Redistributive Capital Taxation Revisited

Özlem Kina, Ctirad Slavík and Hakki Yazici*

November 16, 2019

This paper contributes to the debate on optimal redistributive capital taxation by focusing on a mechanism for taxing capital that has been neglected in the literature. At the heart of our mechanism is the assumption of capital-skill complementarity in the production process, which is the idea that capital is relatively more complementary with skilled labor than it is with unskilled labor. Intuitively, a rise in the capital tax rate depresses capital accumulation, which then decreases the skill premium due to capital-skill complementarity, thereby decreasing the degree of before-tax inequality. To evaluate the importance of this mechanism for the optimal capital tax rate, we build an incomplete markets model with capital-skill complementarity that matches the U.S. economy along several key aggregate and distributional moments. The optimal capital income tax rate is 60%, which is significantly higher than the optimal rate of 47% in an identically calibrated model without capital-skill complementarity. The skill premium falls from the calibrated value of 1.9 in the current steady state to about 1.7 in the final steady state under the optimal tax system.

JEL classification: E25, J31.
Keywords: Capital taxation, capital-skill complementarity, inequality.

*Özlem Kina, European University Institute, Florence, Italy. Email: ozlem.kina@eui.eu.
Ctirad Slavík, CERGE-EI. A joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Czech Academy of Sciences. Politických vězňů 7, Praha 1, 111 21, Czech Republic. Email: ctirad.slavik@cerge-ei.cz.
Hakki Yazici, Sabanci University. Orta Mah. Universite Cad. No:27 34956, Orhanli, Istanbul, Turkey. Email: hakkiyazici@sabanciuniv.edu.
We would like to thank Arpad Abraham, Dirk Krueger, Remzi Kaygusuz, Ramon Marimon, and CEF 2018, CEF 2019, EEA-ESEM 2018, ASSA 2019 seminar participants for helpful suggestions and comments.

1 Introduction

The optimal tax rate on capital has long been debated. The supporters of capital tax cuts stress the efficiency costs associated with slowing down of capital accumulation, and hence, reduced output growth. The proponents of higher capital taxes often bring up their redistributive benefits: wealth is often quite unequally distributed across the population, and therefore, increasing capital taxes in favor of lower labor taxes decreases after-tax inequality. Aiyagari (1995) and Domeij and Heathcote (2004), among others, have shown that redistributive benefits of capital taxation can be large enough to imply significant optimal tax rates on capital. In this paper, we contribute to the debate on the optimal capital tax rate by proposing a mechanism through which capital taxes imply additional redistributive benefits and by quantifying the implications of this mechanism for the optimal capital tax rate. We find that the mechanism we propose implies that the optimal tax rate on capital should be significantly higher than what the standard models deliver.

At the heart of our mechanism is the assumption of capital-skill complementarity in the production process, which is the idea that capital is relatively more complementary with skilled labor than it is with unskilled labor.[1] Intuitively, a rise in the capital tax rate depresses capital accumulation, which then decreases the skill premium - the wages of skilled workers relative to unskilled ones - due to capital-skill complementarity. Since skilled workers normally earn higher wages and have more assets, the decline in skill premium increases social welfare from the perspective of a government that values equality.

We measure the quantitative significance of this mechanism for optimal capital tax rate using a model that embeds capital-skill complementarity into an incomplete markets model a la Aiyagari (1994), where individuals face idiosyncratic wage risk. We choose this model as it allows for sufficiently rich and realistic modeling of earnings and wealth inequality, which

[1] Griliches (1969) was the first to formalize and test the capital-skill complementarity hypothesis. Since then, it has received much attention from economists and has been successful in explaining the evolution of inequality in the returns to education. Among others, see Fallon and Layard (1975), Krusell et al. (2000), Flug and Hercowitz (2000), and Duffy et al. (2004).

are key to accurately assessing redistributive benefits of capital taxation. We consider two versions of the model that differ from each other solely in terms of the aggregate production functions. In the first economy we model capital-skill complementarity by assuming a production function that features a higher degree of complementarity between equipment capital and skilled labor than between equipment capital and unskilled labor, as documented empirically for the U.S. economy by Krusell et al. (2000). We also build a second economy with a standard Cobb-Douglas production function as a benchmark. We synchronize the two model economies by calibrating each one separately to the current U.S. economy along selected dimensions under status-quo capital and labor tax system. Even though we calibrate the economies separately, they are fully comparable since they are calibrated to identical moments. This calibration procedure delivers all structural parameters to be identical across the two economies except for parameters related to the production side of the economy.

When we compute the optimal capital tax rate for each economy taking into account transition and assuming a Utilitarian social welfare function with equal weights on all agents, we find that the optimal capital tax rate for the capital-skill complementarity economy is significantly higher than that in the Cobb-Douglas economy, with respective optimal rates of 60% vs. 47%. Accordingly, the average labor tax rate is smaller in the economy with capital-skill complementarity. Over transition, the skill premium falls from the calibrated value of 1.9 in the current steady state to about 1.7 in the final steady state in the capital-skill complementarity economy while it stays constant in the Cobb-Douglas economy.[2] It is this indirect redistribution benefit of capital taxes that gives rise to such higher optimal taxes in the economy with capital-skill complementarity. Our analysis shows that the debate over the correct tax rate on capital should take into account the presence of capital-skill complementarities in production.

[2]By taxing capital at such high rate in the capital-skill complementarity economy, the government is effectively offseting several decades of rising inequality in the returns to education.

2 Literature Review

This paper is related to the literature on taxation of capital income which is a controversial issue in the macroeconomics literature. In representative-agent paradigm, Chamley (1986) and Judd (1985) have shown that the optimal capital income tax rate is zero in the long-run. Aiyagari (1995) has shown that optimal long-run capital income taxes might be positive when there is uninsured labor income risk due to incomplete markets. He points out that optimal steady state capital income tax is between 25% and 45% depending on parameters.

Imrohoroglu (1998) has shown that in an overlapping generation model with borrowing constraints and uninsurable idiosyncratic earnings risk, optimal tax rate on capital income is 15% at the steady state. Conesa et al. (2009) analyze optimal steady state capital income taxes in an Aiyagari economy. They find that optimal capital income tax rate is positive at 36%. Domeij and Heathcote (2004) compare a representative agent economy with an economy in which agents are exposed to uninsurable idiosyncratic labor income risks. They find that under representative-agent economy, eliminating capital income tax brings large welfare gain. However, when market incompleteness is introduced, most households have welfare losses under zero capital income tax. They find that when markets are incomplete, decreasing capital taxes increases productive efficiency (decrease consumption in the short run and increase it in the long run, and the latter dominates the former) but decreases redistribution (because it is the rich that hold most of the capital). The optimal rate that comes out of this trade off in their analysis is about 40% when transitional dynamics are taken into account.

Jones et al. (1997) provide an important backdrop for the current paper. The authors analyze optimal Ramsey taxation in a growth model with two types of labor, skilled and unskilled, and show that optimal long-run capital tax rate is positive if the labor tax rate on the two types of workers has to be the same. Slavík and Yazici (2014) studies the optimality of differential capital taxation using a model with structure capital and equipment capital in the presence of equipment-skill complementarity. They show that it is optimal to tax equipments at a higher rate than structures since depressing accumulation of equipment capital lowers the skill premium. Slavík and Yazici (2019) analyses the consequences of eliminating capital tax differentials across different capital assets using an incomplete markets model with two types of capital and equipment-skill complementarity. They find that the reform not only increases productive efficiency by reallocating capital from low to high return capital but also improves equality by decreasing the skill premium. However, none of these papers analyzes optimal capital taxes in a Ramsey framework with risk.

3 Model

The economy consists of a unit measure of individuals, a firm, and a government all of whom live forever. In our baseline model, we assume an aggregate production function that features capital-skill complementarity. Later on, for comparison, we also consider an economy that combines capital and labor using a standard Cobb-douglas production function.

Endowments and Preferences. In each period, people are endowed with one unit of time. Ex-ante, they differ in their skill levels: they are born either skilled or unskilled, $i \in \{u, s\}$. Skilled agents can only work in the skilled labor sector and unskilled agents only in the unskilled labor sector. The skill types are permanent. The total mass of the skilled agents is denoted by π_s, the total mass of the unskilled agents is denoted by π_u. In the quantitative analysis, skill types correspond to educational attainment at the time of entering the labor market. Agents who have college education or above are classified as skilled agents and the rest of the agents are classified as unskilled agents.

In addition to heterogeneity between skill groups, we model heterogeneity within each skill group by assuming that agents face idiosyncratic labor productivity shocks over time. The productivity shock, denoted by z, follows a type-specific Markov chain with states $Z_i = \{z_{i,1}, ..., z_{i,I}\}$ and transitions $\Pi_i(z'|z)$. An agent of skill type i and productivity level z who works l units of time produces $l \cdot z$ units of effective i type of labor. As a result, her wage per unit of time is $w_i \cdot z$, where w_i is the wage per effective unit of labor in sector i.

Preferences over sequences of consumption and labor, $(c_{i,t}, l_{i,t})_{t=0}^{\infty}$, are defined using a separable utility function

$$E_i \sum_{t=0}^{\infty} \beta_i^t \Big(u(c_{i,t}) - v(l_{i,t}) \Big),$$

where β_i is the time discount factor, which is allowed to be different across skill types.[3] For each skill type, the unconditional expectation, E_i, is taken with respect to the stochastic processes governing the idiosyncratic labor shock. There are no aggregate shocks.

Technology. There is a constant returns to scale production function: $Y = F(K_s,\ K_e,\ L_s,\ L_u)$, where K_s and K_e refer to aggregate structure capital and equipment capital and L_s and L_u refer to aggregate effective skilled and unskilled labor, respectively. δ_s and δ_e denote the depreciation rates of structure and equipment capital, respectively.

The key feature of technology is equipment-skill complementarity, which means that the degree of complementarity between equipment capital and skilled labor is higher than that between equipment capital and unskilled labor. This implies that an increase in the stock of equipment capital decreases the ratio of the marginal product of unskilled labor to the marginal product of skilled labor. In a world with competitive factor markets, this implies that the skill premium, defined as the ratio of skilled to unskilled wages, is increasing in equipment capital. Structure capital, on the other hand, is assumed to be neutral in terms of its complementarity with skilled and unskilled labor. These assumptions on technology are in line with the empirical evidence provided by Krusell et al. (2000) for the United States.[4] Letting $\partial F / \partial m$ be the partial derivative of function F with respect to variable m, we formalize these assumptions as follows.

Assumption 1. $\frac{\partial F / \partial L_s}{\partial F / \partial L_u}$ *is independent of* K_s.

[3] Attanasio et al. (1999) provide empirical evidence for differences in discount factors across education groups. In our quantitative analysis, we calibrate the discount factors so as to match the observed difference in wealth between skilled and unskilled agents. The calibration implies that the skilled discount factor is slightly larger than the unskilled discount factor, which is in line with the empirical evidence provided by Attanasio et al. (1999).

[4] Flug and Hercowitz (2000) provide evidence for equipment-skill complementarity for a set of countries.

Assumption 2. $\frac{\partial F/\partial L_s}{\partial F/\partial L_u}$ *is strictly increasing in* K_e.

There is a representative firm, which hires the two types of labor and rents the two types of capital to maximize profits in each period. In any period t, its maximization problem reads:

$$\max_{K_{s,t},K_{e,t},L_{s,t},L_{u,t}} \quad F(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - r_{s,t}K_{s,t} - r_{e,t}K_{e,t} - w_{s,t}L_{s,t} - w_{u,t}L_{u,t},$$

where $r_{s,t}$ and $r_{e,t}$ are the rental rates of structure and equipment capital and $w_{u,t}$ and $w_{s,t}$ are the wages rates paid to unskilled and skilled effective labor in period t.

Asset Market Structure. Government debt is the only financial asset in the economy. It has a one period maturity and return R_t in period t. Consumers can also save through the two types of capital. In the absence of aggregate shocks, the returns to savings in the form of the two capital types are certain, as is the return on government bonds. Therefore, all three assets must yield the same after-tax return in equilibrium, $R_t = 1+(r_{s,t}-\delta_s)(1-\tau_t) = 1 + (r_{e,t} - \delta_e)(1 - \tau_t)$. As a result, one does not need to distinguish between savings via different types of assets in the consumer's problem. Consumers' (total) asset holdings will be denoted by a and $\mathcal{A} = [0, \infty)$ denotes the set of possible asset levels that agents can hold. Our assumptions imply that, in every period, the total savings of consumers must be equal to the total borrowing of the government plus the total capital stock in the economy.

Government. The government uses linear taxes on capital income net of depreciation. Let $\{\tau_t\}_{t=0}^{\infty}$ be the sequence of tax rates on capital income. It is irrelevant for our analysis whether capital income is taxed at the consumer or at the corporate level. We assume without loss of generality that all capital income taxes are paid at the consumer level. The government taxes labor income using a sequence of possibly non-linear functions $\{T_t(y)\}_{t=0}^{\infty}$, where y is labor income and $T_t(y)$ are the taxes paid by the consumer. We follow Heathcote et al. (2017) and assume that tax liability given labor income y is defined as:

$$T(y) = \bar{y}\left[\frac{y}{\bar{y}} - \lambda\left(\frac{y}{\bar{y}}\right)^{1-\tau_l}\right], \tag{1}$$

where $\bar{y}$ is the mean labor income in the economy, $1-\lambda$ is the average tax rate of a mean income individual and τ_l controls the progressivity of the tax code. The government uses taxes to finance a stream of expenditure $\{G_t\}_{t=0}^{\infty}$ and repay government debt $\{D_t\}_{t=0}^{\infty}$.

Competitive Equilibrium. Before we provide a formal definition of equilibrium, it is useful to introduce some concepts and notation. The initial state of a worker of type i is fully described by the worker's initial productivity and asset holding. Let $v_0 = (z_0, a_0) \in \mathcal{V}_i = \mathcal{Z}_i \times \mathcal{A}$ denote initial state of a worker of type i. Let $\lambda_i^0(v_0)$ be the exogenously given period 0 distribution of workers of type i across productivities and assets. Denote the partial history of productivity shocks from period 1 up to period t by $z^t \equiv (z_1, ..., z_t)$. Also, denote the conditional probability of z^t for agent of skill type i given period 0 productivity z_0 by $P_{i,t}(z^t|z_0)$. For each agent type, this unconditional probability is achieved by applying the transition probability matrix $\Pi_i(z'|z)$ recursively. We denote by Z_i^t the set in which z^t lies for an agent of type i in period t. At any point t in time, a worker's state is given by (v_0, z^t). It is understood that $z^0 = v_0$ and that $(v_0, z^0) = v_0$.

Definition: Given a couple of initial distributions, $\{\lambda_i^0(v_0)\}_{i=u,s}$, a competitive equilibrium consists of an allocation $\left(\left\{c_{i,t}(v_0, z^t), l_{i,t}(v_0, z^t),\ a_{i,t+1}(v_0, z^t)\right\}_{i\in\{u,s\}, v_0\in\mathcal{V}_i, z^t\in Z_i^t}, K_{s,t},\right.$ $\left.K_{e,t}, L_{s,t}, L_{u,t}\right)_{t=0}^{\infty}$, a policy $\left(T_t(\cdot), \tau_t, D_t, G_t\right)_{t=0}^{\infty}$, and a price system $(r_{s,t}, r_{e,t}, w_{s,t}, w_{u,t}, R_t)_{t=0}^{\infty}$ such that:

1. Given the policy and the price system, for each $i \in \{u, s\}$ and $v_0 \in \mathcal{V}_0$, the allocation

$\left(\left\{c_{i,t}(v_0,z^t), l_{i,t}(v_0,z^t),\ a_{i,t+1}(v_0,z^t)\right\}_{z^t\in Z_i^t}\right)_{t=0}^{\infty}$ solves consumer's problem, i.e.,

$$V_i^0(v_0) = \max_{\left(\left\{c_{i,t}(z^t),l_{i,t}(z^t),a_{i,t+1}(z^t)\right\}_{z^t\in Z_i^t}\right)_{t=0}^{\infty}} \sum_{t=0}^{\infty}\sum_{z^t\in Z_i^t} P_{i,t}(z^t|z_0)\beta_i^t u(c_{i,t}(z^t)) - v(l_{i,t}(z^t)) \qquad s.t.$$

$$\forall t\geq 0, z^t,$$

$$c_{i,t}(z^t) + a_{i,t+1}(z^t) \leq l_{i,t}(z^t)w_{i,t}z_t - T_t(l_{i,t}(z^t)w_{i,t}z_t) + R_t a_{i,t}(z^{t-1}),$$

where z^{-1} is the null history and,

$$\forall t\geq 0, z^t, \qquad c_{i,t}(z^t)\geq 0, a_{i,t+1}(z^t)\in\mathcal{A}, l_{i,t}(z^t)\geq 0.$$

2. In each period $t\geq 0$, taking factor prices as given, $(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})$ solves the following firm's problem:

$$\max_{K_{s,t},K_{e,t},L_{s,t},L_{u,t}} \qquad F(K_{s,t},K_{e,t},L_{s,t},L_{u,t}) - r_{s,t}K_{s,t} - r_{e,t}K_{e,t} - w_{s,t}L_{s,t} - w_{u,t}L_{u,t}.$$

3. Markets for assets, labor and goods clear: for all $t\geq 0$,

$$K_{s,t} + K_{e,t} + D_t = \sum_{i=u,s}\pi_i\int_{\mathcal{V}_i}\sum_{z^{t-1}\in Z_i^{t-1}} P_{i,t-1}(z^{t-1}|z_0)a_{i,t}(v_0,z^{t-1})d\lambda_i^0(v_0),$$

$$L_{i,t} = \pi_i\int_{\mathcal{V}_i}\sum_{z^t\in Z_i^t} P_{i,t}(z^t|z_0)l_{i,t}(v_0,z^t)z_t d\lambda_i^0(v_0),\ \text{for } i=u,s,$$

$$G_t + C_t + K_{s,t+1} + K_{e,t+1} = F(K_{s,t},K_{e,t},L_{s,t},L_{u,t}) + (1-\delta_s)K_{s,t} + (1-\delta_e)K_{e,t},$$

where

$$C_t = \sum_{i=u,s}\pi_i\int_{\mathcal{V}_i}\sum_{z^t\in Z_i^t} P_{i,t}(z^t|z_0)c_{i,t}(v_0,z^t)d\lambda_i^0(v_0)$$

is aggregate consumption in period t.

4. The government's budget constraint is satisfied every period: for all $t \geq 0$,

$$G_t + R_t D_t = D_{t+1} + \sum_{j=s,e} \tau_t (r_{j,t} - \delta_j) K_{j,t} + T_{agg},$$

where

$$T_{agg} = \sum_{i=u,s} \pi_i \int_{\mathcal{V}_i} \sum_{z^t \in Z_i^t} P_{i,t}(z^t|z_0) T_t(l_{i,t}(v_0, z^t) w_{i,t} z_t) d\lambda_i^0(v_0)$$

denotes aggregate labor income tax revenue.

3.1 Cobb-Douglas Economy

Since the main goal of this paper is to understand the effects of capital-skill complementarity on optimal capital income taxation, in our second economy we eliminate capital-skill complementarity from the production function and preserve all other properties of our first model. We do not distinguish between equipment capital and structure capital, we have one type of capital which depreciates every period at rate δ. We use the following Cobb-Douglas production function:

$$Y = F(K, N_s, N_u) = AK^{\theta}(\mu N_s + N_u)^{1-\theta}$$

where A is total factor productivity, θ is the usual Cobb-douglas parameter, and μ is a parameter that allows for skilled labor to be more effective than unskilled labor. Under this production function, the ratio of marginal product of skilled labor to marginal product of unskilled labor, hence the skill premium, is constant and equals to μ. That is, skill premium does not depend on the aggregates in general, and the amount of capial in the economy in particular. The changes in the aggregate capital level do not affect skill premium, therefore here capital income taxation has no impact on wage inequality. The definition of competitive equilibrium for this economy is very similar to the definition given for the capital-skill complementarity economy, and hence is relegated to Appendix A.1.

4 The Optimal Tax Problem

We consider the following optimal fiscal policy reform. The economy is initially at a steady state under a status-quo fiscal policy. Given the initial distribution of workers across productivity-asset space implied by this steady state, the government introduces a ones and for all unannounced change in the tax rate that applies to capital income. At the same time, to ensure that its budget holds under the government spending and bond holding levels given by the initial steady state, government adjusts the parameter that controls the average labor income tax, $\{\lambda_t\}_{t=0}^{\infty}$, along the transition. In our baseline analysis, we assume that the government evaluates the consequences of the refom by aggregating citizens' welfare using a Utilitarian social welfare function that puts an equal weight on all agents in the initial steady state. The optimal tax problem then is to find the tax rate τ on capital income that leads to the competitive equilibrium that achieves the highest social welfare. Formally, government solves the following problem:

$$\max_{\tau} \sum_{i=u,s} \pi_s \int_{\mathcal{V}_i} V_i^0(v_0;\tau) d\lambda_i^0(v_0) \tag{2}$$

such that, for every τ, $V_i^0(v_0;\tau)$ is the value implied by the corresponding competitive equilibrium.

5 Calibration

We first explain how we calibrate our baseline model with capital-skill complementarity to the U.S. economy. To do so, we first fix a number of parameters to values from the data or from the literature. These parameters are summarized in Table 1. We then calibrate the remaining parameters so that the steady state of the model economy under status-quo U.S. fiscal policy matches the U.S. data along selected key dimensions. Our calibration procedure is summarized in Table 2. For data availability reasons, we focus on working age males, when we compare the model with data. This concerns the skill premium and educational attainment as well as the idiosyncratic productivity processes.

Preferences and Demographics. One period in our model corresponds to one year. We assume that the period utility function takes the form

$$u(c) - v(l) = \frac{c^{1-\sigma}}{1-\sigma} - \phi\frac{l^{1+\gamma}}{1+\gamma}.$$

Here, σ equals the coefficient of relative risk aversion while γ controls Frisch elasticity of labor supply. In the benchmark case, we use $\sigma = 2$ and $\gamma = 1$. These are within the range of values that have been considered in the literature. We calibrate ϕ to match the average labor supply. The discount rate for each type, β_i, is calibrated internally as explained below.

The fraction of skilled agents is calculated to be 0.3544 using the Current Population Survey (CPS) data for 2018. We focus on males who are 25 years old or older and who have earnings. To be consistent with Krusell et al. (2000), skilled people are defined as those who have at least 16 years of schooling.

Technology. In the baseline economy with capital-skill complementarity, we assume that the production function takes the same form as in Krusell et al. (2000):

$$Y = F(K_s, K_e, L_s, L_u) = K_s^{\alpha}\left(\nu\left[\omega K_e^{\rho} + (1-\omega)L_s^{\rho}\right]^{\frac{\eta}{\rho}} + (1-\nu)L_u^{\eta}\right)^{\frac{1-\alpha}{\eta}}. \tag{3}$$

In this formula, ρ controls the degree of complementarity between equipment capital and skilled labor while η controls the degree of complementarity between equipment capital and unskilled labor. Krusell et al. (2000) estimate ρ and η, and we use their estimates. Their estimates of these two parameters imply that equipment capital is more complementary with skilled than unskilled labor. The parameter α gives the income share of structure capital. The other two parameters in this production function, ω and ν jointly control the income shares of equipment capital, skilled labor and unskilled labor. These three parameters are calibrated internally, as explained in detail later.

Government. We assume that the government consumption-to-output ratio equals 16%, which is close to the average ratio in the United States during the period 1980 – 2012, as reported in the National Income and Product Accounts (NIPA) data. We assume a government debt of 60% of GDP, which approximates the U.S. government debt between 1990 and the Great Recession (see the FRED series GFDEGDQ188S).

We follow Trabandt and Uhlig (2011) and assume that the status quo tax rate on capital income is $\tau_k = 36\%$. As for labor income taxes, modelling the progressivity of the U.S. tax system may be important for our exercise since progressive tax systems can already provide substantial redistribution from skilled workers to unskilled workers, dwarfing the importance of taxing capital for indirect redistribution. To approximate the progressive U.S. labor tax code, we follow Heathcote et al. (2017). Using the PSID data for 2000 – 2006 and the TAXSIM program, they estimate $\tau_l = 0.18$. We use their estimate and calibrate λ to clear the government budget, following their procedure.[5]

Wage Risk. We cannot identify the mean levels of the idiosyncratic labor productivity shock z for the two types of agents separately from the remaining parameters of the production function and therefore set $E[z] = 1$ for both skilled and unskilled. This assumption implies that w_i corresponds to the average wage rate of agents of skill type i. Thus, skill premium in the model economy is given by w_s/w_u. We assume that the processes for z differ across the two types of agents. Specifically, we assume that for all $i \in \{u, s\} : \log z_{t+1} = \rho_i \log z_t + \varepsilon_{i,t}$. Following Krueger and Ludwig (2015), we set $\rho_s = 0.9690, var(\varepsilon_s) = 0.0100$, $\rho_u = 0.9280$, $var(\varepsilon_u) = 0.0192$. We approximate these processes by finite number Markov chains using the Rouwenhorst method described in Kopecky and Suen (2010).

[5] It is likely that estimations using different data sets might give different estimates. See, for instance, Bakis et al. (2015) who use CPS data for the period 1979-2009 and finds $\tau_l = 0.17$.

Table 1: Benchmark Parameters

Parameter	Symbol	Value	Source
Preferences			
Relative risk aversion parameter	σ	2	
Inverse Frisch elasticity	γ	1	
Technology			
Structure capital depreciation rate	δ_s	0.056	GHK
Equipment capital depreciation rate	δ_e	0.124	GHK
Measure of elasticity of substitution between equipment capital K_e and unskilled labor L_u	η	0.401	KORV
Measure of elasticity of substitution between equipment capital K_e and skilled labor L_s	ρ	-0.495	KORV
Technology (Cobb-Douglass Model)			
Capital's share of output	θ	1/3	
Skill premium	μ	1.9	CPS
Depreciation rate of capital	δ	0.0787	
Relative supply of skilled workers	π_s	0.3544	CPS
Productivity			
Productivity persistence of skilled workers	ρ_s	0.9690	KL
Productivity volatility of skilled workers	$var(\varepsilon_s)$	0.0100	KL
Productivity persistence of unskilled workers	ρ_u	0.9280	KL
Productivity volatility of unskilled workers	$var(\varepsilon_u)$	0.0192	KL
Government polices			
Labor tax progressivity	τ_l	0.18	HSV
Linear tax rate on capital income	τ_k	0.36	TU
Government consumption	G/Y	0.16	NIPA
Government debt	D/Y	0.60	FRED

This table reports the benchmark parameters that we take directly from the literature or the data. The acronyms GHK, KORV, HSV, KL and TU stand for Greenwood et al. (1997), Krusell et al. (2000), Heathcote et al. (2017), Krueger and Ludwig (2015), and Trabandt and Uhlig (2011) respectively. NIPA stands for the National Income and Product Accounts, CPS for Current Population Survey and FRED for the FRED database of the Federal Reserve Bank of St. Louis.

Table 2: Benchmark Calibration Procedure

Parameter	Symbol	CSC	Cobb-Douglas	Target	Source
prod. func. parameter	ω	0.3332	-	labor share (2/3)	NIPA
prod. func. parameter	ν	0.6205	-	skill premium ($\frac{w_s}{w_u}$=1.9)	CPS
prod. func. parameter	α	0.1920	-	share of equipments ($\frac{K_e}{K}$=1/3)	
total factor productivity	A	1	0.4037	output level of CSC economy	
skilled discount factor	β_s	0.9415	0.9415	capital to output ratio ($\frac{K}{Y}=2$)	NIPA, FAT
unskilled discount factor	β_u	0.9365	0.9365	rel. skilled wealth=2.68	US Census
tax function parameter	λ	0.8142	0.8142	govn. budget balance	
disutility of labor	ϕ	65.9705	65.9705	labor supply (1/3)	

This table reports our benchmark calibration procedure. The production function parameters α,ν and ω control the income share of structure capital, equipment capital, skilled and unskilled labor in output. The tax function parameter λ controls the labor income tax rate of the mean income agent. Relative skilled wealth refers to the ratio of the average skilled asset holdings to the average unskilled asset holdings. The acronym NIPA stands for the National Income and Product Accounts and FAT stands for the Fixed Asset Tables.

Internal Calibration. There are still seven parameter values left to be determined: these are the three production function parameters, α, ω and ν, the labor disutility parameter ϕ, the discount factors β_s and β_u and the parameter governing the overall level of taxes in the tax function, λ. The income shares of equipment capital, skilled labor and unskilled labor are governed by ω and ν, and α governs the income share of structure capital. We calibrate α, ω and ν so that (i) the share of equipment capital in total capital is 1/3 as in the data, (ii) the labor share equals 2/3 (approximately the average labor share in 1980 – 2010 as reported in the NIPA data), and (iii) the skill premium w_s/w_u equals 1.9 (computed using CPS data for 2010s). We choose ϕ so that the aggregate labor supply in steady state equals 1/3 (as is commonly assumed in the macro literature). We calibrate β_s and β_u so that: (a) The capital-to-output ratio in the model equals 2. This number is calculated using the NIPA and Fixed Asset Tables as the average over the period 1967 – 2010. Krusell et al. (2000) exclude housing from both capital stock and output time series when they estimate the parameters of the production function. Since we use their estimates, we also exclude housing from both capital stock and output when we calculate the capital-to-output ratio.[6] (b) The asset holdings of an average skilled agent are 2.68 times those of an average unskilled agent (as in the 2010 U.S. Census). Finally, following Heathcote et al. (2017), we choose λ to clear the government budget constraint in equilibrium. Table 2 summarizes our calibration procedure.

5.1 Cobb-Douglas Economy

In our second economy, we eliminate capital-skill complementarity, and use the following production function:

$$Y = F(K, L_s, L_u) = AK^{\theta}(\mu L_s + L_u)^{1-\theta}.$$

The parameter θ governs the income share of capital; we assume $\theta = 1/3$ as common in the literature, which is also in line with the labor share target of the complementarity economy. The parameter μ equals to skill premium and allows us to capture the wage inequality between skilled and unskilled types with the Cobb-Douglas specification. We set μ=1.9 so that under the status quo policies two economies are the same in terms of wage inequality between skill groups. The parameter A stands for total factor productivity, which is internally calibrated to synchronize the two economies as explained below. We depreciation rate of capital, δ, is assumed to equal the weighted average of depreciation rates of structure capital and that of equipment capital. These exogenously calibrated technology parameters for the Cobb-Douglas economy is summarized in Table 1. The rest of the externally calibrated parameters in the Cobb-Douglas economy are chosen identically to the complementarity economy and are also summarized in the same table.

[6] Output is defined using Table 1.5.5 in NIPA as GDP (line 1) net of Housing and utilities (line 16) and Residential investment (line 41). Capital stock is calculated using the Fixed Asset Tables (FAT), Table 1.1 as the sum of the stocks of private and government structure and equipment capital (line 5 + line 6 + line 11 + line 12). The resulting annual capital-output ratio varies between 1.8 and 2.4 during the period of 1967-2010. To abstract from short-term fluctuations, the capital-output ratio value of 2, which we use as a calibration target, is computed by taking the average of annual capital-output ratios over this period.

The internal calibration procedure in the Cobb-Douglas economy is identical to that in the complementarity economy except that there are only five parameter values left to be determined. The first parameter is the total factor productivity parameter, A, which is calibrated so that the Cobb-Douglas economy has the same total output as the complementarity economy in the status-quo steady state. The calibrated value for A is reported in Table 2. The remaining four parameters are the labor disutility parameter ϕ, the discount factors β_s and β_u and the parameter governing the overall level of taxes in the tax function, λ. We calibrate these parameters to match the exact same targets as in the complementarity economy. The calibrated parameter values for these four are identical to those in the complementarity economy, which are given in the last four rows of Table 2.

6 Results

6.1 Optimal Policies vs. Benchmark Policies

In this section, we compute optimal capital and labor taxes for the economies calibrated in Section 5. Table 3 displays our findings. The first and third columns of the table summarizes the current U.S. economy under status-quo tax system with Cobb-Douglas and KORV production functions, respectively. The first two rows of the second and the fourth columns display optimal capital and average labor taxes under the assumptions of Cobb-Douglas and KORV aggregate production functions. The main finding of this section is that the optimal capital tax rate in the capital-skill complementarity economy is significantly larger than that in the Cobb-Douglas economy, 60% vs. 47%, which allows for a lower average labor tax rate.

In the standard Cobb-Douglas economy, increasing capital tax rate has the benefit of decreasing consumption inequality since it increases the tax burden on asset and labor income rich citizens and decreases the burden on low asset workers since it allows for lowering average labor taxes. However, taxing capital also entails the usual cost of discouraging capital accumulation and hence decreasing output. The fact that the optimal capital tax rate can be positive and large, 47% in our calculation, comes out of this trade off. Similar results have been shown previously by Domeij and Heathcote (2004) and other in the literature.

What is perhaps more interesting is the finding that under capital-skill complementarity, capital tax rate should be set significantly higher. The reason for this difference is that, in the capital-skill complementarity economy, in addition to the trade off explained above, increasing capital taxes has an additional redistributive benefit. Higher capital taxes slow down aggregate capital accumulation, and in particular accumulation of equipment capital, which diminishes skill premium. This way increasing capital taxes provide an indirect redistribution from skilled to unskilled agents. To the extent that unskilled agents have lower assets and wages, they have higher marginal utility from consumption, and hence, this redistribution increases the Utilitarian social welfare function. The striking nature of our finding is that the optimal capital tax difference is as high as 13 percentage points across the two economies that are fairly synchronized and calibrated in a standard way. The welfare gains from an optimal tax reform are 0.78% in consumption equivalent units in the economy with capital-skill complementarity while it is 0.18% in the standard Cobb-Douglas economy.

Sensitivity with Respect to Elasticity of Labor Supply. In our benchmark exercise, we take the parameter that controls the Frisch elasticity of labor supply to be $\gamma = 1$. This implies an elasticity of 1. As a sensitivity check, we calculate optimal capital tax rate when $\gamma = 2$, in other words, when Frisch elasticity is equals 0.5. The optimal capital tax rate equals 64% in the economy with capital-skill complementarity while it is 51% in the Cobb-Douglas economy. The fact that optimal tax differential is about the same when $\gamma = 1$ and $\gamma = 2$ suggests that the magnitude of the additional tax on capital coming from capital-skill complementarity is not affected by Frisch elasticity, at least around the region of plausible elasticities.

Rawlsian Social Welfare. In our baseline analysis, we assume that the government evaluates the consequences of the refom by aggregating citizens' welfare using a Utilitarian social

Table 3: Optimal Policies

	Calibrated CD	**Optimal CD**	**Calibrated CSC**	**Optimal CSC**
τ_k	0.36	0.47	0.36	0.60
λ	0.8142	0.8401	0.8142	0.8760
K/Y	2	1.8118	2	1.5770
w_s/w_u	1.90	1.90	1.90	1.6716
a_s/a_u	2.78	2.6424	2.78	2.3260
c_s/c_u	1.5499	1.5302	1.5499	1.4058

welfare function that puts an equal weight on all agents in the initial steady state. In this section, we consider another, significantly more redistributive, social welfare function, which follows the Rawlsian social welfare criterion. This social welfare function maximizes the welfare of the least forunate member of the society. The optimal tax problem then is to find the tax rate τ on capital income that leads to the competitive equilibrium that achieves the highest social welfare for the agent with the lowest social welfare. Formally, government solves the following problem:

$$\max_{\tau} \min_{v_0, i} V_i^0(v_0; \tau) \tag{4}$$

such that, for every τ, $V_i^0(v_0; \tau)$ is the value implied by the corresponding competitive equilibrium.

We find that the optimal capital tax rate is 85% in the economy with capital-skill complementarity while it is 64% in the Cobb-Douglas economy. Since redistributive considerations are much larger under the Rawlsian social welfare criterion and the least fortunate agent is an unskilled one, the government uses capital taxes heavily both to tax the asset-rich agents and to reduce the skill premium to as low as 1.1274.

7 Conclusion

This paper shows that capital-skill complementarity can generate a quantitative significant new rationale for taxing capital for redistributive governments. We find that it is optimal to rely much more on capital income tax and less on labor income tax in an economy where production process takes into account capital-skill complementarity relative to an economy with simple Cobb-Douglas production function. The welfare gains of an optimal tax reform are also significantly larger in the economy with capital-skill complementarity. Our analysis suggests that the debate over the correct tax rate on capital income should take into account the presence of capital-skill complementarities in production.

References

Aiyagari, S. R. (1994, August). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics 109*(3), 659–84.

Aiyagari, S. R. (1995, December). Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. *Journal of Political Economy 103*(6), 1158–75.

Attanasio, O. P., J. Banks, C. Meghir, and G. Weber (1999, January). Humps and bumps in lifetime consumption. *Journal of Business & Economic Statistics 17*(1), 22–35.

Bakis, O., B. Kaymak, and M. Poschke (2015, July). Transitional Dynamics and the Optimal Progressivity of Income Redistribution. *Review of Economic Dynamics 18*(3), 679–693.

Chamley, C. (1986, May). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica 54*(3), 607–622.

Conesa, J. C., S. Kitao, and D. Krueger (2009, March). Taxing capital? not a bad idea after all! *American Economic Review 99*(1), 25–48.

Domeij, D. and J. Heathcote (2004, 05). On the distributional effects of reducing capital taxes. *International Economic Review 45*(2), 523–554.

Duffy, J., C. Papageorgiou, and F. Perez-Sebastian (2004, 02). Capital-skill complementarity? evidence from a panel of countries. *Review of Economics and Statistics 86*(1), 327–344.

Fallon, P. R. and P. R. G. Layard (1975, 04). Capital-skill complementarity, income distribution, and output accounting. *Journal of Political Economy 83*(2), 279–302.

Flug, K. and Z. Hercowitz (2000, 07). Equipment investment and the relative demand for skilled labor: International evidence. *Review of Economic Dynamics 3*, 461–485.

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

Griliches, Z. (1969, 11). Capital-skill complementarity. *Review of Economics and Statistics 51*(4), 465–468.

Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal tax progressivity: An analytical framework. *forthcoming, Quarterly Journal of Economics*.

Imrohoroglu, S. (1998, May). A Quantitative Analysis of Capital Income Taxation. *International Economic Review 39*(2), 307–28.

Jones, L., R. Manuelli, and P. Rossi (1997). On the optimal taxation of capital income. *Journal of Economic Theory 73*(1), 93–117.

Judd, K. L. (1985, October). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics 28*(1), 59–83.

Kopecky, K. and R. Suen (2010, July). Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics 13*(3), 701–714.

Krueger, D. and A. Ludwig (2015). On the optimal provision of social insurance. Working paper.

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

Slavík, C. and H. Yazici (2014). Machines, buildings, and optimal dynamic taxes. *Journal of Monetary Economics 66*, 47 – 61.

Slavík, C. and H. Yazici (2019). On the consequences of eliminating capital tax differentials. *Canadian Journal of Economics 52*(1), 225 – 252.

Trabandt, M. and H. Uhlig (2011, 7). The laffer curve revisited. *Journal of Monetary Economics 58*(4), 305–327.

Appendix

A Equilibrium Definitions

A.1 Definition of Competitive Equilibrium for Cobb-Douglas Economy

Definition: Given a couple of initial distributions, $\{\lambda_i^0(v_0)\}_{i=u,s}$, a *competitive equilibrium* consists of an allocation $\left(\left\{c_{i,t}(v_0, z^t), l_{i,t}(v_0, z^t),\ a_{i,t+1}(v_0, z^t)\right\}_{i\in\{u,s\}, v_0\in\mathcal{V}_i, z^t\in Z_i^t}, K,\ L_{s,t},\ L_{u,t}\right)_{t=0}^{\infty}$, a policy $\left(T_t(\cdot), \tau_t,\ D_t, G_t\right)_{t=0}^{\infty}$, and a price system $(r_t, w_{s,t}, w_{u,t}, R_t)_{t=0}^{\infty}$ such that:

1. Given the policy and the price system, for each $i \in \{u, s\}$ and $v_0 \in \mathcal{V}_0$, the allocation $\left(\left\{c_{i,t}(v_0, z^t), l_{i,t}(v_0, z^t),\ a_{i,t+1}(v_0, z^t)\right\}_{z^t\in Z_i^t}\right)_{t=0}^{\infty}$ solves consumer's problem, i.e.,

$$V_i^0(v_0) = \max_{\left(\left\{c_{i,t}(z^t), l_{i,t}(z^t), a_{i,t+1}(z^t)\right\}_{z^t\in Z_i^t}\right)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{z^t\in Z_i^t} P_{i,t}(z^t|z_0)\beta_i^t u(c_{i,t}(z^t)) - v(l_{i,t}(z^t)) \qquad s.t.$$

$$\forall t \geq 0, z^t,$$

$$c_{i,t}(z^t) + a_{i,t+1}(z^t) \leq l_{i,t}(z^t)w_{i,t}z_t - T_t(l_{i,t}(z^t)w_{i,t}z_t) + R_t a_{i,t}(z^{t-1}),$$

where z^{-1} is the null history and,

$$\forall t \geq 0, z^t, \qquad c_{i,t}(z^t) \geq 0, a_{i,t+1}(z^t) \in \mathcal{A}, l_{i,t}(z^t) \geq 0.$$

2. In each period $t \geq 0$, taking factor prices as given, $(K_t, L_{s,t}, L_{u,t})$ solves the following firm's problem:

$$\max_{K_t, L_{s,t}, L_{u,t}} \qquad F(K_t, L_{s,t}, L_{u,t}) - r_t K_t - w_{s,t}L_{s,t} - w_{u,t}L_{u,t}.$$

3. Markets for assets, labor and goods clear: for all $t \geq 0$,

$$K_t + D_t = \sum_{i=u,s} \pi_i \int_{\mathcal{V}_i} \sum_{z^{t-1} \in Z_i^{t-1}} P_{i,t-1}(z^{t-1}|z_0) a_{i,t}(v_0, z^{t-1}) d\lambda_i^0(v_0),$$

$$L_{i,t} = \pi_i \int_{\mathcal{V}_i} \sum_{z^t \in Z_i^t} P_{i,t}(z^t|z_0) l_{i,t}(v_0, z^t) z_t d\lambda_i^0(v_0), \text{ for } i = u, s,$$

$$G_t + C_t + K_{t+1} = F(K_t, L_{s,t}, L_{u,t}) + (1-\delta)K_t,$$

where

$$C_t = \sum_{i=u,s} \pi_i \int_{\mathcal{V}_i} \sum_{z^t \in Z_i^t} P_{i,t}(z^t|z_0) c_{i,t}(v_0, z^t) d\lambda_i^0(v_0)$$

is aggregate consumption in period t.

4. The government's budget constraint is satisfied every period: for all $t \geq 0$,

$$G_t + R_t D_t = D_{t+1} + \tau_t (r_t - \delta) K_t + T_{agg},$$

where

$$T_{agg} = \sum_{i=u,s} \pi_i \int_{\mathcal{V}_i} \sum_{z^t \in Z_i^t} P_{i,t}(z^t|z_0) T_t(l_{i,t}(v_0, z^t) w_{i,t} z_t) d\lambda_i^0(v_0)$$

denotes aggregate labor income tax revenue.

A.2 Definition of Stationary Recursive Competitive Equilibrium for Capital-Skill Complementarity Economy

In order to define a stationary equilibrium, we assume that policies (government expenditure, debt and taxes) do not change over time.

Stationary Recursive Competitive Equilibrium (SRCE). SRCE is two value functions $\{V_u, V_s\}$, policy functions $\{c_u, c_s, l_u, l_s, a'_u, a'_s\}$, the firm's decision rules $\{K_s, K_e, L_s, L_u\}$, government policies $\{\tau_k, T(\cdot), D, G\}$, two distributions over productivity-asset types $\{\lambda_u(z, a),$

$\lambda_s(z,a)\}$ and prices $\{w_u, w_s, r_s, r_e, R\}$ such that

1. The value functions and the policy functions solve the consumer problem given prices and government policies, i.e., for all $i \in \{u,s\}$:

$$\begin{aligned} V_i(z,a) = \max_{(c_i,l_i,a_i')\geq 0} \quad & u(c_i) - v(l_i) + \beta_i \sum_{z'} \Pi_i(z'|z) V_i(z', a_i') \qquad \text{s.t.} \\ c_i + a_i' \quad \leq \quad & w_i z l_i - T(w_i z l_i) + Ra, \end{aligned}$$

   where $R = 1 + (r_s - \delta_s)(1-\tau_k) = 1 + (r_e - \delta_e)(1-\tau_k)$ is the after-tax asset return.

2. The firm solves the profit maximization problem each period.

3. The distribution λ_i is stationary for each type, i.e. $\forall i : \lambda_i'(z,a) = \lambda_i(z,a)$. This means:

$$\lambda_i(\bar{z},\bar{a}) \quad = \quad \sum_{z \in Z_i} \Pi_i(\bar{z}|z) \int_{a:a_i'(z,a)\leq \bar{a}} d\lambda_i(z,a), \qquad \forall (\bar{z},\bar{a}).$$

4. Markets clear:

$$\begin{aligned} \sum_i \pi_i \int_z \int_a a \cdot d\lambda_i(z,a) &= K_s + K_e + D, \\ \pi_s \int_z \int_a z l_s(z,a) \cdot d\lambda_s(z,a) &= L_s, \\ \pi_u \int_z \int_a z l_u(z,a) \cdot d\lambda_u(z,a) &= L_u, \\ C + G + K_s + K_e &= F(K_s, K_e, L_s, L_u) + (1-\delta_s)K_s + (1-\delta_e)K_e, \end{aligned}$$

   where $C = \sum_{i=u,s} \pi_i \int_z \int_a c_i(z,a) \cdot d\lambda_i(z,a)$ denotes aggregate consumption.

5. Government budget constraint is satisfied.

$$RD + G = D + \tau_k(r_e - \delta_e)K_e + \tau_k(r_s - \delta_s)K_s + T_{agg},$$

   $T_{agg} = \sum_{i=u,s} \pi_i \int_z \int_a T(w_i z l_i(z,a)) \cdot d\lambda_i(z,a)$ denotes aggregate labor tax revenue.