogenous fiscal policy. We finally set the value of government expenditures, $G^c = 0.0320$, to obtain a steady-state debt to output ratio, $b/Y = 53\%$, which is equal to the average debt to GDP ratio obtained in the data.[12]

Under exogenous fiscal policy we solve the decentralized competitive equilibrium system of equations (58)-(69) in Appendix keeping the tax rates at their calibrated values in Table 1. Table 2 presents the steady-state results of the exogenous fiscal policy model together with the U.S. data averages for 1970-2011.

Table 2: Steady-state

	model	data
C/Y	0.563	0.660
K/Y	1.945	1.895
I/Y	0.181	0.146
E^g/Y	0.018	0.021
b/Y	0.530	0.530
h^s/h^u	0.798	0.763
G^c/Y	0.238	0.203
w^s/w^u	1.640	1.640
$\overline{w}^s/\overline{w}^u$	1.538	1.538
$\overline{r}$	0.042	0.042
$\frac{e}{h^s+h^u+e}$	0.053	0.051

The steady-state presented in Table 2 confirms that the model is close to the data as described above.[13]

4 Optimal fiscal policy

In this section we derive the optimal Ramsey plan, where it is assumed that the government chooses the series of taxes, subsidies and debt to finance exogenously determined public spending, with the objective to maximise the

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

[12] The source of this time series is FRED Economic Data on Gross Federal Debt as a percentage of GDP, 1970-2011.

[13] Note that the barred values in Table 2 are defined as follows: $\overline{w}^s = (1-\tau^s)\,w^s$, $\overline{w}^u = (1-\tau^u)\,w^u$ and $\overline{r} = \left(1-\tau^i\right)r^i + \left(1-\delta^i\right) = R^b$, where $i = s, u$.

welfare of the household.[14] The government, in other words, wishes to minimise the welfare costs of taxation. To obtain the *second-best* allocations, it is assumed that the government has access to a commitment technology. To solve the Ramsey problem we use the primal approach and first derive the present discounted value (PDV) of the household's lifetime budget constraint making use of the no-arbitrage and transversality conditions for the three assets as well as the Arrow-Debreu price of the bond. Second, we derive the implementability constraint by substituting out prices and tax rates from the household's PDV budget constraint using the household's and firm's first-order conditions. Finally, we derive the optimal Ramsey plan by maximising the planner's objective function subject to the implementability, skill acquisition and aggregate resource constraint.

4.1 Implementability constraint

Summing the household's budget constraint (5) successively forward from $t = 0$ to $t = \infty$ and imposing the no-arbitrage (16)-(17) and transversality conditions (18)-(19) gives the household's PDV or lifetime budget constraint:

$$\begin{aligned}&\sum_{t=0}^{\infty}\left[\prod_{i=0}^{t-1}\left(R_i^b\right)^{-1}\right]\left[C_t+\left(1-s_t^g\right)E_t^g\right]=\sum_{t=0}^{\infty}\left[\prod_{i=0}^{t-1}\left(R_i^b\right)^{-1}\right]\times\\&\times\{\left(1-\tau_t^s\right)w_t^s h_t^s+(1-\tau_t^u)w_t^u h_t^u\}+b_0+\{(1-\tau_0^p)r_0^p+\\&+(1-\delta^p)\}K_0^p+\{(1-\tau_0^e)r_0^e+(1-\delta^e)\}K_0^e.\end{aligned}\tag{35}$$

Following Ljungqvist and Sargent (2012), the Arrow-Debreu price is defined as: $q_t^0 = \prod_{i=0}^{t-1}\left(R_i^b\right)^{-1}$, $\forall t \geq 1$, with $q_0^0 = 1$, which implies that (35) can be rewritten as:

$$\begin{aligned}&\sum_{t=0}^{\infty}q_t^0\left[C_t+\left(1-s_t^g\right)E_t^g\right]=\sum_{t=0}^{\infty}q_t^0\{\left(1-\tau_t^s\right)w_t^s h_t^s+(1-\tau_t^u)w_t^u h_t^u\}+\\&+b_0+\{(1-\tau_0^p)r_0^p+(1-\delta^p)\}K_0^p+\{(1-\tau_0^e)r_0^e+(1-\delta^e)\}K_0^e.\end{aligned}\tag{36}$$

Notice that the Arrow-Debreu price satisfies the recursion:

$$q_{t+1}^0=\left(R_t^b\right)^{-1}q_t^0.\tag{37}$$

[14] Note that following the optimal fiscal policy literature, we keep the level of G^c fixed over time to the value obtained under exogenous fiscal policy. Note also that the subsidy to skill creation expenditure is added to ensure that all margins relating to the household's decision making are taxed/subsidised, so that the optimal policy problem is indeed Ramsey.

Using the first-order conditions (7) and (14), the above recursion can be written as:

$$\begin{aligned} q_{t+1}^0 &= \beta^{t+1}\frac{U_{C_{t+1}}}{U_{C_0}} \\ \Rightarrow q_t^0 &= \beta^t \frac{U_{C_t}}{U_{C_0}}. \end{aligned} \tag{38}$$

Substituting: (i) (38) into (36) for q_t^0; (ii) the first-order conditions of the firm, (24)-(25) into (36) for r_0^p and r_0^e respectively; and (iii) the first-order conditions of the household, (7)–(11) into (36) for $(1-\tau_t^s)\,w_t^s$, $(1-\tau_t^u)w_t^u$ and $(1-s_t^g)$ gives the household's implementability constraint:

$$\sum_{t=0}^{\infty}\beta^t\left[U_{C_t}C_t - \left(\frac{U_{e_t}}{g_{e_t}}g_{E_t^g}\right)E_t^g + \left(U_{h_t^s} + \frac{U_{e_t}}{g_{e_t}}\right)h_t^s + U_{h_t^u}h_t^u\right] = A_0 \tag{39}$$

where, $A_0 = U_{C_0}\{b_0 + \left[(1-\tau_0^p)f_{K_0^p} + 1 - \delta^p\right]K_0^p + \left[(1-\tau_0^e)f_{K_0^e} + 1 - \delta^e\right]K_0^e$. Note that $f_{K_0^p}$ and $f_{K_0^e}$ are obtained by substituting the market clearing conditions (28)-(29) into $f_{\widetilde{K}_t^p}$ and $f_{\widetilde{K}_t^e}$ respectively.

4.2 The primal approach

Under the primal approach the government maximises the following objective function:

$$\max \sum_{t=0}^{\infty}\beta^t U\left(C_t, l_t\right) \tag{40}$$

subject to the skill acquisition (3), the aggregate resource (27) and the implementability (39) constraints by choosing: $\{C_t, h_t^s, h_t^u, e_t,\ E_t^g, K_{t+1}^p, K_{t+1}^e\}_{t=0}^{\infty}$ given $\{\tau_0^p, \tau_0^e, b_0, K_0^p, K_0^e\}$.[15] Following Ljungqvist and Sargent (2012) we define a pseudo-value function and assume that Φ is the Lagrange multiplier with respect to the implementability constraint:

$$\begin{aligned} &V\left(C_t, h_{s,t}, h_{u,t}, e_t; \Phi\right) = U\left(C_t, l_t\right) + \Phi\{U_{C_t}C_t - \\ &-\left(\tfrac{U_{e_t}}{g_{e_t}}g_{E_t^g}\right)E_t^g + \left(U_{h_t^s} + \tfrac{U_{e_t}}{g_{e_t}}\right)h_t^s + U_{h_t^u}h_t^u\}. \end{aligned} \tag{41}$$

We can then write the Lagrangian under the primal approach as:

$$\begin{aligned} J = \sum_{t=0}^{\infty}\beta^t V\left(C_t, h_{s,t}, h_{u,t}, e_t; \Phi\right) + \theta_t[Y_t - G_t^c - C_t - E_t^g - K_{t+1}^p + \\ + (1-\delta^p)K_t^p - K_{t+1}^e + (1-\delta^e)K_t^e] + \zeta_t\left[h_t^s - g(E_t^g, e_t)\right] - \Phi A_0 \end{aligned} \tag{42}$$

[15] Following the literature, we do not examine the problem of initial capital taxation and thus not allow the government to optimally choose the capital income taxes at $t = 0$.

where Y_t is given by equation (32) above and $\theta_t, \zeta_t \geq 0\ \forall t$, are sequences of Lagrange multipliers with respect to the aggregate resource constraint and the skill acquisition constraint respectively. Given the initial values of capital taxes, debt and the two stocks of capital, equation J is maximised with respect to $\{C_t, h_t^s, h_t^u, e_t, E_t^g, K_{t+1}^p, K_{t+1}^e\}_{t=1}^{\infty}$ and for $t = 0$ equation J is maximised with respect to $\{C_0, h_0^s, h_0^u, e_0, I_0^g\}$ yielding the following first-order conditions respectively:

$$V_{C_t} = \theta_t,\ t \geq 1 \tag{43}$$

$$V_{h_t^s} = -\theta_t Y_{h_t^s} - \zeta_t,\ t \geq 1 \tag{44}$$

$$V_{h_t^u} = -\theta_t Y_{h_t^u},\ t \geq 1 \tag{45}$$

$$V_{e_t} = \zeta_t g_{e_t},\ t \geq 1 \tag{46}$$

$$V_{E_t^g} = \theta_t + \zeta_t g_{E_t^g},\ t \geq 0 \tag{47}$$

$$\theta_t = \beta\theta_{t+1}\left[Y_{K_{t+1}^p} + 1 - \delta^p\right],\ t \geq 0 \tag{48}$$

$$\theta_t = \beta\theta_{t+1}\left[Y_{K_{t+1}^e} + 1 - \delta^e\right],\ t \geq 0 \tag{49}$$

$$V_{C_0} = \theta_0 + \Phi A_C \tag{50}$$

$$V_{h_0^s} = -\theta_0 Y_{h_0^s} - \zeta_0 + \Phi A_{h_s} \tag{51}$$

$$V_{h_0^u} = -\theta_0 Y_{h_0^u} + \Phi A_{h_u} \tag{52}$$

$$V_{e_0} = \zeta_0 g_{e_0} + \Phi A_e \tag{53}$$

where $Y_{x,t}$ is the derivative of Y_t, given by equation (32), with respect to the relevant variable x at time t.

The above system of first-order conditions implies that the system to be solved will be different for $t = 0$ and $t = 1, 2, 3...T-1$ and $t = T$. This is reflected in equations (70)-(93) reported in the Appendix. To solve this system, we initially guess a value for Φ and solve equations (70)-(93) for an allocation $\{C_t, h_t^s, h_t^u, e_t, E_t^g, \zeta_t, K_{t+1}^p, K_{t+1}^e\}_{t=0}^{\infty}$.[16] The system has $[(8 \times T) + 1]$ equations and is solved using standard non-linear numerical methods (see, e.g. Garcia-Milà *et al.* (2010) and Adjemian *et al.* (2011)). Then we test if the implementability constraint (39) is binding and we increase or decrease accordingly the value of Φ until the implementability constraint is satisfied. We set the initial conditions for debt, the two stocks of capital and the two capital income taxes equal to their exogenous steady-state, to calculate the dynamic transition path from the exogenous to optimal fiscal policy steady-state. To ensure that the variables converge to the optimal fiscal policy

[16]Note that the multiplier θ_t has been substituted out of the system presented in the Appendix.

steady-state, we set the value of $T = 250$. The results indicate that model convergence is achieved after 150 periods.

4.3 Optimal allocations and policy

We first analyse the steady-state under optimal policy and compare outcomes with the current economy. We then evaluate the transition paths that the policymaker would choose if, starting from the current economy, economic policy was chosen optimally by working as described in the previous subsection.

4.3.1 Ramsey policy in the steady-state

In Table 3, we present the Ramsey optimal resource allocations and policy choices in the steady-state. The Table also includes the steady-state outcomes of the economy under exogenous policy that is calibrated to the data averages as explained in the previous section. The first result which can be confirmed in Table 3 is that, consistent with the literature on optimal taxation, capital taxes are zero in the long-run, for both capital stocks. In contrast, labour income taxes, are positive and, in fact, significantly more progressive compared with the calibrated economy under exogenous policy. In particular, in the steady state, skilled labour income is taxed about 40% more than unskilled labour income. The tax revenue generated by these taxes, in addition to the revenue from the assets that the government holds optimally in the steady-state, finance the exogenous stream of public spending as well as optimal subsidies to skill acquisition expenditure.

Table 3: Exogenous and Ramsey steady-states

	Exogenous policy	Ramsey policy		Exogenous policy	Ramsey policy
Y	0.134	0.169	$\overline{w}^s/\overline{w}^u$	1.538	1.456
C	0.076	0.090	τ^p	0.310	0.000
K^p	0.091	0.167	τ^e	0.310	0.000
K^e	0.170	0.289	τ^s	0.250	0.318
h^s	0.118	0.135	τ^u	0.200	0.233
h^u	0.148	0.133	s^g	0.000	0.451
e	0.015	0.014	b/Y	0.530	-0.475
E^g	0.002	0.005	U	-76.438	-75.010
w^s/w^u	1.640	1.638	ψ	-	5.243%

The optimal allocations in turn reflect the changes in the policy instruments compared with the calibrated economy under exogenous policy. In particular, capital accumulation increases, following the elimination of the capital taxes. The rise in skill-biased capital stock tends to increase the skill premium, which acts to raise the relative skill supply. The latter is further supported by the subsidy to skill acquisition expenditure. As a result, the relative skill supply increases, so that the skill premium under Ramsey policy is effectively the same as the skill premium under exogenous policy. However, the increase in the progressivity of labour income taxation implies that wage inequality, as captured by the skill premium net of taxes, is reduced. Overall, optimal policy reduces the distortions associated with the tax system. This is evident by the increase in output and consumption under Ramsey policy and by the welfare gains, in terms of consumption, obtained by moving from exogenous to optimal policy. In particular, the welfare gains measured by the compensating consumption supplement, ψ, are roughly 5.2%.[17]

4.3.2 Optimal progressive labour income taxes

The most striking result regarding optimal policy in Table 3 is that the labour income taxes should optimally be progressive. What makes this result notable is that it is obtained in an economy without market failures and without redistribution incentives for the policymaker.[18] To understand this finding, we start by considering the main principle of Ramsey taxation, which suggests that, to minimise efficiency distortions, taxes should be higher for more inelastic tax bases.[19] Our finding, that skilled labour income should be taxed more than unskilled, is consistent with this principle, since we find that skilled work time is more inelastic than unskilled work time. To demonstrate this, in Figure 1 we plot the percent deviations for h^s and h^u from their steady-state under exogenous policy in Table 3, after a permanent 1% change in either τ^s (solid lines) or τ^u (dashed lines). Subplots (1,1) and (1,2) respectively show the elasticities of skilled and unskilled work time and the elasticities of skilled and unskilled labour income with respect to the tax

[17] The welfare gains are obtained as the compensating consumption supplement that would make the economy under exogenous policy as well off as the economy under optimal policy.

[18] Note that optimal labour income progressivity is not driven by the subsidy to expenditure on skill acquisition, although the latter does affect its magnitude. In particular, if we restrict the government from having access to this policy instrument, τ^s is still higher than τ^u, but the difference is smaller.

[19] The importance of the elasticities for labour income tax progressivity has been highlighted in the recent optimal taxation literature (see e.g. Diamond (1998) and Saez (2001)). These studies also demonstrate the importance of the shape of the income (or skills) distribution and of the social weights in the objective function of the planner in setups with heterogeneous households.

rates. As can be seen, skilled work time and income are more inelastic with respect to the relevant income tax rate, compared with unskilled work time and labour income. Accordingly, the policymaker finds it optimal to tax skilled labour income more than unskilled, so that labour income taxation is progressive.

[Figure 1 here]

Note that in this economy, the opportunity cost of skilled work time is foregone: (i) utility-augmenting leisure; (ii) income from unskilled work time; and (iii) income due to higher education expenditure to compensate for the loss in education time required for skill creation (see 15)). Therefore, the elasticity of skilled work time with respect to the tax rate is affected by channels that operate via unskilled work time and skill acquisition, in addition to usual substitution and income effects. To illustrate the importance of these channels and explain how they affect the skilled work time elasticity, relative to unskilled work time elasticity, we next further investigate the factors that determine the elasticities in our setup.

The elasticity of skill supply with respect to the tax rate on skilled labour income is defined as $\epsilon^s \equiv \frac{dh_t^s}{d\tau_t^s}\frac{\tau_t^s}{h_t^s}$. Using the household's optimality conditions and the functional forms for the utility function assumed above, we have:

$$h_t^s = 1 - h_t^u - e_t - \frac{\left[(1-\gamma) + \frac{(1-\gamma)}{g_{e_t}}\right] C_t}{\gamma(1-\tau_t^s)w_t^s} \tag{54}$$

which implies that the total derivative $\frac{dh_t^s}{d\tau_t^s}$ is given by:

$$\begin{aligned} \frac{dh_t^s}{d\tau_t^s} &= \frac{\partial h_t^s}{\partial \tau_t^s} + \frac{\partial h_t^s}{\partial C_t}\frac{dC_t}{d\tau_t^s} + \frac{\partial h_t^s}{\partial w_t^s}\frac{dw_t^s}{d\tau_t^s} + \frac{\partial h_t^s}{\partial h_t^u}\frac{dh_t^u}{d\tau_t^s} + \frac{\partial h_t^s}{\partial E_t^g}\frac{dE_t^g}{d\tau_t^s} + \frac{\partial h_t^s}{\partial e_t}\frac{de_t}{d\tau_t^s} \\ &= -\frac{(1-\gamma)\left[1+\frac{1}{g_{e_t}}\right]}{\gamma w_t^s(1-\tau_t^s)}\left[\frac{C_t}{(1-\tau_t^s)} + \frac{dC_t}{d\tau_t^s} - \frac{C_t}{w_t^s}\frac{dw_t^s}{d\tau_t^s}\right] - \\ &\quad -\frac{dh_t^u}{d\tau_t^s} + \frac{\frac{(1-\gamma)}{(g_{e_t})^2}C_t g_{e_t E_t^g}}{\gamma(1-\tau_t^s)w_t^s}\frac{dE_t^g}{d\tau_t^s} + \left[\frac{\frac{(1-\gamma)}{(g_{e_t})^2}C_t g_{e_t e_t}}{\gamma(1-\tau_t^s)w_t^s} - 1\right]\frac{de_t}{d\tau_t^s} \end{aligned} \tag{55}$$

where $g_{e_t} > 0$, $g_{e_t e_t} < 0$ and $g_{e_t E_t^g} > 0$ denote the respective first-partial, second-partial and cross-partial derivatives of the skill acquisition function.

The elasticity of unskilled work time with respect to the tax rate on unskilled labour income is defined as $\epsilon^u \equiv \frac{dh_t^u}{d\tau_t^u}\frac{\tau_t^u}{h_t^u}$. Using the household's optimality conditions and the functional forms for the utility function assumed above, we have:

$$h_t^u = 1 - h_t^s - \frac{(1-\gamma)\,C_t}{\gamma(1-\tau_t^u)w_t^u} \tag{56}$$

which implies that the total derivative $\frac{dh_t^u}{d\tau_t^u}$ is given by:

$$\begin{aligned}\frac{dh_t^u}{d\tau_t^u} &= \frac{\partial h_t^u}{\partial \tau_t^u} + \frac{\partial h_t^u}{\partial C_t}\frac{dC_t}{d\tau_t^u} + \frac{\partial h_t^u}{\partial w_t^u}\frac{dw_t^u}{d\tau_t^u} + \frac{\partial h_t^u}{\partial h_t^s}\frac{dh_t^s}{d\tau_t^u} \\ &= -\frac{(1-\gamma)C_t}{\gamma w_t^u(1-\tau_t^u)^2} - \frac{(1-\gamma)}{\gamma(1-\tau_t^u)w_t^u}\frac{dC_t}{d\tau_t^u} + \frac{(1-\gamma)C_t}{\gamma(1-\tau_t^u)(w_t^u)^2}\frac{dw_t^u}{d\tau_t^u} - \frac{dh_t^s}{d\tau_t^u}.\end{aligned} \tag{57}$$

Equation (55) shows that there are six terms determining the responsiveness of skilled work time to the tax rate on skilled labour income. These terms incorporate the following marginal effects of τ_t^s: $\frac{dh_t^s}{d\tau_t^s}$, $\frac{dC_t}{d\tau_t^s}$, $\frac{dw_t^s}{d\tau_t^s}$, $\frac{dh_t^u}{d\tau_t^s}$, $\frac{dE_t^g}{d\tau_t^s}$ and $\frac{de_t}{d\tau_t^s}$. To facilitate the analysis we plot these derivatives in Figure 2 for a permanent 1% increase in τ_t^s.[20]

[Figure 2 here]

With the aid of Figure 2 we can sign the effect of each term in (55) on $\frac{dh_t^s}{d\tau_t^s}$. These include the:

(i) substitution effect: $-\frac{(1-\gamma)\left[1+\frac{1}{g_{e_t}}\right]C_t}{\gamma w_t^s(1-\tau_t^s)^2} < 0$, capturing the reduction in the net return to an hour of work, due to a higher τ_t^s, which tends to decrease h_t^s;

(ii) consumption (income) effect: $-\frac{(1-\gamma)\left[1+\frac{1}{g_{e_t}}\right]}{\gamma(1-\tau_t^s)w_t^s}\frac{dC_t}{d\tau_t^s} > 0$, since $\frac{dC_t}{d\tau_t^s} < 0$ (see subplot (1,2)), capturing the reduction in consumption, due to a higher τ_t^s, which tends to increase h_t^s to compensate for the loss in income;

(iii) wage effect: $\frac{(1-\gamma)\left[1+\frac{1}{g_{e_t}}\right]C_t}{\gamma(1-\tau_t^s)(w_t^s)^2}\frac{dw_t^s}{d\tau_t^s} > 0$, since $\frac{dw_t^s}{d\tau_t^s} > 0$ (see subplot (2,1)), capturing the higher gross equilibrium wage rate, due to the higher τ_t^s, which tends to increase h_t^s;

(iv) cross-work time effect: $-\frac{dh_t^u}{d\tau_t^s} < 0$, since $\frac{dh_t^u}{d\tau_t^s} > 0$ (see subplot (2,2)), capturing the increase in the supply of unskilled labour, due the fall in the return of skilled relative to unskilled labour, which tends to decrease h_t^s.

(v) resource allocation effect from E_t^g: $\frac{\frac{(1-\gamma)}{(g_{e_t})^2}C_t g_{e_t E_t^g}}{\gamma(1-\tau_t^s)w_t^s}\frac{dE_t^g}{d\tau_t^s} < 0$, since $\frac{dE_t^g}{d\tau_t^s} < 0$ (see subplot (3,1)), capturing the reduction in expenditure for skill acquisition, due to the fall in the net return to skilled labour, which tends to decrease h_t^s because of the increase in disposable income;

[20] Note that the derivatives for a permanent 1% increase in τ_t^u are also plotted in Figure 2 to help sign deriviative terms in (57).

(vi) resource allocation effect from e_t: $\left[\frac{\frac{(1-\gamma)}{(g_{e_t})^2} C_t g_{e_t e_t}}{\gamma(1-\tau_t^s) w_t^s} - 1\right] \frac{de_t}{d\tau_t^s} > 0$ since $\frac{de_t}{d\tau_t^s} < 0$ (see subplot (3,2)), capturing the reduction in education time for skill acquisition, due to the fall in the net return to skilled labour, which tends to increase h_t^s since there is more time available to work.

To understand why the response of skilled labour to a tax change is smaller than that of unskilled labour, i.e. $|\epsilon^s| < |\epsilon^u|$, we next consider the elasticity of unskilled work time with respect to the tax rate on unskilled labour income in (57). In particular, note that the two resource allocation effects from skill creation, i.e. (v) and (vi) are absent from the total derivative expression for unskilled work time. Referring to Figure 1, it can be seen that the total effect on unskilled work time in the presence of effects (i)-(iv) only is strongly negative. Figure 2 also confirms that each of the derivatives in equation (57) has the same signs as the respective derivatives in the skilled labour elasticity equation (55). Hence, it appears that the additional effects associated with skill creation in equation (55) play an important role in decreasing $|\epsilon^s|$ relative to $|\epsilon^u|$, since the positive resource allocation effect from e_t dominates that from E_t^g.

4.3.3 Optimal transition paths

We can now examine the optimal transition of the economy from the steady-state of exogenous policy, as summarised in Table 3, to the optimal steady-state under Ramsey policy in the same Table. We plot the optimal transition paths for the policy instruments and the economic allocations in Figure 3.

[Figure 3 here]

Consistent with the analysis of optimal capital income taxes (see e.g. Chamley (1986) and Chari *et al.* (1994)), the capital taxes, τ_t^p and τ_t^e, are set very high in the first period, switch to zero in the second period and remain at this level. This tax policy allows the government to generate high tax revenue in the first period, accumulate assets by lending to the household and thus generate a stream of revenue from the returns to these assets in the future (see path of b_t/Y_t). The pattern of the capital tax rates is reflected on the dynamics of the capital stocks, K_t^p and K_t^e, which initially decrease and then increase. In turn, these dynamics influence the movements of the wage rates, since the marginal products of labour are positive functions of the capital stocks. However, given the skill-biased role of equipment capital, the skill premium, w_t^s/w_t^u, increases initially, before returning effectively, in the long-run, to its original steady-state.

The initial rise in capital taxes (and fall in capital stocks) is met by a reduction in both labour income taxes, τ_t^s and τ_t^u, (which become subsidies in the first period) and a subsidy to skill expenditure, s_t^g, which help to increase work time, h_t^s and h_t^u, despite the fall in marginal labour productivity. This policy mediates the negative effects on the economy in the very short-run brought about by the capital tax and capital stock movements. Both labour income taxes return to positive magnitudes after the first period, but it is interesting to note that the labour tax system is optimally progressive from the first period onwards. This progressivity in turn works to reduce post-tax wage inequality, $\overline{w}_t^s/\overline{w}_t^u$.

Note that the response of skilled labour hours to optimal policy is much smoother than the response of unskilled labour hours, despite the higher volatility in the skilled wage rate compared with that of the unskilled. This is consistent with a more inelastic skilled work time, as discussed above. Moreover, it is consistent with empirical evidence which suggests that in the USA, over 1979-2002, the standard deviation of unskilled labour hours was on average 1.3 times higher than the standard deviation of skilled labour hours, despite the standard deviation of the skilled wage being 1.2 times higher than the standard deviation of the unskilled wage.[21]

Optimal policy leads to a more efficient economy with higher output, Y_t, and consumption, C_t. Hence, the implicit relative cost of goods versus time in creating skill is lower for the Ramsey planner, relative to the exogenous policy case. This, in turn, results in a subsidy for skill acquisition expenditure, making it cheaper for the household to use goods relative to time when creating skill, and this is reflected in the movements of education time, e_t, and skill expenditure, E_t^g, which decrease and increase respectively.

5 Conclusions

This paper evaluated quantitatively the extent of the progressivity of optimal labour income taxation, under skill heterogeneity, endogenous skill acquisition and a production sector characterised by capital-skill complementarity. We isolated optimal taxation from incentives for income redistribution by working with a representative agent framework and considered the problem of a Ramsey planner, who had access to a full instrument set to minimise the distortions associated with taxation in an economy with perfect capital and labour markets.

[21] These statistics are obtained using the quarterly data on skilled and unskilled hours and wages in Lindquist (2004). We thank Matthew Lindquist for making these series available to us.

In this framework, the household decided how to allocate its expenditure into investment in the two types of capital stock and into goods for creating skilled labour. Moreover, it decided how to allocate its time endowment into leisure, work time in skill and unskilled jobs, as well as in education for creating skilled labour. The resource requirements associated with skill acquisition in turn implied that there was a wage premium accruing to skilled labour to compensate for these costs.

We found that wage inequality in this setup implied progressive labour income taxation, because the work time elasticity for skilled labour was lower relative to that of unskilled. The resource implications for creating skilled labour established effects on the skilled work time elasticity, driven changes in the household's disposable income and in its available total time when the tax on skilled labour income changed. These additional effects worked to increase and decrease, respectively, the elasticity of skilled work time with respect to the tax rate. When the model was calibrated to U.S. data, the skilled work time elasticity was lower relative to that for unskilled labour, thus leading to optimal progressive labour income taxes.

We further found that the Ramsey plan required that capital taxes were set very high in the first period and then rapidly decreased towards zero, as is common in the literature on optimal capital taxation. Moreover, the government ran big surpluses in the first period, which allowed it to create a stock of assets, which were in turn used to finance primary deficits in the future. The changes in taxation along the optimal path implied a sharp increase in wage inequality in the first periods, before the skill premium returned effectively to its initial level. However, since the tax system became progressive from the first period, net wage inequality was reduced under Ramsey policy over the entire transition path.

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] Adjemian, S., Bastani, H., Karamé, F., Juillard, M., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M. and S. Villemot (2011). 'Dynare: Reference Manual Version 4', Dynare Working Papers 1, CEPREMAP, revised Jan 2014.

[3] Angelopoulos, K., Asimakopoulos, S. and J. Malley (2014). 'Tax smoothing in a business cycle model with capital-skill complementarity', CESifo Working Paper Series, No. 4744.

[4] Angelopoulos, K., Malley, J. and A. Philippopoulos (2013). 'Human Capital, Social Mobility and the Skill Premium', CESifo Working Paper Series 4388.

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

[6] Chari, V. and P. Kehoe (1999). 'Optimal fiscal and monetary policy', in Handbook of Macroeconomics, Vol. 1A, edited by J.B. Taylor and M. Woodford, North-Holland.

[7] 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.

[8] Diamond, P. (1998). 'Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates', American Economic Review, 88 (1), 83-95.

[9] Drazen, A. (2000). 'Political economy in macroeconomics', Princeton University Press.

[10] Garcia-Milà, T., Marcet, A. and E. Ventura (2010). 'Supply Side Interventions and Redistribution', Economic Journal, 120 (543), 105-130.

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

[12] Greenwood, J., Hercowitz, Z. and P. Krusell (1997). 'Long-run implications of investment-specific technological change', American Economic Review, 87 (3), 342–362.

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

[14] Heckman, J.J. (1976). 'A life-cycle model of earnings, learning and consumption', Journal of Political Economy, 84 (4), S11-S44.

[15] 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.

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

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

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

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

[20] Martinez-Mongay, C. (2000). 'ECFIN's effective tax rates. Properties and comparisons with other tax indicators', Economic Paper No. 146, (October), Brussels, European Commission.

[21] Mirlees, J.A. (1971). 'An exploration in the theory of optimum income taxation', The Review of Economic Studies, 38 (2), 175-208.

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

[23] Saez, E. (2001). 'Using elasticities to derive optimal income taxes', Review of Economic Studies, 68, 205-229.

[24] Stokey, N.L. (1996). 'Free trade, factor returns, and factor accumulation', Journal of Economic Growth, 1 (4), 421-447.

6 Appendix

6.1 DCE system of equations

The DCE, when the Lagrange multipliers Λ_t, and M_t have been substituted can be written as follows:

$$-\frac{U_{h_t^s}}{U_{C_t}} - (1-\tau_t^s)w_t^s - \frac{U_{e_t}}{U_{C_t} g_{e_t}} = 0 \tag{58}$$

$$-\frac{U_{h_t^u}}{U_{C_t}} - (1-\tau_t^u)w_t^u = 0 \tag{59}$$

$$-\frac{U_{e_t}}{U_{C,t}} - (1-s_t^g)\,\frac{g_{e_t}}{g_{E_t^g}} = 0 \tag{60}$$

$$\frac{U_{C_t}}{\beta U_{C_{t+1}}} - r_{t+1}^p(1-\tau_{t+1}^p) - (1-\delta^p) = 0 \tag{61}$$

$$\frac{U_{C_t}}{\beta U_{C_{t+1}}} - r_{t+1}^e(1-\tau_{t+1}^e) - (1-\delta^e) = 0 \tag{62}$$

$$\frac{U_{C_t}}{\beta U_{C_{t+1}}} - R_t^b = 0 \tag{63}$$

$$w_t^s - f_{\widetilde{h}_t^s} = 0 \tag{64}$$

$$w_t^u - f_{\widetilde{h}_t^u} = 0 \tag{65}$$

$$r_t^p - f_{\widetilde{K}_t} = 0 \tag{66}$$

$$r_t^e - f_{\widetilde{K}_t^e} = 0 \tag{67}$$

$$G_t^c + s_t^g E_t^g + b_t - \tau_t^s w_t^s h_t^s - \tau_t^u w_t^u h_t^u - \tau_t^p r_t^p K_t^p - \tau_t^e r_t^e K_t^e - \frac{b_{t+1}}{R_t^b} = 0 \tag{68}$$

$$Y_t - G^c - C_t - I_t^p - I_t^e - E_t^g = 0. \tag{69}$$

6.2 First order conditions of optimal policy

The first order conditions for the government's problem are:

- for $t = 0$:

$$V_{h_0^s} = -(V_{C_0} - \Phi A_C)Y_{h_0^s} - \zeta_0 + \Phi A_{h_s} \tag{70}$$

$$V_{h_0^u} = -(V_{C_0} - \Phi A_C)Y_{h_0^u} + \Phi A_{h_u} \tag{71}$$

$$V_{e_0} = \zeta_0 g_{e_0} + \Phi A_e \tag{72}$$

$$V_{E_0^g} = \zeta_0 g_{E_0^g} + V_{C_0} - \Phi A_C \tag{73}$$

$$V_{C_0} - \Phi A_C = \beta V_{C_1} \left[Y_{K_1^p} + 1 - \delta^p \right] \tag{74}$$

$$V_{C_0} - \Phi A_C = \beta V_{C_1} \left[Y_{K_1^e} + 1 - \delta^e \right] \tag{75}$$

$$Y_0 = G^c + C_0 + I_0^p + I_0^e + E_0^g \tag{76}$$

$$h_0^s = g(E_0^g, e_0) \tag{77}$$

- for $t = 1, 2, 3...T - 1$:

$$V_{h_t^s} = -V_{Ct} Y_{h_t^s} - \zeta_t \tag{78}$$

$$V_{h_t^u} = -V_{C_t} Y_{h_t^u} \tag{79}$$

$$V_{e_t} = \zeta_t g_{e_t} \tag{80}$$

$$V_{E_t^g} = \zeta_t g_{E_t^g} + V_{C_t} \tag{81}$$

$$V_{Ct} = \beta V_{C_{t+1}} \left[Y_{K_{t+1}^p} + 1 - \delta^p \right] \tag{82}$$

$$V_{Ct} = \beta V_{C_{t+1}} \left[Y_{K_{t+1}^e} + 1 - \delta^e \right] \tag{83}$$

$$Y_t = G^c + C_t + I_t^p + I_t^e + E_t^g \tag{84}$$

$$h_t^s = g(E_t^g, e_t) \tag{85}$$

- for $t = T$:

$$V_{h_T^s} = -V_{C_T} Y_{h_T^s} - \zeta_T \tag{86}$$

$$V_{h_T^u} = -V_{C_T} Y_{h_T^u} \tag{87}$$

$$V_{e_T} = \zeta_T g_{e_T} \tag{88}$$

$$V_{E_T^g} = \zeta_T g_{E_T^g} + V_{C_T} \tag{89}$$

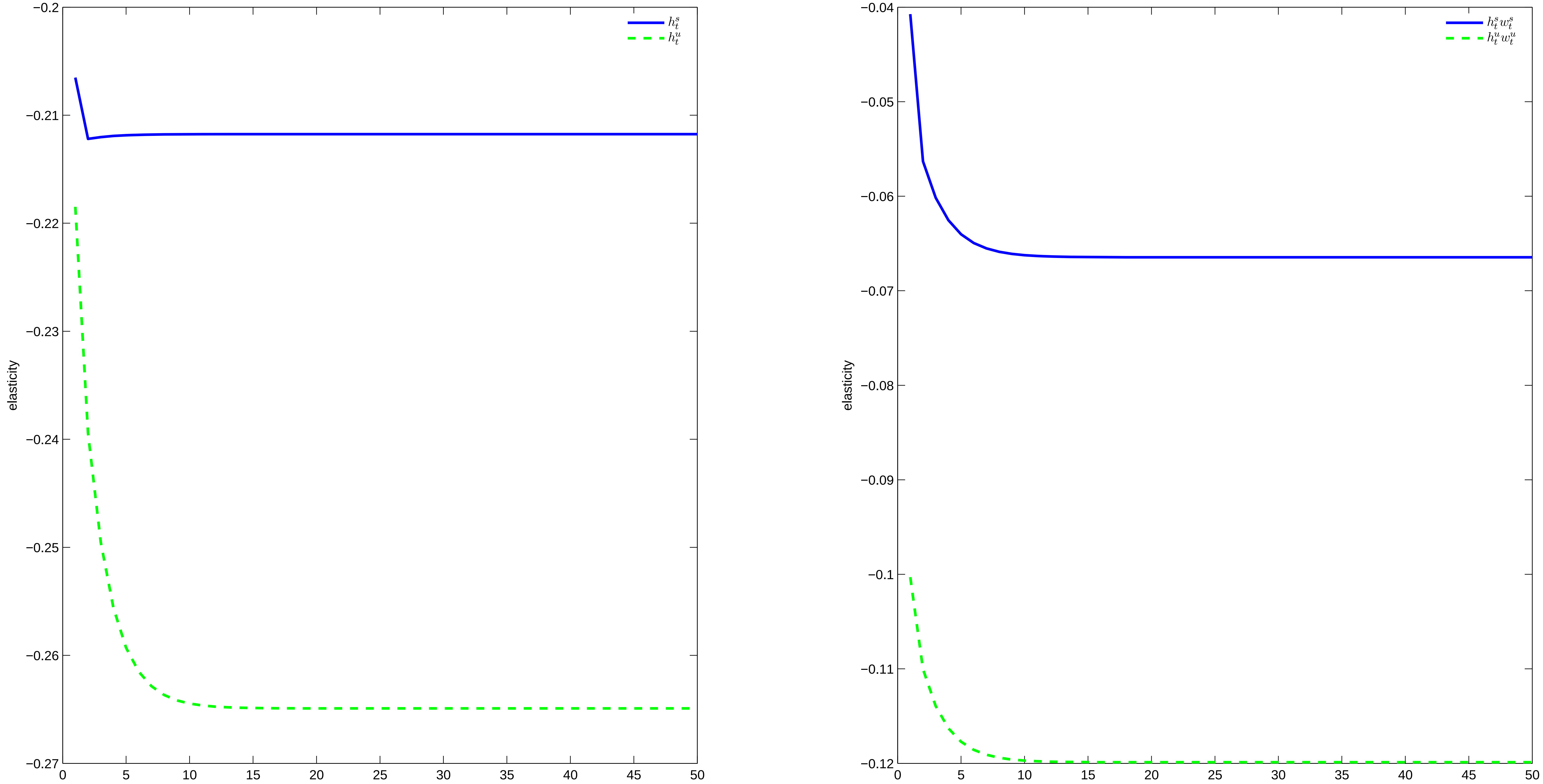
$$1 = \beta \left[Y_{K_T^p} + 1 - \delta^p \right] \tag{90}$$

$$1 = \beta \left[Y_{K_T^e} + 1 - \delta^e \right] \tag{91}$$

$$Y_T = G^c + C_T + I_T^p + I_T^e + E_T^g \tag{92}$$

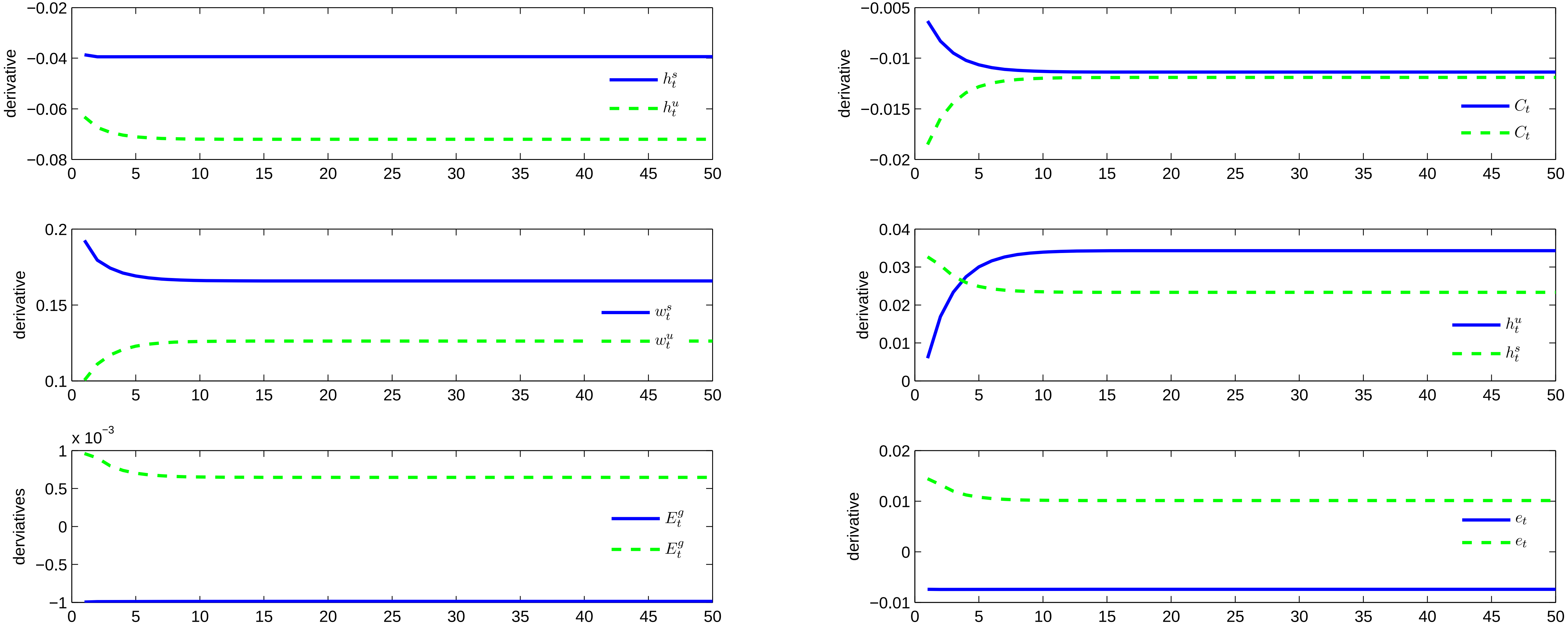
$$h_T^s = g(E_T^g, e_T) \tag{93}$$

Figure 1: Impulse responses for work time and income



Solid lines are for 1% perm. shock to τ_t^s and dashed lines are for 1% perm. shock to τ_t^u

Figure 2: Impulse responses derivatives



Solid lines are for 1% perm. shock to τ_t^s and dashed lines are for 1% perm. shock to τ_t^u

Figure 3: Transition paths from exogenous to optimal steady−state