Ramsey Taxation with Capital-Skill Complementarity and Investment-Specific Technological Change

Yi-Chan Tsai*

Department of Economics, National Taiwan University

C.C. Yang

Institute of Economics, Academia Sinica

Department of Public Finance, National Chengchi University

Department of Public Finance, Feng Chia University

Hsin Jung Yu

Institute of Economics, Academia Sinica

April 2018

Abstract

The relative price of capital has been falling over time in the presence of investment-specific technological change, which leads to an increase in the stock of capital. With capital-skill complementarity, a higher capital stock pushes up the demand for skilled relative to unskilled labor, and thus contributes to the rising wage inequality between skilled and unskilled workers. This paper qualitatively and quantitatively characterizes the dynamics of Ramsey taxation in response. It is shown that:

At a point in time, intertemporal distortions on capital are positive; and (ii)

Over time, intertemporal distortions on capital equipment display a decreasing pattern, while intratemporal distortions on labor display an increasing trend though at a diminishing rate. Interestingly, we find that the time path of the U.S. tax structure is close to that of Ramsey taxation.

*Corresponding author, Yi-Chan Tsai. Email: yichantsai@ntu.edu.tw

1 Introduction

The U.S.economy since the early 1980 has witnessed a dramatic rise in the log skill premium (defined as the mean of the natural logarithm of weekly wages for college graduates relative to high school graduates); see Figure 1. This significally upward trend on the skill premium has aroused serious concerns because it gives rise to large income inequality and disparity of economic well-being between skilled and unskilled workers.[1]

Krusell et al. (2000) forcefully argued that the force driving the rise of the skill premium lies in the simultaneous presence of "capital-skill complementarity" and "investment-specific technological change."

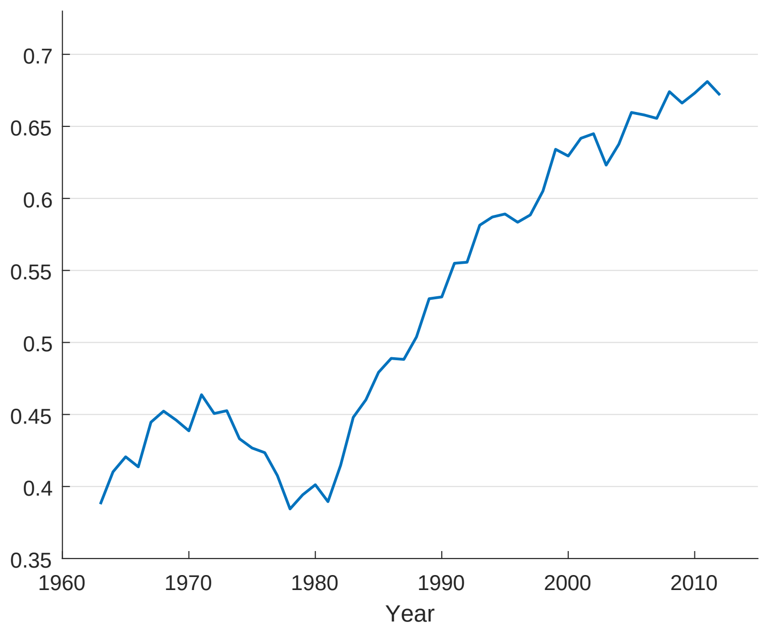

Figure 1: Log skill premium, 1963-2012

Note: Data are from Autor (2014).

Consider a production technology in which capital is more substitutionable for or less complementary to unskilled labor than skilled labor. This property is known as "capital-skill complementarity" in the literature. A critical implication of capital-skill complementarity is that a higher stock of capital will raise the marginal product of skilled labor but lower the marginal product of unskilled labor; see Krusell et al. (2000). There has been a steady, dramatic decline in the relative price of capital due to investment-specific technological change in the U.S.economy; see Figure 2.[2] With plausible differences in the elasticities of substitution

[1]Katz and Murphy (1992) is the seminal work on the issue. For a recent review, see Autor (2014).

[2]We provide details on the time series of Figure 2 in Section 4. The inverse of the pice of capital represents investment-specific technological change; see Eqs. (2).

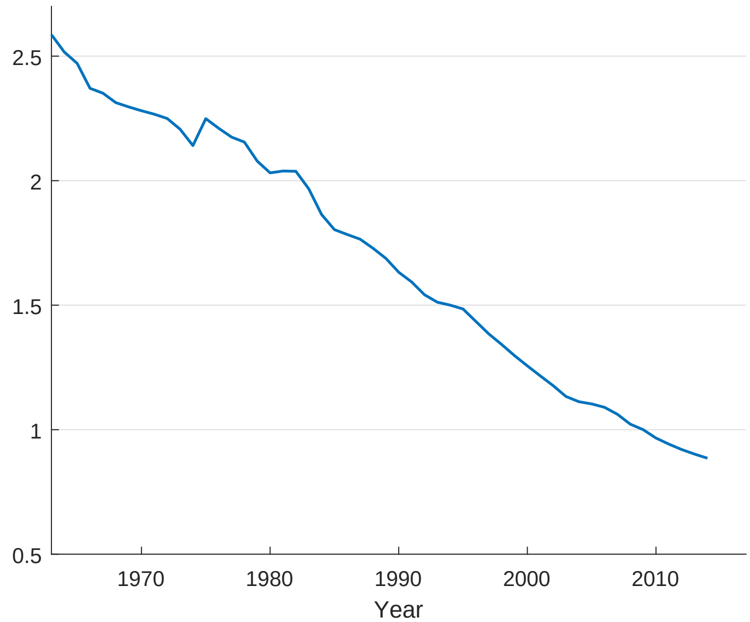

Figure 2: Price of capital relative to consumption, 1963-2014

Note: Capital includes both structures and equipment.
Source: Federal Reserve Bank Economic Data (FRED), series PIRIC, 1963-2014.

between capital and skilled vs. unskilled labor that support capital-skill complementarity, Krusell et al. (2000) found that the rising skill premium since the early 1980s can be well explained through the secular cheapening of capital associated with investment-specific technological change, even though there was a substantial increase in the relative supply of college skills during the sample period they studied.

In light of the important finding of Krusell et al. (2000), a sensible and interesting question about tax policy naturally arises: how should a government set taxes on capital versus labor over time in response to the simultaneous presence of capital-skill complementarity and investment-specific technological change? Formulating the question as a Ramsey problem and employing an empirically plausible capital-skill complementarity form of production function as suggested by Krusell et al. (2000), this paper quantitatively characterize the equilibrium dynamics of the optimal tax structure in the face of the secular decline in the relative price of capital equipment associated with investment-specific technological change.

Our optimal tax prescribes: (i) *at a point in time*, intertemporal distortions on capital are always positive; and (ii) *over time*, intertemporal distortions on capital equipment display a decreasing pattern, while intratemporal distortions on labor display an increasing trend though at a diminishing rate. To understand why the optimal tax for capital is positive, we consider a special case that allows the government to set the discriminatory taxes conditional on the types of workers. We find that the government will set higher tax rates for the skilled workers relative to the unskilled workers; The reason that the government wants to set a different tax rates for different types of workers is due to the differences in their labor supply elasticity (and the redistribution concern?) In particular, the labor supply elasticity for skilled labor is lower, and thus the typical inverse elasticity taxation rule applies, which requires the government to set higher tax rates for the skilled workers. Since the skilled workers have higher income, the optimal tax will prescribe the progressive labor tax. When setting discriminatory tax rate for different types of labor is plausible, the optimal capital tax rate is zero. However, when progressive taxes on skill are not allowed, capital equipment tax will be positive, which can imitate a progressive labor income tax with respect to skill. This results can be viewed as an application of Corlett Hague rule. In particular, when setting a higher tax rate for skilled worker is implausible, the government can impose tax on the complementary good of the skilled worker, i.e., the capital equipment, to mimic a progressive tax rates for skill. [3]

There is a large literature investigating how the government as a Ramsey planner should set taxes on capital versus labor in the face of aggregate shocks.[4]

Of them, the important contribution of Werning (2007) is closely related to our paper. Unlike the typical representative-agent framework in the literature, he explicitly modeled distributional concerns by envisioning a heterogeneous-agents economy in which agent types in terms of their labor productivity are permanently fixed. Distortionary taxes naturally arise from the trade-off between redistribution and efficiency in this heterogeneous-agents economy.[5] However, there is a crucial departure: while all types of labor are equally complementary with capital in production in Werning (2007), skilled labor is more complementary with capital equipment than unskilled labor in our setting. This departure is of critical importance in light of the Krusell et al.(2000) finding that capital-skill complementarity is key to the rise of skill premium. Indeed, our tax policy prescriptions deviate substantially from those prescribed by Werning (2007), which prescribes that capital should go untaxed and labor taxes should be perfectly smoothed.

Although there is a large literature on capital-skill complementarity, Jones et al. (1997), Ctirad Slavik and Hakki Yazici (2013) and Angelopoulos et al. (2015) are the only three papers ever addressing the issue of optimal taxation with capital-skill complementarity ac-

[3] The result is also similar to age-dependent tax proposed by Erosa and Gervais (2002). When it is impossible to setting a discriminatory for elder people, the government can instead taxing capital as the elder hold more capital.

[4] See Ljungqvist and Sargent (2012, chapter 16) for a review of the literature.

[5] It should be noted that although the focus of Werning (2007) was on Ramsey taxation, the paper addressed Mirrleesian taxation as well.

cording to our knowledge,.[6]

Jones et al. (1997) utilized capital-skill complementarity as an example to highlight a point: if the tax rate on skilled and unskilled labor is required to be equal, the startling finding by Chamley (1986) and Judd (1985) that capital should go untaxed in the steady state may no longer hold. Similar to the findings of Jones et al. (1997) that optimal tax for capital should be positive, we extend their analysis in the dynamic environment. In particular, we are interested in the following question:

In response to the investment-specific technological change as shown in Figure xx, what will the equilibrium dynamics of the optimal tax structure be?

Assuming that labor being skilled or unskilled is observable, Angelopoulos et al. (2015) studied tax smoothing in a business cycle with capital-skill complementarity and endogenous skill formation. One of their main findings is that with capital-skill complementarity, the cyclical properties of optimal labor taxes significantly depend on whether the relative supply of skill is restricted or flexible. Both Jones et al. (1997) and Angelopoulos et al. (2015) addressed their problems in the representative-agent framework, and so both abstract from redistribution between the skilled and the unskilled. By contrast, we are in the context of rising wage inequality between skilled and unskilled workers. While Angelopoulos et al. (2015) also find that the optimal tax for capital is zero or not depending on whether the discrminatory tax on labor is allowed or not, the main difference between our paper and Angelopoulos et al. (2015) is as follows. (or, while the reason that why optimal capital tax is zero or not is qualitatively identical to Angelopoulos et al. (2015)) In contrast to assumptions of Angelopoulos et al. (2015) that the relative price of capital follows the stationary AR(1) process, we solves the optimal tax over the transitional dynamics adopting the empirical relevant relative price of capital. With the secular trend of fall in the relative price of capital, our paper is able to address the issue of optimal taxation in response to the secular rise in the wage inequality between skilled and unskilled workers. In contrast, Angelopoulos et al. (2015) can only investigate the fluctuation in wage premium over the business cycles given the assumed the stationary AR(1) process for the relative price of capital.

Ctirad Slavik and Hakki Yazici (2013) instead adopted the Mirrless approach to dynamic taxation, while we adopt the Ramsey approach. A key difference between Ramsey and Mirrlees is that tax schemes are exogenously specified in the former, while the only restiction in the latter is that taxes cannot condition directly on people's skills (agent types).[7] Although

[6]He and Liu (2008) and Angelopoulos et al. (2014) addressed the quantitative effects of some hypothetical tax-policy reforms with capital-skill complementarity.

[7]See Golosov et al. (2006) and Kocherlakota (2010) for more discussion on the differences between Mirrleesian and Ramsey taxation.

less restrictive, the Mirrlees approach often ends up with highly nonlinear tax schemes at the optimum. In a contrast, because tax schemes are specified a priori in Ramsey, they typically take a simpler form. It is arguable that simplicity of tax schemes is important for administering any tax system in the real world (Slemrod and Bakija, 2004). In addition, Ctirad Slavik and Hakki Yazici (2013) assumes the relative price of capital to be constant, while we allow the relative price of capital to change in response to investment-specific technological change.

The rest of the paper is organized as follows. Section 2 introduces the benchmark model. Section 3 addresses Mirrleesin taxation and Section 4 Ramsey taxation. Section 5 considers extensions. We explore quantitative implications of both approaches in Section 6. Section 7 concludes.

2 Benchmark model

We consider a dynamic economy with heterogeneous agent and thus the redistribution motive similar to Werning (2007). The main departures are: (i) production technology is not standard neoclassical but takes the form of empirically plausible capital-skill complementarity as suggested by Krusell et al. (2000), and (ii) aggregate shocks are specific in terms of variations in the relative price of capital equipment associated with investment-specific technological change. In order to highlight the impact of the secular decline in the relative price of capital equipment, we assume that: (i) agents are perfect foresight about variations in the relative price of capital equipment, (ii) the skill composition of the labor force remains constant, and (iii) the labor productivity of the skilled and that of the unskilled are both unchanged.[8] We revisit these assumptions in the extension. There are three types of agents in the economy: heterogeneous households, a representative firm and the government. We describe their optimization problem in turn.

2.1 Economic environment

Time is discrete and the horizon is infinite, indexd by $t = 0, 1, 2, ...$ The economy is popluated by a continuum of infinitely-lived agents with measure one. They are divided into two groups: the skilled (s) and the unskilled (u), with fixed population fraction π^s and π^u, $\pi^s + \pi^u = 1$.

All workers share the same preference and can be represented by a discounted infinite stream of the instantaneous flow of utility derived from consumption c_{it} and raw labor supply

[8] In the benchmark model we set labor productivity to unity for both classes of workers and hence the difference between the skilled and the unskilled is driven solely by whether they are complementary or substitutable with capital equipment in production.

n_{it} for worker $i \in \{s, u\}$:

$$\sum_{t=0}^{\infty} \beta^t U(c_{it}, n_{it}) \tag{1}$$

where $\beta \in (0,1)$ is discount factor, and U is weakly concave and continuous differentiable, $U_c \geq 0$, $U_{cc} \geq 0$, $U_n \leq 0$, $U_{nn} \leq 0$, and $U_{cn} = 0$. The workers own capital and rent it to the representative firm for the use. Let $i \in \{s, u\}$. Capital owned by agent i is denoted k_{it} and its laws of motion is governed by

$$k_{it+1} = (1-\delta)k_{it} + q_t I_t^i, \tag{2}$$

where δ is the depreciation rate; I_t^i is investment at time t; k_0^i is given. Following the literature, we interpret the variable q_t in (2) as the investment-specific technological change that enhances the productivity of newly formed capital relative to prior vintages of capital. One can interpret its inverse, $1/q_t$, as the relative price of new capital; see Hornstein et al. (2005). From Figure 2, we see that $1/q_t$ has been falling over most of the postwar period.

We specify a priori linear tax system for the Ramsey planner as in Werning (2007). In particular, $\tau_{L,t}$ is a time-varying proprotional tax rate on labor income, $\tau_{K,t}$ on rental income from capital, and T_t a lump-sum transfer uniformly across all agents that can be either positive or negative. Note that a common labor tax rate $\tau_{L,t}$ is applied to both skilled and unskilled labor at any point in time, which is the same restriction as imposed by Werning (2007). It can be shown that imposing separate proportional tax rates on skilled and unskilled labor is not incentive compatible with a uniform lump-sum transfer.

The household faces the flow budget constraint as

$$\begin{aligned} c_{it} + \frac{k_{it+1}}{q_t} + b_{it+1} \quad = \quad & (1-\tau_{L,t}) w_{it} n_{it} + \left(\frac{1}{q_t} + (1-\tau_{Kt}) \left(\frac{r_t - \delta}{q_t} \right) \right) k_{it} \\ & + (1 + r_{bt}) b_{it} + T_t \end{aligned} \tag{3}$$

where b_t^i is the one-period, risk-free governemnt bond held by the agents (b_0^i is given), r_{bt} is its rate of returns, r_t denote the before tax rental rate of the capital, and w_{it} is the wage rate for type i agent.

The household's problem is to choose sequences of c_{it}, n_{it}, k_{it+1}, and b_{it+1} to maximize its lifetime utility (1), subject to the laws of motion (2) and a sequence of budget constraints (3). In addition, b_0^i, k_0^i are exogenously given.

The first order conditions for each t (period) can be written as

$$
\begin{aligned}
-U_n(c_{it}, n_{it}) &= U_c(c_{it}, n_{it})(1-\tau_{Lt})w_{it}x_{it} && \text{(4a)}\\
\frac{U_c(c_{it}, n_{it})}{q_t} &= \beta U_c(c_{i,t+1}, n_{i,t+1})\left[\frac{1}{q_{t+1}} + (1-\tau_{K,t+1})\left(\frac{r_{et+1}-\delta}{q_{t+1}}\right)\right] && \text{(4b)}\\
U_c(c_{it}, n_{it}) &= \beta U_c(c_{i,t+1}, n_{i,t+1})[1+r_{b,t+1}] && \text{(4c)}
\end{aligned}
$$

These equations describe workers' optimal behaviors given the factor price $\{r_{Kt}, w_{st}, w_{ut}\}$ and tax policy $\mathcal{T}_t$. Equation (4a) characterizes intratemporal substitution between consumption and labor supply, and the equation (4b) characterize the intertemporal substitution between two consecutive periods. No arbitrage between bonds and capital implies that their after tax return must be the same:

$$
\frac{q_t - q_{t+1}}{q_{t+1}} + q_t\left(1-\tau_{K,t+1}\right)\left[r_{K,t+1} - \frac{\delta}{q_{t+1}}\right] = 1 + r_{b,t+1}
$$

From household's perspective, all assets are perfect substitutes in equilibrium and we don't need to track the individual portfolio allocation within a period.

We define the period 0 price of consumption at time t as:

$$
p_t = \frac{1}{\prod_{s=1}^{t}(1+r_{b,s})}
$$

We normalize $p_0 = 1$.

Furthermore, imposing the Non Ponzi game condition, we can solve the period budget constraint forward and obtain the life-time budget constraint:

$$
\sum_{t=0}^{\infty} p_t\left[c_t^i - (1-\tau_{Lt})w_{it}x_{it}n_t^i\right] = A_0^i + T_0
$$

where $A_0^i = \left[\frac{1}{q_0} + (1-\tau_{K,0})\left(\frac{r_{K,0}-\delta}{q_0}\right)\right]k_0^i + b_0^i$ denotes the wealth of the household in period 0 after tax and interest rate payment are made, which includes both financial asset and capital holdings. Note that the wealth is a sufficient state variable for the household's problem conditional on tax rates and prices.

2.2 Firms

There is a representative firm producing the final good with the production function of the form at time t:

$$Y_t = F\left(K_t,\ N_{st},\ N_{ut}\right) = \left[\mu N_{ut}^{\sigma} + (1-\mu)\left[\lambda K_t^{\rho} + (1-\lambda) N_{st}^{\rho}\right]^{\frac{\sigma}{\rho}}\right]^{\frac{1}{\sigma}}, \tag{5}$$

where Y_t denotes output, K_t capital, N_{st} skilled labor input, and N_{ut} unskilled labor input, with σ, $\rho < 1$. All of them are aggregate variables. For example, $N_{st} = \pi^s n_t^s$ and $N_{ut} = \pi^u n_t^u$. This three-factor production function is similar to that in Krusell et al. (2000). A key feature of this production function is that it allows for different elasticities of substitution between capital and the two types of labor. In particular, the elasticity of subsitution between capital and unskilled labor equals $1/(1-\sigma)$, while the elasticity of subsitution between capital and skilled labor equals $1/(1-\rho)$. There is the so-called "capital-skill complementarity" in production if $\sigma > \rho$.

All markets are competitive and we let the final good be the numeraire. Subject to the production technology (5), the representative firm maximizes its profit

$$\Pi_t = Y_t - w_{st} N_{st} - w_{ut} N_{ut} - r_{Kt} K_t, \tag{6}$$

where w_{st}, w_{ut}, and r_{Kt} are the prices for N_{st}, N_{ut}, and K_t, respectively.

The first order conditions are

$$\frac{\partial Y}{\partial K} = r_K = (1-\mu)\lambda A^{\frac{1-\sigma}{\sigma}} B^{\frac{\sigma}{\rho}-1} K^{\rho-1},$$

$$\frac{\partial Y}{\partial N_s} = w_s = (1-\mu)(1-\lambda) A^{\frac{1-\sigma}{\sigma}} B^{\frac{\sigma-\rho}{\rho}} N_s^{\rho-1},$$

$$\frac{\partial Y}{\partial N_u} = w_u = \mu A^{\frac{1-\alpha-\sigma}{\sigma}} N_u^{\sigma-1},$$

where $B = \lambda K^{\rho} + (1-\lambda) N_s^{\rho}$ and $A = \mu N_u^{\sigma} + (1-\mu) B^{\frac{\sigma}{\rho}}$. Skill premium is then defined to be

$$\xi \equiv \frac{w_s}{w_u} = \frac{(1-\mu)(1-\lambda)}{\mu} \left[\lambda \left(\frac{K}{N_s}\right)^{\rho} + (1-\lambda)\right]^{\frac{\sigma-\rho}{\rho}} \left(\frac{N_u}{N_s}\right)^{1-\sigma}. \tag{7}$$

2.3 Government and fiscal policies

The government provides lump sum transfers to households and pays for exogenous stream of expenditure $\{G_t\}$ with revenues from taxing labor and capital income, as well as issuing a one-period government bond b_t with interest rate r_{bt}. The tax scheme of the government, $\mathcal{T}_t(.)$, is given by

$$\mathcal{T}_t(.) = \tau_{L,t}\left(w_{ut}N_{ut} + \tau_{L,t}w_{st}N_{st}\right) + \tau_{K,t}(r_t - \frac{\delta}{q})K_t - T_t, \tag{8}$$

where $K_t = \sum k_t^i$.

Therefore, the government budget constraint is given by

$$G_t + \left(1 + r_{bt}\right)b_t = \mathcal{T}_t(.) + b_{t+1}, \forall t, \tag{9}$$

We assume $\lim_{t\to\infty}\beta^t b_t = 0$ to rule out Ponzi scheme. Note that the labor tax rates for both types of workers are identical. Later on, we also consider a special case that allows the government to impose different tax rates conditional on the type of workers. In this case, we will assume τ_{st} and τ_{ut} are marginal labor tax rates on skilled and unskilled workers respectively, and the tax revenue collected from $\mathcal{T}_t(.)$ can be written as

$$\mathcal{T}_t(.) = \tau_{st}w_{st}N_{st} + \tau_{ut}w_{ut}N_{ut} + \tau_{K,t}\left(r_{Kt} - \delta\right)K_t - T_t, \tag{10}$$

The tax system $\mathcal{T}_t$ resemble the special case of Erosa and Gervais (2002) that tax authorities are able to set different labor tax rates conditioned on workers' age. These tax systems implies the assumption that government has full information of workers' traits and is able to use this information to design its fiscal policy. This case is implausible, in that a separate-rate tax scheme with the uniform transfer is not incentive compatible. Nevertheless, the case shed light on the structure of optimal capital tax rate that we will discuss later.

2.4 Aggregate resource constraint

We can sum the budget constraint of both types of households and the government to obtain an aggregate resource constraint for the economy:

$$Y_t = C_t + I_t + G_t. \tag{11}$$

2.5 Competitive equilibrium

Different government policies result in different competitive equilibria. The Ramsey problem is to choose a competitive equilibrium through policies that attains the maximum according to some welfare criterion. To pave the way for the analysis of the Ramsey problem, we first describe the competitive equilibrium of our model economy. The following definition of competitive equilibrium (CE) is standard.

Definition 1 *Given initial government bond holdings* $\{b_0^i\}_{i\in\{s,u\}}$ *and initial capital holdings* $\{k_0^i\}_{i\in\{s,u\}}$, *a sequence of investment specific technology changes* $\{q_t\}$, *government purchases* $\{G_t\}$ *and tax schemes* $\mathcal{T}_t(.)$, *a competitive equilibrium is a sequence of market prices* $\{r_{Kt}, w_{ut}, w_{st}\}$ *and non-negative quantities* $\left\{c_t^i, n_t^i, b_{t+1}^i, k_{t+1}^i\right\}_{i\in\{s,u\}}$ *such that*

1. *Given* $\{q_t\}$, $\{G_t\}$,$\{\mathcal{T}_t(.)\}$ *and* $\{r_{K,t}, w_{ut}, w_{st}\}$, *both skilled and unskilled workers maximize their lifetime utility subject to the laws of motion (2) and budget constraint (3).*

2. *Given* $\{q_t\}$, $\{G_t\}$,$\{\mathcal{T}_t(.)\}$ *and* $\{r_{Kt}, w_{ut}, w_{st}\}$, *the representative firm maximizes its profit (6) subject to the production technology (5).*

3. *The government's budget constraints given by (9) hold.*

4. *All markets clear:*

$$\begin{aligned} K_t &= \sum_{i\in\{s,u\}} k_t^i \;\; ; \; C_t = \sum_{i\in\{s,u\}} c_t^i \\ N_{st} &= x_{st}n_{st} \;\; ; \;\; N_{ut} = x_{ut}n_{ut}; \\ Y_t &= C_t + I_t + G_t. \end{aligned}$$

3 Ramsey taxation

The Ramsey taxation problem consist of setting a specific tax program $\mathcal{T}$ so that allocations and prices determined in competitive market can maximize a given social welfare criterion. We assume the Ramsey planner (government) maximizes the utilitarian social welfare function:

$$SWF = \sum_{t=0}^{\infty} \beta^t \sum_{i\in\{s,u\}} \psi^i U\left(c_{it}, n_{it}\right). \tag{12}$$

where $\psi^i \geq 0$ denotes the Pareto weight and assume $\psi^s = \psi^u$. To solve the Ramsey taxation problem, we adopt the primal approach, which follows the standard procedure as described in Werning (2007) and Ljungqvist and Sargent (2012, chapter 16). The primal approach consists of a Ramsey planner that chooses an allocation that can be implemented as a competitive equilibrium defined in definition 1. To derive the implementability condition of each workers, we use agents' first-order conditions to substitute out for the prices and taxes that appear in agents' intertemporal budget constraints.

From the FOCS, $1+r_{b,t+1}=\frac{u'(c_{it})}{\beta u'(c_{i,t+1})}$, so $p_t=\frac{1}{\prod_{s=0}^{t}(1+r_{b,s})}=\frac{\beta^t u'(c_{it})}{u'(c_{i0})}$.
Moreover, $p_t(1-\tau_{it})w_{it}=\frac{(1-\tau_{it})w_{it}}{\prod_{s=0}^{t}(1+r_{b,s})}=\frac{\beta^t u'(c_{it})(1-\tau_{it})w_{it}}{u'(c_{i0})}=-\frac{\beta^t v'(1-n_{it})}{u'(c_{i0})}$.

$$\sum_{t=0}^{\infty}\beta^t\left[U_c(c_{it},n_{it})c_{it}+U_n(c_{it},n_{it})n_t^i\right]-U_c(c_{i0},n_{i0})\left(A_{i0}+T\right)\geq 0\ ,\ i\in\{s,u\}, \tag{13}$$

With these two additional implementabitily constraints together with feasibility constraint, planner can solve planning problem which is equivalent to a set of $\{\mathcal{T}_t(.),B_{t+1}\}$ that maximize (12) in decentralized ecnonomy.

It is important to recognize that when using agents' intratemporal conditions to derive the implementability conditions (13), there is no restriction that heterogeneous agents must face the same proportional labor tax rate as we impose in (10). To account for this restriction, the Ramsey planner must also respect the following equality between agents' intratemporal conditions at each time t

$$\frac{-U_{nt}^s}{U_c^s w_{st}x_{st}}=\frac{-U_{nt}^u}{U_c^u w_{ut}x_{ut}} \tag{14}$$

which guarantees that both skilled and unskilled workers face the same proportional labor tax rate for all t. We can further rewrite this constraint by taking log on both sides and obtain the following conditions :

$$\log\frac{U_{n,t}^s}{U_{n,t}^u}-\log\frac{U_{c,t}^s}{U_{c,t}^u}-\log\xi_t-\log\left(\frac{x_{st}}{x_{ut}}\right)=0 \tag{15}$$

where $U_{l,t}^i=-U_{n,t}^i$. When

In the following analysis, we will first investigate the case that government can set different tax rates on both workers and explain the reason when should tax skilled workers more heavily than unskilled. We then characterize optimal fiscal policy when imposing single labor tax rates on both workers in a economy with capital skill complementarity.

Let θ^i, $\{\beta^t\Gamma_t\}$, $\{\beta^t\Upsilon_t\}$ denote the Lagrange multipliers on the implementability condition (13), the resource constraints (11), and the same proportional labor tax constraint (14) and define $\mathbf{1}_{t=0}$ is an indicator function attached on initial asset term

$$\mathcal{A}^0=\sum_{i\in\{s,u\}}\pi^i\theta^i U_{c0}^i\left(A_{i0}+T\right)$$

the planning problem can be written as

$$
\begin{aligned}
L = \max_{\{c_{st}, c_{u,t}, n_{s,t}, n_{u,t}, K_{t+1}, T\}} &= \sum_{t=0}^{\infty} \beta^t \sum_{i \in \{s,u\}} \left[\psi^i U\left(c_{it}, n_{it}\right) + \theta^i \left(U_c c_{it} + U_n n_{it}\right) \right] \\
&+ \sum_{t=0}^{\infty} \beta^t \Gamma_t \left[f\left(K_t, N_{st,} N_{ut}\right) + \frac{(1-\delta)}{q_t} K_t - C_t - K_{t+1} - G_t \right] \\
&+ \sum_{t=0}^{\infty} \beta^t \Upsilon_t \left[\log \frac{U^s_{n,t}}{U^u_{n,t}} - \log \frac{U^s_{c,t}}{U^u_{c,t}} - \log \xi_t - \log \left(\frac{x_{st}}{x_{ut}} \right) \right] \\
&- \sum_{i \in \{s,u\}} \theta^i U_c(c_{i0}) \left(A_{i0} + T\right)
\end{aligned}
$$

the first order conditions on (15) terms are

$$
\begin{aligned}
\left[\psi^s + \theta^s \left(1 + \frac{U^s_{cc,t} c_{st}}{U^s_{c,t}} \right) - \Upsilon_t \frac{U^s_{cc,t}}{\left(U^s_{c,t}\right)^2} \right] U^s_{c,t} &= \Gamma_t + \mathbf{1}_{t=0} \mathcal{A}^0_{c_s} && (16) \\
\left[\psi^u + \theta^u \left(1 + \frac{U^u_{cc,t} c_{ut}}{U^u_{c,t}} \right) + \Upsilon_t \frac{U^u_{cc,t}}{\left(U^u_{c,t}\right)^2} \right] U^u_{c,t} &= \Gamma_t + \mathbf{1}_{t=0} \mathcal{A}^0_{c_u} && (17) \\
-\left[\psi^s + \theta^s \left(1 + \frac{U^s_{nn,t} n_{st}}{U^s_{n,t}} \right) - \Upsilon_t \frac{U^s_{nn,t}}{\left(U^s_{n,t}\right)^2} \right] U^s_{nt} &= \Gamma_t w_{it} x_{it} - \frac{\Upsilon_t}{\xi_t} \frac{\partial \xi_t}{\partial n_{st}} + \mathbf{1}_{t=0} \mathcal{A}^0_{N_i} && (18) \\
-\left[\psi^u + \theta^u \left(1 + \frac{U^u_{nn,t} n_{ut}}{U^u_{n,t}} \right) + \Upsilon_t \frac{U^u_{nn,t}}{\left(U^u_{n,t}\right)^2} \right] U^u_{nt} &= \Gamma_t w_{it} x_{it} - \frac{\Upsilon_t}{\xi_t} \frac{\partial \xi_t}{\partial n_{ut}} + \mathbf{1}_{t=0} \mathcal{A}^0_{N_i} && (19) \\
-\frac{\Gamma_t}{q_t} + \beta \Gamma_{t+1} \left[r_{t+1} + \frac{1-\delta}{q_{t+1}} - \Upsilon_{t+1} \left[\frac{1}{\xi_{t+1}} \frac{\partial \xi_{t+1}}{\partial K_{et+1}} \right] \right] &= 0 && (20) \\
\sum_{i \in \{s,u\}} \pi^i \theta^i U^i_{c0} &= 0 && (21)
\end{aligned}
$$

These conditions characterize the allocation of competitive equilibrium under the optimal tax rates $\{\tau_{Lt}, \tau_{Kt}, T_t\}$ designed by the Ramsey planner. We can back out the tax rates by comparing the wedge between marginal rate of substitution and marginal rates of transformation.

Optimal tax rates

Since both skilled and unskilled workers face the same marginal labor tax rates, we can combine equation (16)- (19) for $t \geq 1$, $i \in \{s, u\}$, we can obtain intratemporal condition between worker $i's$ consumption and labor supply,

$$\begin{aligned}\frac{-U^s_{n,t}}{w_{st}x_{st}U^s_{c,t}} &= \frac{H^s_{c,t}-\Phi^s_{ct}}{H^s_{nt}-\Phi^s_{nt}}-\Delta^s_t \\ \frac{-U^u_{n,t}}{w_{ut}x_{ut}U^u_{c,t}} &= \frac{H^u_{ct}+\Phi^u_{ct}}{H^u_{nt}+\Phi^u_{nt}}-\Delta^u_t\end{aligned}$$

where

$$\begin{aligned}H^i_{c,t} &= \left[\psi^i+\theta^i\left(1+\frac{U^i_{cc,t}c_{it}}{U^i_{c,t}}\right)\right], & H^i_{n,t} &= \left[\psi^i+\theta^i\left(1+\frac{U^i_{nn,t}n_{it}}{U^i_{n,t}}\right)\right] \\ \Phi^i_{ct} &= \Upsilon_t\frac{U^i_{cc,t}}{\left(U^i_{c,t}\right)^2} & \Phi^i_{nt} &= \frac{U^i_{nn,t}}{\left(U^i_{n,t}\right)^2} \\ \Delta^s_t &= \frac{\frac{\Upsilon_t}{\xi_t}\frac{\partial\xi_t}{\partial n_{st}}}{w_{st}x_{st}U^s_{ct}(H^s_{nt}-\Phi^s_{nt})} & \Delta^u_t &= \frac{\frac{\Upsilon_t}{\xi_t}\frac{\partial\xi_t}{\partial n_{ut}}}{w_{ut}x_{ut}U^u_{ct}(H^u_{nt}+\Phi^u_{nt})}\end{aligned}$$

Since equation (18) and (19) have to be consistent with worker's intratemporal conditions (4a), we can compare these two conditions and back out worker $i's$ labor tax rates:

$$\begin{aligned}\tau_{L,t} &= 1+\frac{U^i_{n,t}}{w_{it}z_{it}U^i_{c,t}} \\ &= 1+\Delta^s_t-\frac{H^s_{c,t}-\Phi^s_{ct}}{H^s_{nt}-\Phi^s_{nt}} = 1+\Delta^u_t-\frac{H^u_{ct}+\Phi^u_{ct}}{H^u_{nt}+\Phi^u_{nt}}\end{aligned} \tag{22}$$

Similarly, when $t \geq 1$, the intertemporal conditions of Ramsey planner in (20) can be expressed as follows

$$\begin{aligned}\frac{U^s_{c,t+1}}{q_t\beta U^s_{c,t+1}} &= \left[\frac{H^s_{c,t+1}-\Phi^s_{ct}}{H^s_{c,t}-\Phi^s_{ct}}\right]\left[\frac{1-\delta}{q_{t+1}}+r_{t+1}-\Delta^K_{t+1}\right] \\ \frac{U^u_{c,t+1}}{q_t\beta U^u_{c,t+1}} &= \left[\frac{H^u_{c,t+1}+\Phi^u_{ct}}{H^u_{c,t}+\Phi^u_{ct}}\right]\left[\frac{1-\delta}{q_{t+1}}+r_{e,t+1}-\Delta^K_{t+1}\right]\end{aligned}$$

where

$$\Delta^K_{t+1} = \frac{\Upsilon_{t+1}}{\xi_{t+1}}\frac{\partial\xi_{t+1}}{\partial K_{et+1}}$$

comparing the intertemporal conditon (4b), one can obtain the optimal tax rates of capital income at period

$$\tau_K = \frac{\Delta_K}{r_e-\frac{\delta}{q}}$$

at steady state. This result implies optimal tax rates $\tau_K > 0$ if $\frac{\partial\xi_{t+1}}{\partial K_{t+1}} > 0$ from capital-skill complementarity and the multipliers $\Upsilon_t > 0$, which implies the conditions

$$\frac{-U^s_{nt}}{U^s_c w_{st}x_{st}} \geq \frac{-U^u_{nt}}{U^u_c w_{ut}x_{ut}}$$

binds at each period. In the following discussion, we will investigate under what conditions that the conditions (15) will be bindings.

3.1 Intuition

To understand the intuition of $\tau_K > 0$, we first return to the case where different proportional tax rates $\{\tau_{st}, \tau_{ut}\}$ are avaiable to the government. Under capital skill complementarity technology specified in (5), we have the following relationship between $\{\tau_{st}, \tau_{ut}, \tau_{K_e,t}\}$ in competitive equilibrium:

$$\begin{aligned}
\left(\frac{U^s_{n,t}}{U^u_{n,t}}\right)\left(\frac{U^u_{c,t}}{U^s_{c,t}}\right) &= \frac{(1-\tau_{st})}{(1-\tau_{ut})}\left(\frac{w_{st}}{w_{ut}}\right) \\
&= \frac{(1-\tau_{st})}{(1-\tau_{ut})\,w_{ut}}\left(\frac{K_{et}}{N_{st}}\right)^{1-\rho} r_{et} \\
&= \frac{(1-\tau_{st})}{(1-\tau_{ut})\,w_{ut}}\left(\frac{K_{et}}{N_{st}}\right)^{1-\rho}\left\{\frac{\left[\frac{U^i_{c,t-1}}{q_{t-1}\beta U^i_{c,t}} - \frac{1}{q_t}\right]}{1-\tau_{K_e,t}} + \frac{\delta_e}{q_t}\right\}
\end{aligned} \tag{23}$$

which derived from (4a) and (4b). When $\{\tau_{st}, \tau_{ut}\}$ can be imposed differently, the government will set zero capital equipment tax and let $\tau_s > \tau_u$ for efficiency and redistributive motive. On the other hand, when the government's tax code was limited to impose $\tau_s = \tau_u$, (23) shows that a positive capital income tax rates on equipment imitates a labor income tax rising with skill. This finding is the application of the Corlett and Hague (1953) to capital-skill complementarity technology. When skilled workers cannot be taxed more heavily than unskilled workers, government can still tax capital equipment that are complementary with skilled worker supply. This is achieved with non-zero income tax on equipment and differential tax rates between structure and equipment.

In a life-cycle model, Erosa and Gervais (2002) found that a zero capital tax result if age-dependent labor taxes were available. Their result suggests that the arise of a positive capital tax typically found in the life-cycle model is meant to mimic age-dependent labor taxes. In offering an intuition for their quantitative result of finding a high optimal capital tax rate, Conesa et al. (2009, p. 41) explicitly explained: "in a life-cycle model in which household labor supply changes with age, if the government cannot condition the tax function on age, it optimally uses the capital income tax to mimic age-dependent labor income taxes." Analogously, endogenous labor supply coupled with capital skill complementarity technology implies a robust role for positive capital income taxation as long as the government cannot condition the tax code on workers' skill type. Our results can be viewed as an application of their findings.

Another implication of our result is to provide a reaon for the optimality of differential capital taxation. When skill-dependent taxation rates are not allowed, capital income tax on equipment should be positive but zero on structure. This result challenges the generality of uniform commodity taxation theorem proposed by Diamond and Mirrlees (1971) and Werning (2007), suggesting tax authority should impose different capital tax rate by asset type when labor tax rates cannot be imposed differently. Our result is also similar to the finding from Hakki and Yazici (2013), where they consider a similar environment under incomplete information. We summarize our finding in the following proposition:

Proposition 1 *At a point of time, Ramsey taxation prescribes that capital structures should go untaxed but capital equipment should be taxed to mimic separate taxes on skilled and unskilled labor.*

3.2 Skill-dependent tax case

This section we remove the constraint (15) and consider the allocations in competitive equilibrium with tax program $\{\tau_{st}, \tau_{ut}, \tau_{Kt}\}$. In this case, the labor tax rate τ_{st}, τ_{ut} can be written as

$$\tau_{it} = 1 + \frac{U^i_{n,t}}{w_{it} z_{it} U^i_{c,t}} = \frac{\theta^i \left[\frac{U^i_{nn,t} n_{it}}{U^i_{n,t}} - \frac{U^i_{cc,t} c_{it}}{U^i_{c,t}} \right]}{H^i_{n,t}},\ i \in \{s, u\}$$

and $\tau_K = 0$ due to $\Delta^K = 0$.

Since we have assumed $U_{cn} = 0$, the term $\frac{U^i_{nn,t} n_{it}}{U^i_{n,t}}$, $\frac{-U^i_{cc} c_{it}}{U^i_{ct}}$ are the inverse Frisch elasicity and curvature of utility function on consumption of worker i, which is exactly equal to inverse of intertemporal substitution elasicity in CRRA preference case. This conditions allows us to check the progressivity of tax rates between τ_s and τ_u. In the following section, we apply the general analysis and formulas laid out above to examples and illustrate how efficiency and redistributive concern affect the shape of skilled and unskilled labor tax rates.

3.2.1 Two examples solved

Since optimal capital income tax in skill-dependent tax is zero in steady state, it is innocuous to set , $\sigma = 1$, $\rho = 0$ and set labor productivity $x_s = x_u = 1$ to retain capital-skill complementarity but simplify the model. In this case, the production function can be expressed as

$$f(K_t, N_{st}, N_{ut}) = \mu N_{ut} + (1 - \mu) K^{\lambda}_{et} N^{1-\lambda}_{st} \tag{24}$$

implying factor prices

$$\begin{aligned} w_u &= \mu && \text{(25a)} \\ w_s &= (1-\mu)(1-\lambda)\left(\frac{K_e}{N_s}\right)^{\lambda} && \text{(25b)} \\ r_e &= (1-\mu)\lambda\left(\frac{K_e}{N_s}\right)^{\lambda-1} && \text{(25c)} \end{aligned}$$

For the preference of workers, We assume $U(c,n)$ is separable and isoelastic between consumption and leisure:

$$U(c_{it},n_{it}) = \frac{c_{it}^{1-\gamma_c}-1}{1-\gamma_c} + \chi\frac{(1-n_{it})^{1-\gamma_n}}{1-\gamma_n} \tag{26}$$

To highlight how Frisch elasticity affect optimal tax rates τ_{st} and τ_{ut}, we shut down government's redistributive concern by setting $\gamma_c = 0$, but $\gamma_n = 1$, and instantaneous utility function reduce to $U(c_{it},n_{it}) = c_{it} + \chi \log(1-n)$, where χ is a number such that $\chi < \mu$ to guarantee the interrior solution $n < 1$. Given this specification, workers Frisch elasicity is

$$\varepsilon_i^F = U_n^i / \left(U_{nn}^i n_i\right) = (1-n_i)/n_i$$

When $t \geq 1$ and government set equal weight between workers ($\psi^s = \psi^u = 1$) and lump-sum transfer is excluded, the condition (16) and 17 imply multipliers of (13) are equal across workers, $\theta^s = \theta^u = \theta$. In addition, the intratemporal conditions (18) combined with (4a) can obtain worker $i's$ tax rates

$$\tau_i = \frac{\theta}{(1+\theta)\varepsilon_i^F + \theta} \tag{27}$$

which shows labor tax rate τ_i inversely related with Frisch elasticites workers.

Given θ and w_i, we can invoke (18 and 19) to derive labor supply function $n(w;\theta)$

$$\begin{aligned} \left[1+\theta\left(\frac{1}{1-n_{it}}\right)\right]\frac{\chi}{1-n_{it}} &= (1+\theta)\,w_{it} \\ 1-n_i(w_i;\theta) &= \frac{1+\sqrt{1+4\theta\,(w_i\,(1+\theta)\,/\chi)}}{2\,((1+\theta)\,w_i/\chi)} \\ &< 1 \end{aligned}$$

In steady state, all factor prices will be exogeneous, denoted as

$$\begin{aligned} r_e^* &= \frac{1}{q}\left[\frac{1}{\beta}+1-\delta_e\right] \\ w_s^* &= (1-\mu)(1-\lambda)\left(\frac{\frac{1}{\beta}-1+\delta_e}{q\,(1-\mu)}\right)^{\frac{\lambda}{\lambda-1}} \\ w_u^* &= \mu \end{aligned}$$

Therefore, if q is large enough such that $w_s^* > w_u^*$, higher skill worker's wage will imply $n_s^* > n_u^*$ and make skilled worker bear lower Frisch elasicity and higher tax burden $\tau_s^* > \tau_u^*$. This example sheds some light which workers should tax more as we shut down the redistribution rule for the government. When investment specific shock rises, skilled worker's wage will go up and increase labor supply, this result make Frisch elasicity of skilled worker lower and bears higher labor tax burden.

Redistributive concerns

Redistributive motive is rmarkable for designing skill-dependent tax under heterogeneous agent framework. To redistribute income between workers, we allow government can impose lump-sum tax T to redistribute wealth. In this example, we still use previous production parameters but change the preference into the following form:

$$U(c_{it}, n_{it}) = \frac{c_{it}^{1-\gamma_c}-1}{1-\gamma_c} - \chi\frac{n_{it}^{1+\gamma_n}}{1+\gamma_n}$$

Here labor elasicity is constant across workers so that we can rule out inverse elasicity rule of public finance in this example. On the other hand, we restrict marginal utility of consumption be strictly concave so that government would like to narrow the gap of consumption between different workers. To simplify the computation without loss of generality, we set $\gamma_c = \gamma_n = 1$ so instantaneous utility function is of the form $U\,(c,n) = \log c - \chi\frac{n^2}{2}$, making labor tax rates equal to

$$\tau_i = \frac{2\theta^i}{(1+2\theta^i)} \quad i \in \{s,u\}$$

in (22). When the government is allowed to set lump-sum taxation, (21) implies that either θ^s or θ^u be negative under optimal taxation scheme:

$$\frac{\theta^s}{c_{s0}} + \frac{\theta^u}{c_{u0}} = 0$$

due to positive values $\pi^s, \pi^u, c_{s0}, c_{u0}$.

Given the specification above, if $A_{s0} \geq A_{u0}$ and q large enough, we can obtain the result that $\theta^s > 0 > \theta^u$, implying $\tau_u < 0 < \tau_s$.We relegate the proof to the appendix.

In sum, when government is able to design a tax scheme dependent on agent's skill in a economy with capital skill complementarity, optimal fiscal policy has the following features:

1. Optimal tax rates on capital tax rates will be zero.
2. Tax authorities would set higher tax rates on skilled workers due to efficiency and redistributive motive.
3. Optimal labor income tax rates on skilled workers τ_{st} will increase with investment specific shocks q_t.

Feature 1 and 2 have discussed in previous context. The intuition behind feature 3 comes from technology of capital skilled complentarity. Since higher investment specific shock would induce higher capital equipment investment, causing lower elasiticity and higher inequality in the economy at the same time. To reduce efficiency loss and improve social inequality, government would therefore impose higher tax burdens rates on skilled workers. The relationship between capital skill complementarity and progressivity labor tax also shed some light on the rationale of positive capital euipment tax rates if government cannot impose workers separately.

This section quantitatively characterizes the equilibrium dynamics of the optimal tax rates over time. The quantitative optimal tax solution we obtain maximizes the social welfare along the transition between the initial steady state and an endogenously determined final steady state. Note that a balanced growth path does not exist with the production function suggested by Krusell et al. (2000) if the investment-specific technological change q_t exhibits a trend; see He and Liu (2008). Therefore, the typical solution methods that involve log-linearizing around a balanced growth path are not applicable to our model. We instead compute the transitional dynamics from the initial steady state to the new steady state by a non-linear solution method.

3.3 Calibration

To carry out quantitative explorations, we calibrate the model parameters to match some key features of the U.S. economy. We take one period in the model to be one calendar year in the data. The details are as follows.

First, considering the availability of relevant data and that the q_t series are relatively stationary before 1960, we choose the year 1963 as our initial steady state so as to line up with the skill premium data shown in Figure 1. Thus, we let $q_{t=1963} \equiv q_0 = 1$.

Using national account statistics as a primal source, McDaniel (2007) calculated series of average tax rates on labor income and capital income for 15 OECD countries for the period 1950-2003. McDaniel's calculation focuses on the taxes part and leaves out the transfers part; as such, her obtained average tax rates can be viewed as the proportional or marginal tax rates of a linear income tax system; see McDaniel (2007) for a formal argument. Browning and Johnson (1984) argued that only the net effect of taxes and transfers is crucial for redistribution, and they provided evidence in support of the hypothesis that a linear income tax can have distributional implications similar to those resulting from the actual tax plus transfer system. Figure 3 shows the evolution of McDaniel's calculated tax rates on labor and capital in the U.S. economy for the period 1963-2013.[9] Since the year 1963 serves as our initial steady state, we simply let McDaniel's (2007) calculated tax rates at 1963 to be the initial U.S. tax rates. The corresponding steady state ratio of government expenditures to GDP is set to 17.5%, which is from NIPA at 1963.

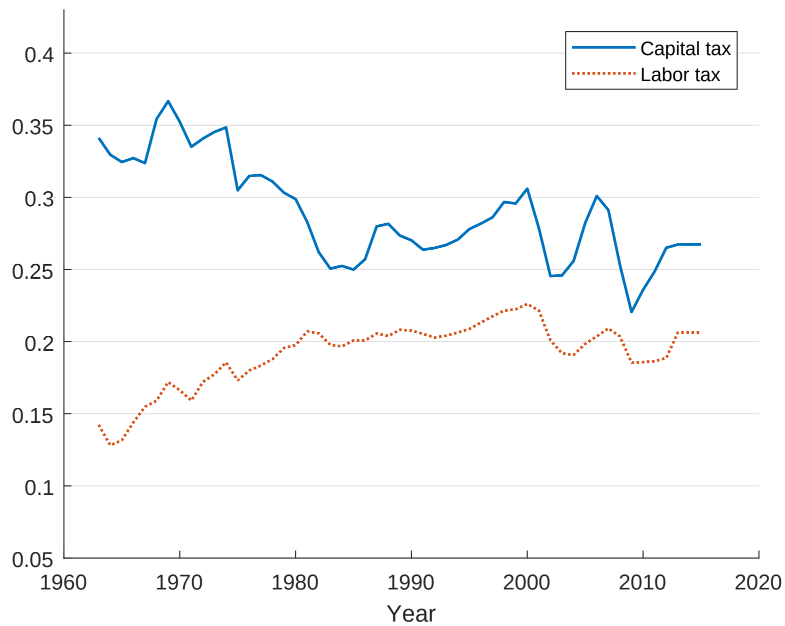

Figure 3: Tax rates in the U.S. calculated by McDaniel (2007)

For the parameters associated with preferences, we set γ_c and γ_n to be 1.5 and 2, and the discount factor β to be 0.98. These are standard in the literature. On the production side, based on Greenwood et al. (1997), we set the depreciation rate for capital to be the weighted average between structures and equipment, which equals $\delta = 0.096$. The parameter σ is chosen to be 0.401, and the parameter ρ is equal to -0.495 . As a result, the elasticity of substitution between capital and skilled labor is about $\frac{1}{1-\rho} \approx 0.67$, while that between

[9] The tax series have been updated to 2013 by McDaniel.

capital and unskilled labor is about $\frac{1}{1-\sigma} \approx 1.67$. Both are consistent with the findings in Krusell et al. (2000).

There are three parameters that remain to be calibrated, which are μ, λ and χ. Parameters μ and λ are related to the production function (5), and χ is the relative weight between consumption and leisure in utility. We employ the method of simulated moments to calibrate these three parameters by matching the following moments of the U.S. economy in 1963 (the initial steady state):

1. The skill premium ξ, when defined as the average annual wage of college graduates relative to that of high-school graduates, equals 1.474 (Autor, 2014).

2. The capital output ratio, $\frac{K}{Y}$, is about 2.27 (NIPA).

3. The ratio of consumption (excluding durable goods) to GDP in 1963 is equal to 0.6 (NIPA).

4. The average income share of capital, which includes both capital structures and capital equipment, i.e.,$\frac{rK}{Y}$, is around 0.3 in 1963 (OECD.Stat).

5. The ratio of gross domestic investment to GDP in 1963 is equal to 0.247 (NIPA).

Finally, we need to specify the distribution of capital between the skilled and the unskilled workers in the initial steady state. We match it to the ratio of the unskilled's wealth relative to the skilled's. Since the data on this ratio are only available from 1989 on, we simply use the 1989 datum, which is equal to 0.57.[10] Thus, we set $K^u/K^s = 0.57$. Although admittedly unsatisfactory, it is the earliest datum we can find with regard to it.

Table 1 summarizes all of our parameter values that are directly set, while Table 2 reports the results of our moment matching. The resulting steady-state capital in competitive equilibrium will serve as the initial capital for our dynamic economy.

We apply the obtained parameter values above to both the benchmark model and the extended model in our quantitative study.

3.4 Time-series data

There are four sets of time-series data that are key to our quantitative study: the tax rates series, the skill premium $\{\xi_t\}$, the investment-specific technological change $\{q_t\}$, and the series $\{x_t = x_t^s/x_t^u\}$ and $\{z_t = z_t^s/z_t^u\}$, which are related to the skill composition and labor productivity of the economy.[11] We briefly describe how we have obtained these data.

[10] The data are from the United States Census Bureau, Asset ownership of households.

[11] We normalize $x_t^u = 1$ and $z_t^u = 1$ to focus on the relative values x_t^s/x_t^u and z_t^s/z_t^u.

Parameter	Symbol	Value	Source
Discount factor	β	0.98	
Elasticity of intertemporal substitution	γ_c	1.5	
Elasticity of leisure	γ_n	2	
Elasticity of N_u and N_s, K composite	σ	0.401	Krusell et al. (2000)
Elasticity of N_s and K	ρ	-0.495	Krusell et al. (2000)
Depreciation rate of K	δ	0.096	Greenwood et al. (1997)
Capital tax rate	τ_K		McDaniel (2007)
Labor income tax rate	τ_s, τ_u		McDaniel (2007)
Government expenditure to GDP ratio	G	0.175	NIPA

Table 1: Parameter values set

Parameter	Symbol	Value	Target	Model (Data)
Income share of N_u	μ	0.327	Skill premium	1.474 (1.474)
Income share of K	λ	0.567	Capital-output ratio	2.27 (2.27)
Utility weight of leisure	χ	17.9	Consumption-output ratio	0.60 (0.6)
			Capital income share	0.288 (0.3)
			Gross investment-output ratio	0.218 (0.247)

Table 2: Parameter values calibrated

Tax rate series The tax rate series are obtained directly from McDaniel (2007). We have already described them. These tax series will be viewed as the representation of the U.S. tax system in our model. For the years after 2013, we let the tax rates remain the same as those in the year 2013 so that the economy can converge to the new steady state. In computing the competitive equilibrium of the U.S. economy, they are applied uniformly to skilled and unskilled labor.

Skill premium $\{\xi_t\}$ Acemoglu and Autor (2011) used data sources including the March CPS to calculate the college/high-school skill premium for full-time, full-year workers for the period 1963-2008. Their approach is sophisticated, in that they managed to hold constant the relative employment shares of the demographic group (including gender, education, and potential experience) across all years of their sample. Autor (2014) extended the data sequence to the year 2012, which is the representation in our Figure 1.[12]

Investment-specific technological change $\{q_t\}$ Gordon (1990) is the seminal work on measuring investment-specific technological change. DiCecio (2009) constructed the relative price of capital by chainweighting the deflator for equipment and software from NIPA. DiCe-

[12]The data are available from Autor's website.

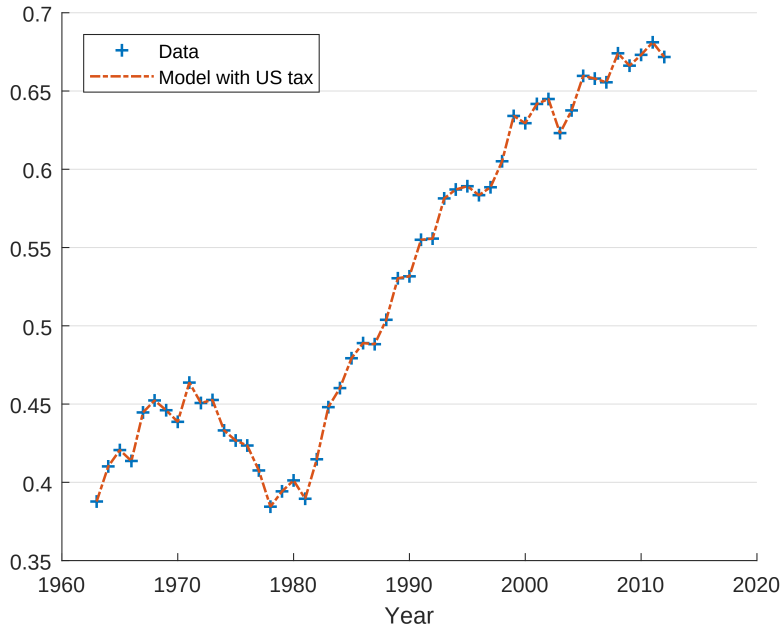

Figure 4: Log skill premium: Data versus Model with US tax

cio (2009)'s data sequence is updated at Federal Reserve Economic Data (FRED), a database maintained by the Federal Reserve Bank of St. Louis. The time series on the relative price of capital shown in Figure 2 are directly taken from the Data. Given the investment-specific technological change q_t, the relative price of capital is equal to $1/q_t$. Thus, the time series on investment-specific technological change $\{q_t\}$ are simply the reciproal of the time series shown in Figure 2.

To obtain q_t beyond the data shown in Figure 2, we first compute the average growth rate of q_t from 2010 to 2014 and let it serve as the growth rate of q_t from 2014 to 2015. We then follow He and Liu's (2008) method by assuming that the growth rates of $\{q_t\}$ beyond year 2014 slow down linearly to zero from 2015 to 2072 and reach a steady state at 2072 and then remain constant from 2072 to 2142. In this way, we construct a time series sequence of $\{q_t\}$ for a length of 180 years consisting of 52 years of data (1963-2014) from FRED and 128 years of artificial data (2015-2142) with 2072 to 2142 being the new steady state.

Varying skill composition/labor productivity $\{x_t\}$ and $\{z_t\}$ Given the "deep" parameters of the model as we have calibrated, one can compute the transitional dynamics of competitive equilibrium of the U.S. economy from the initial steady state to a new steady state, once the tax rate series, $\{q_t\}$, and $\{x_t\}$ are given. We obtain the sequences $\{x_t\}$ and $\{z_t\}$ according to the following algorithm:

1. Starting with $\{x_t(1)\}$, compute the transitional dynamics of competitive equilibrium

and thereby obtain from the model the allocation $\left\{\hat{n}_t^s, \hat{n}_t^u, \hat{K}_t, \hat{N}_{st}, \hat{N}_{ut}\right\}$.

2. Using the real-world data $\{\xi_t\}$, we then calculate $\{z_t\}$ by applying (7), that is,

$$z_t = \frac{\xi_t}{\frac{(1-\mu)(1-\lambda)}{\mu}\left[\lambda\left(\frac{\hat{K}_t}{\hat{N}_{st}}\right)^{\rho} + (1-\lambda)\right]^{\frac{\sigma-\rho}{\rho}}\left(\frac{\hat{N}_{ut}}{\hat{N}_{st}}\right)^{1-\sigma}}.$$

3. By definition, $(N_t^s/N_t^u) = \pi_t^s \cdot n_t^s/\left(\pi_t^u \cdot n_t^u\right)$. Using the obtained $\left\{\hat{n}_t^s, \hat{n}_t^u, \hat{N}_{st}, \hat{N}_{ut}\right\}$ from step 1, we calculate

$$\left(\frac{\pi_t^s}{\pi_t^u}\right) = \frac{\left(\hat{N}_t^s/\hat{N}_t^u\right)}{\left(\hat{n}_t^s/\hat{n}_t^u\right)}.$$

4. Given $\{z_t\}$ obtained from step 2 and $\{\pi_t^s/\pi_t^u\}$ obtained from step 3, we calculate

$$x_t(2) = z_t \cdot \left(\frac{\pi_t^s}{\pi_t^u}\right).$$

5. Iterate until $\{x_t(2)\} \approx \{x_t(1)\}$.

Figure 4 shows the match between the real-world data $\{\xi_t\}$ and the model data on the skill premium under the U.S. tax system. Note that the match is very well.

On the basis of the calibrated parameters and the obtained time-series data, we compute the transitional dynamics of optimal taxation using a non-linear solution method in the spirit of Conesa and Krueger (1999) and He and Liu (2008). The details of the algorithm are relegated to the Appendix.

3.5 Dynamic Ramsey taxation

Figure (5) reports the equilibrium dynamics of Ramsey taxation when government can impose lump-sum tax to redistributive income and design discriminating tax rates on skilled and unskilled workers. When the economy experienced investment specific shocks depicted in (2), we find both marginal tax rates τ_{st} and τ_{ut} increase gradually over time. During 1963-2013, marginal tax of skilled workers is slightly decreasing from 40.9% to 39.6% , while tax rates of unskilled worker rises from -29.6% to -22.5%, implying subsidizing undkilled workers is less efficient if there is an technological improvement on equipment in an economy with capital skill complementarity. However, subsidizing unskilled worker is necessary over

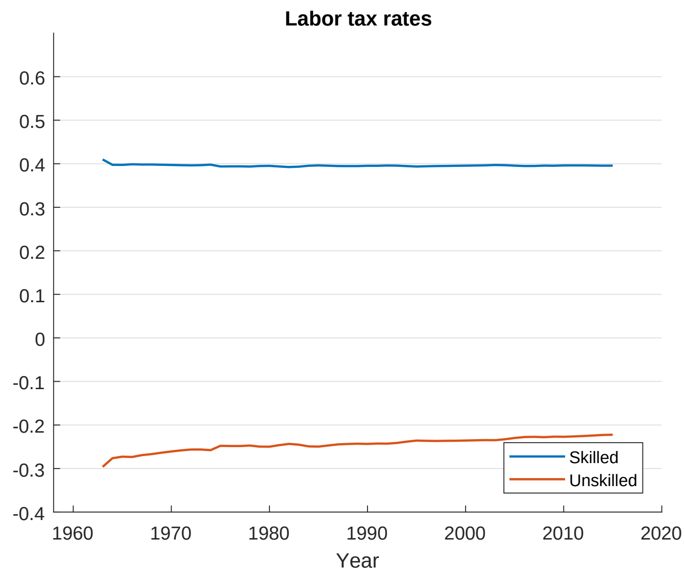

Figure 5: Discriminating labor income tax case

the whole period. Given progressive labor tax design, there is no need for imposing capital tax rates under separating tax program. Besides, due to the increasing pattern of τ_{ut}, indicating the difference between τ_s and τ_u is narrowing.

Figure 6 and 7 illustrate the equilibrium dynamics of of labor and capital income tax rates when government is only able to impose single labor income tax rates between two workers. We also compare our results with U.S data estimated by McDaniel (2007) and displayed in 3. Under single labor tax scheme, we find labor tax rates in our model displays similar pattern with U.S data before 2000, increase from 14.6% to 26.7% during 1963-2012. Meanwhile, income tax on capital, which surges 86.7% then drop to 50% and decrease to to 37.1% is slightly higher than captial tax rate data in U.S.

4 Conclusion

This paper studies optimal taxation in an heterogeneous-agent economy with capital-skill complementarity proposed by Krusell et al. (2000). We firstly show that if government is able to impose separate labor income tax on skilled and unskilled workers, tax rates on skilled workers will higher than unskilled workers due to efficiency and redistributive motive. This effect will be larger when investment specific shock increase, making higher income inequality and lower elasicity on skilled labor supply. On the other hand, when labor tax instrument was restricted on single flat tax rates, uniform commodity taxation results are not likely to hold in this framework, giving rise to a role for interest taxation on equipment but zero on structure. The principles underlying optimal taxation of equipment income in life-cycle

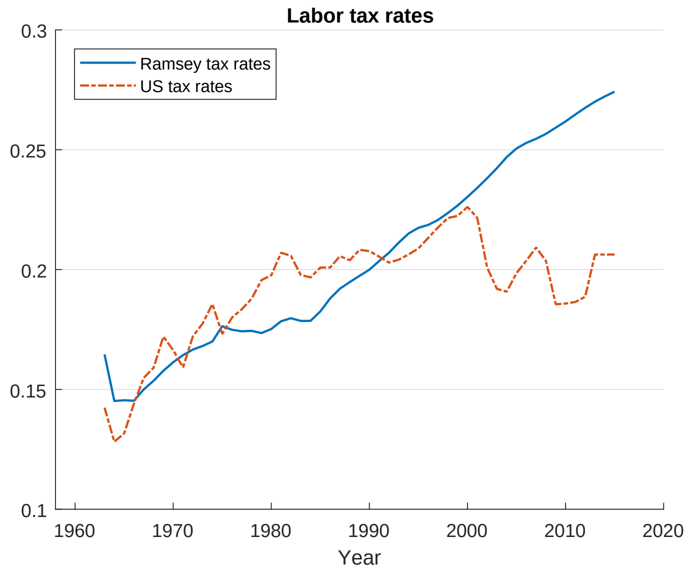

Figure 6: Single labor income tax case

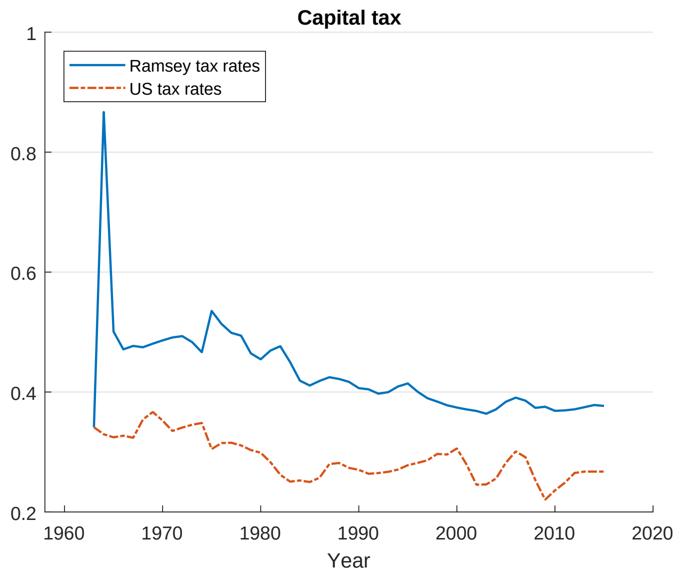

Figure 7: Optimal capital income tax rates

economies relate to the Corlett-Hague intuition: optimal tax rates on capital equipment are set in order to imitate a progressive taxation on skilled workers. We also explore our model quantatatively and apply our theoretical result and propose a tax prescription to tax authority.

5 Appendix

References

Acemoglu, D., Autor, D., 2011. Skills, tasks and technologies: implications for employment and earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. Vol. 4. Elsevier, Ch. 12, pp. 1043–1171.

Autor, D. H., 2014. Skills, education, and the rise of earnings inequality among the other 99 percent. Science 344, 843–851.

Browning, E. K., Johnson, W. R., 1984. The trade-off between equality and efficiency. J. Polit. Econ. 92 (2), 175–203.

Conesa, J. C., Krueger, D., 1999. Social security reform with heterogeneous agents. Rev. Econ. Dyn. 2 (4), 757–795.

Erosa, A., Gervais, M., 2002. Optimal taxation in life-cycle economies. J. Econ Theory 105 (2), 338–369.

Gordon, R. J., 1990. The measurement of durable goods prices. No. 9780226304601. NBER Books.

Greenwood, J., Hercowitz, Z., Krusell, P., 1997. Long-run implications of investment-specific technological change. Am. Econ. Rev. 87 (3), 342–362.

He, H., Liu, Z., 2008. Investment-specific technological change, skill accumulation, and wage inequality. Rev. Econ. Dyn. 11 (2), 314–334.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., Violante, G. L., 2000. Capital-skill complementarity and inequality: a macroeconomic analysis. Econometrica 68 (5), 1029–1054.

McDaniel, C., 2007. Average tax rates on consumption, investment, labor and capital in the OECD: 1950-2003, Working Paper.