

Redistributive Capital Taxation Revisited

Özlem Kına, Ctirad Slavík, Hakki Yazici

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Redistributive Capital Taxation Revisited

Abstract

This paper shows that capital-skill complementarity provides a quantitatively significant rationale to tax capital for redistributive governments. The optimal capital income tax rate is 60%, which is significantly higher than the optimal rate of 48% in an identically calibrated model without capital-skill complementarity. The skill premium falls from 1.9 to 1.67 along the transition following the optimal reform in the capital-skill complementarity model, implying substantial indirect redistribution from skilled to unskilled workers. These results show that a government that cares about redistribution should take into account capital-skill complementarity in production when setting the tax rate on capital income.

JEL-Codes: E250, J310.

Keywords: capital taxation, capital-skill complementarity, inequality, redistribution.

Özlem Kina
European University Institute
Italy – 50014 San Domenico di Fiesole
ozlem.kina@eui.eu

Čtírad Slavík
CERGE-EI
Czech Republic - 111 21 Prague
ctirad.slavik@cerge-ei.cz

Hakki Yazici
University of Bristol
United Kingdom – Bristol BS8 1TU
hakki.yazici@bristol.ac.uk

October 7, 2020

We would like to thank Árpád Ábrahám, Ivo Bakota, Marek Kapička, Dirk Krueger, Remzi Kaygusuz, Baris Kaymak, Alex Ludwig, Ramon Marimon, Alex Monge-Naranjo, Gonul Sengul, Faisal Sohail, Dušan Stojanović, Hitoshi Tsujiyama, Gianluca Violante, and participants of CEF 2018, CEF 2019, EEA 2018, ASSA 2019, 2020 World Congress of the Econometric Society for helpful suggestions and comments. We would also like to thank Andrea Downing for excellent English editing. Slavík acknowledges support from the Czech Science Foundation (grant project 17-27676S).

1 Introduction

The optimal tax rate on capital income has long been debated. The supporters of capital tax cuts stress the efficiency costs associated with capital taxation: the 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, show that redistributive benefits of capital taxation can be large enough to imply significant optimal tax rates on capital income. In this paper, we contribute to the debate on optimal capital taxation by putting forward 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 proposed mechanism implies that the optimal tax rate on capital income should be considerably higher than the conventional economic models tell us.

At the heart of this 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 relative demand for skilled workers due to capital-skill complementarity. As a result, the skill premium - the wages of the skilled workers relative to those of the unskilled workers - declines. Since skilled workers normally earn higher wages and have more assets, this decline in the skill premium increases social welfare from the perspective of a government that values equality.

We measure the quantitative significance of this mechanism for the optimal capital tax rate using a model that embeds capital-skill complementarity into an incomplete markets

¹Capital-skill complementarity was first empirically documented by Griliches (1969). It has received much attention from economists and has been successfully used in explaining the evolution of inequality in the returns to education. Among others, see Fallon and Layard (1975), Krusell, Ohanian, Ríos-Rull, and Violante (2000), Flug and Hercowitz (2000), and Duffy, Papageorgiou, and Perez-Sebastian (2004).

model à la Bewley (1986), Imrohoroglu (1989), Huggett (1993) and Aiyagari (1994), where individuals face idiosyncratic wage risk. We choose this model as it allows for sufficiently rich modelling of earnings and wealth inequality, which is key to accurately assessing the redistributive benefits of capital taxation. We consider two versions of the model that differ from each other only 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, Ohanian, Ríos-Rull, and Violante (2000). As a benchmark for comparison, we also build a second economy with a standard Cobb-Douglas production function that does not feature capital-skill complementarity. We make the two model economies comparable by calibrating each one separately to the current U.S. economy along selected dimensions under the status-quo capital and labor tax system.

We consider the problem of a government which chooses a linear tax rate on capital income to maximize a Utilitarian social welfare function with equal weights on all agents. The government takes into account the effect of tax changes on people's welfare over the transition to the new steady state. 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. 48%. Accordingly, the average labor income tax is smaller in the economy with capital-skill complementarity. In response to the optimal tax reform, the skill premium falls over transition from the calibrated value of 1.9 in the initial steady state to 1.67 in the final steady state in the capital-skill complementarity economy, while it remains virtually unchanged in the Cobb-Douglas economy. Since labor income taxes are distortionary, this indirect redistribution channel is valuable to the government and gives rise to the higher optimal capital tax rate in the economy with capital-skill complementarity. This finding shows that the debate over the correct tax rate on capital income should take into account the presence of capital-skill complementarities in production.

Under the Utilitarian social welfare function, the welfare gains of the reform are equivalent to those from increasing consumption of all agents by 0.78% at every date and state in the economy with capital-skill complementarity. The corresponding welfare gains amount to 0.23% in the Cobb-Douglas economy. This difference in welfare gains implies that carrying out the optimal capital tax reform is considerably more important once capital-skill complementarity is taken into account.

Through an extensive sensitivity analysis, we show that our results are quantitatively robust to alternative degrees of capital-skill complementarity estimated using different data sets, to a lower level of elasticity of labor supply, to the absence of uninsured individual wage risk, to modeling top wealth inequality and to alternative specifications of the social welfare function.

In the main body of the paper, we focus on the effect of capital-skill complementarity on the optimal capital tax rate in the context of a tax reform where the government is allowed to only choose the capital tax rate. As an extension, in Section 6, we analyze the effect of capital-skill complementarity on the optimal capital tax rate in the context of a comprehensive tax reform where, in addition to the capital tax rate, the government also chooses the degree of progressivity of the labor income tax code. We find that the optimal capital tax rate is still 12 percentage points higher in the model with capital-skill complementarity relative to the Cobb-Douglas model.

Related Literature. Taxation of capital income is a controversial topic in the macroeconomics literature. In the representative-agent paradigm, Chamley (1986) and Judd (1985) show that it is optimal not to tax capital at all in the long run.² Aiyagari (1995) shows that the optimal long-run capital income tax might be positive when we model heterogeneity across agents, arising from uninsured labor income risk and incomplete markets. He points

²Straub and Werning (2020) provide a set of conditions under which the optimality of zero taxes on capital in the long run does not hold. Chari, Nicolini, and Teles (2019) show that with a richer set of tax instruments and under the assumption that initial confiscation of wealth is restricted, one recovers the long-run optimality of zero capital taxes.

out that the optimal steady state capital income tax is between 25% and 45% depending on the values of various model parameters.³ Domeij and Heathcote (2004) investigate the quantitative importance of heterogeneity and idiosyncratic labor income risk for capital taxation using an Aiyagari (1994) model. They consider the problem of a redistributive government which needs to choose constant (time-independent) tax rates on capital and labor income. They find that eliminating capital income taxes altogether brings large welfare gains if they assume a representative-agent economy. However, when there is heterogeneity and risk, the optimal capital tax rate can be quite high, namely 40% according to their calculations. Imrohoroglu (1998) and Conesa, Kitao, and Krueger (2009) also analyze optimal capital taxation in quantitative models with rich heterogeneity, and in particular, a life cycle structure. Conesa, Kitao, and Krueger (2009) find that, due to the life-cycle structure, optimal capital taxes can be significantly positive at 36% even when the government maximizes steady-state welfare.⁴ We add to this literature by assessing the quantitative impact of capital-skill complementarity on optimal capital taxation.

There is also a more recent and growing literature on taxation of capital in the presence of capital-skill complementarity. Jones, Manuelli, and Rossi (1997) provide an important backdrop to this literature. In an extension section, the authors analyze optimal linear taxation in a growth model with two types of labor, skilled and unskilled, and show that the optimal long-run capital tax rate may be positive if the labor income tax rate is not allowed to depend on skill type and there is capital-skill complementarity. The key difference of the current paper from Jones, Manuelli, and Rossi (1997) is that we evaluate the effect of capital-skill complementarity for optimal capital tax rate quantitatively in a model that allows for a rich modelling of earnings and wealth inequality, whereas they use a simple, representative agent

³These numerical results are not included in the published version of the paper, and are only available in a working paper version. This version is available as Minneapolis Fed Working Paper Series #508. Moreover, Aiyagari (1995) only reports optimal taxes at the steady state. Recently, Acikgoz, Hagedorn, Holter, and Wang (2018) and Dyrda and Pedroni (2020) calculate time-varying paths of optimal capital and labor taxes in environments with uninsurable wage risk.

⁴See also the New Dynamic Public Finance literature, which has followed the seminal contribution of Golosov, Kocherlakota, and Tsyvinski (2003), for investigations of optimal capital taxation in dynamic Mirrlesian private information models with idiosyncratic labor income shocks.

framework to make a qualitative statement. Slavík and Yazici (2014) also build a model with capital-skill complementarity, but they use it to study the optimality of differential capital taxation. They find that in their private information Mirrleesian model it is optimal to tax equipment at a higher rate than structures.⁵ Bhattarai, Lee, Park, and Yang (2020) analyze macroeconomic effects of specific capital tax reforms under capital-skill complementarity. Angelopoulos, Asimakopoulos, and Malley (2015) use a representative agent model to evaluate the optimality of labor tax smoothing under capital-skill complementarity, while Dolado, Motyovszki, and Pappa (2020) analyze monetary policy and its redistributive implications in a New Keynesian model with capital-skill complementarity.⁶

The rest of the paper is organized as follows. Section 2 lays out the model while Section 3 describes the optimal tax problem formally. Section 4 explains the calibration strategy and Section 5 discusses the main quantitative results. Section 6 explores the optimal comprehensive tax reform. Finally, Section 7 concludes.

2 Model

The economy consists of a unit measure of individuals, a firm, and a government all of whom live forever. In the baseline model, the aggregate production function 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. Each period people are endowed with one unit of time. People are permanently different with respect to their skill levels: they are either skilled or unskilled, $i \in \{u, s\}$. Skilled agents can only work in the skilled labor sector and unskilled

⁵There is a growing literature which analyzes the optimal taxation of robots, see e.g. Guerreiro, Rebelo, and Teles (2017), Costinot and Werning (2018) and Thuemmel (2020).

⁶Krueger and Ludwig (2016) and Heathcote, Storesletten, and Violante (2017) are also related to the current paper in the sense that they analyze optimal taxation in models with imperfect substitutability between skilled and unskilled labor in which there are general equilibrium effects of taxation on skill prices. Importantly, neither of these papers models capital-skill complementarity.

agents only in the unskilled labor sector. The total mass of type i workers is denoted by π_i . In the quantitative analysis, skill types correspond to educational attainment at the time of entering the labor market. Workers who have at least a bachelor degree are classified as skilled agents and the rest of the agents are classified as unskilled.

There is also ex-post heterogeneity within each skill group arising from workers facing idiosyncratic labor productivity shocks over time. The productivity shock is denoted by z and follows a type-specific Markov chain with states $Z_i = \{z_{i,1}, \dots, z_{i,I}\}$ and transitions $\Pi_i(z'|z)$. When a skill type i worker draws productivity level z and works l units in a period, she produces $l \cdot z$ units of effective type i labor. 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 utility function which is separable between consumption and labor and over time

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

where, for each worker type i , the expectation, E_i , is taken over productivity shocks and β_i is the time discount factor.⁷

Technology. The production process is summarized by a constant returns to scale production function: $Y = F(K_s, K_e, L_s, L_u)$, where K_s , K_e , L_s and L_u refer to the aggregate levels of structure capital, equipment capital, effective skilled labor supply and effective unskilled labor supply, respectively. The stocks of structure and equipment capital depreciate at rates δ_s and δ_e , respectively.

⁷In the quantitative analysis, we calibrate the discount factors so as to match the observed difference in wealth between skilled and unskilled agents. The calibration implies a slightly higher discount factor for skilled workers, which is in line with the empirical evidence provided by Attanasio, Banks, Meghir, and Weber (1999), who estimate discount factors for different education groups. Importantly, as shown in Section 5.1, the quantitative importance of capital-skill complementarity for the optimal capital tax rate does not depend on the assumption of heterogeneous discount factors.

We assume that there is capital-skill complementarity in the production process. More specifically, technology features 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 that of skilled labor. Under the assumption of 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 consistent with the estimation results of Krusell, Ohanian, Ríos-Rull, and Violante (2000).

Production is carried out by a representative firm, which, in each period, rents the two types of capital and hires the two types of labor to maximize profits. The firm solves the following maximization problem in period t :

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

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 .

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, Storesletten, and Violante (2017) and assume that tax liability given labor income y is defined as:

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

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. When $\tau_l > 0$, labor taxes are progressive and the tax function implies transfers to people with sufficiently low income. The government uses taxes to finance a stream of expenditure $\{G_t\}_{t=0}^{\infty}$ and repay government debt $\{D_t\}_{t=0}^{\infty}$.

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.

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 $z^t \equiv (z_1, \dots, z_t)$. 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) . In the definition that follows, it is understood 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 $\left(\left\{c_{i,t}(v_0, z^t), l_{i,t}(v_0, z^t), a_{i,t+1}(v_0, z^t)\right\}_{i \in \{u,s\}, v_0 \in \mathcal{V}_i, z^t \in Z_i^t}, K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}\right)_{t=0}^\infty$, 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 the consumer's problem:⁸

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

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,

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

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:

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

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

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

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

⁸In the consumer's utility maximization problem below, the discounted infinite sum of period utilities is scaled by $(1 - \beta_i)$ to ensure that the infinite sum of weights is equal across all agents.

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

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 $t \geq 0$,

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

where

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

denotes aggregate labor income tax revenue in period t .

2.1 Cobb-Douglas Economy

To assess the quantitative significance of capital-skill complementarity for optimal capital taxes, we consider a second, benchmark, economy in which the production function does not feature capital-skill complementarity. In this economy, we do not distinguish between equipment capital and structure capital; there is only one type of capital, which depreciates every period at rate δ . First, the skilled and unskilled labor inputs are combined to aggregate labor N . The details of how the two types of labor are combined will be discussed in Section 4. Next, capital and labor are combined to produce aggregate output using a standard Cobb-Douglas production function $Y = AK^\theta N^{1-\theta}$. We preserve all the other properties of the first model.

Importantly, under this production function, the ratio of the marginal product of skilled labor to that of unskilled labor, hence the skill premium, is independent of the amount of capital in the economy. The changes in the aggregate capital level do not affect the skill

premium, therefore, capital income taxation has no direct impact on wage inequality. The definition of competitive equilibrium for this economy is very similar to that given for the capital-skill complementarity economy, and hence is relegated to Appendix A.1.

3 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 the productivity-asset space implied by this steady state, the government introduces a once 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 spending and bond holding levels given by the initial steady state, the government adjusts the parameter that controls the average labor income tax, $\{\lambda_t\}_{t=0}^{\infty}$, along the transition to the new steady state.

In the baseline analysis, we assume that the government does not change the progressivity of the labor tax function. We do so because, perhaps due to political constraints, it is difficult for governments to carry out comprehensive reforms in which capital and labor tax codes are changed substantially at the same time. In Section 6, we analyze the effect of capital-skill complementarity on optimal capital taxes in the presence of such a comprehensive reform. The government evaluates the consequences of the reform by aggregating agents' welfare using a Utilitarian social welfare function that puts an equal weight on all agents in the initial steady state. Importantly, the government takes into account the effect of the tax reform on people's welfare over the transition. 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, the government solves the following problem:

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

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

4 Calibration

This section first explains how we calibrate the baseline model with capital-skill complementarity to the U.S. economy. 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 stationary recursive competitive equilibrium of the model economy matches the U.S. economy around the mid 2010s along selected dimensions that are key for our investigation.⁹ 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. The details of our data work are included in Appendix B.

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

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

where σ equals the coefficient of relative risk aversion while γ controls the 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 2017. We focus on males who are 25 years old or older and who have earnings. To be consistent with Krusell, Ohanian, Ríos-Rull, and Violante (2000), skilled people are defined as those who have at least a bachelor's degree.

⁹The definition of stationary recursive competitive equilibrium is relegated to Appendix A.2 for brevity. We choose the mid 2010s U.S. economy as the calibration target because we want to focus on the U.S. economy before the capital tax reform of President Trump's administration, which entered into effect on January 1, 2018, as the full effects of that reform may not have taken place yet.

Technology. In the baseline economy with capital-skill complementarity, the production function takes the same form as in Krusell, Ohanian, Ríos-Rull, and Violante (2000):

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

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, Ohanian, Ríos-Rull, and Violante (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. As reported in the National Income and Product Accounts (NIPA), the government consumption-to-output ratio has been fairly stable with an average of about 16% since the 1980's. This is the value we use. We assume a government debt of 60% of GDP, which equals the federal debt held by private investors over GDP in 2015 according to the Federal Reserve Bank of Saint Louis FRED database.

We follow Trabandt and Uhlig (2011) and assume that the current tax rate on capital income is $\tau = 36\%$. 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 to unskilled workers, dwarfing the importance of taxing capital for indirect redistribution. To approximate the progressive U.S. labor tax code, we follow Heathcote, Storesletten, and Violante (2017). Using the PSID data for 2000 – 2006 and the TAXSIM program, they estimate $\tau_l = 0.18$. We use their estimate and

Table 1: Benchmark Parameters

Parameter	Symbol	Value	Source
<i>Technology (Capital-skill complementarity model)</i>			
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
<i>Technology (Cobb-Douglas model)</i>			
Capital's share of output	θ	1/3	
Measure of elasticity of substitution between skilled labor L_s and unskilled labor L_u	ε	0.2908	KM
Depreciation rate of capital	δ	0.0787	
<i>Common parameters</i>			
Relative risk aversion parameter	σ	2	
Inverse Frisch elasticity	γ	1	
Relative supply of skilled workers	π_s	0.3544	CPS
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
Labor tax progressivity	τ_l	0.18	HSV
Linear tax rate on capital income	τ	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, KM and TU stand for Greenwood, Hercowitz, and Krusell (1997), Krusell, Ohanian, Ríos-Rull, and Violante (2000), Heathcote, Storesletten, and Violante (2017), Krueger and Ludwig (2016), Katz and Murphy (1992) 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.

calibrate λ , which controls the average labor tax in the economy, to clear the government budget, following their procedure.¹⁰

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 workers. This assumption implies that w_i corresponds to the average wage rate of agents of skill type i . As a result, the 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 (2016), we set $\rho_s = 0.9690, \text{var}(\varepsilon_s) = 0.0100, \rho_u = 0.9280, \text{var}(\varepsilon_u) = 0.0192$. We approximate these processes by finite number Markov chains using the Rouwenhorst method described in Kopecky and Suen (2010).

Internal Calibration. There are still seven parameter values left to be determined: 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 we calculate using Fixed Asset Tables (FAT) data for mid 2010s, (ii) the labor share equals 2/3, and (iii) the skill premium equals 1.9 as reported by Heathcote, Perri, and Violante (2010).¹¹ We choose ϕ so that the aggregate labor supply in steady state equals 1/3 as commonly assumed in the macro literature. We calibrate β_s and β_u so that: (i) The capital-to-output ratio in the model

¹⁰Bakis, Kaymak, and Poschke (2015) use CPS data for the period 1979-2009 and find a similar value for the progressivity parameter $\tau_l = 0.17$.

¹¹Heathcote, Perri, and Violante (2010) use CPS data and compute the skill premium for the period 1967-2005 for males between ages of 25 and 60, working at least 260 hours a year. In subsequent work, they update skill premium data series until 2016. They find that the skill premium has been stable around 1.9 during 2005-2016 period.

Table 2: Benchmark Calibration Procedure

Parameter	Symbol	Value	Target	Source
<i>Technology (CSC)</i>				
Production parameter	ω	0.3332	Labor share = 2/3	NIPA
Production parameter	ν	0.6205	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K^e}{K} = 1/3$	FAT
<i>Technology (CD)</i>				
Total factor productivity	A	0.7830	Output level of CSC economy	
Production parameter	κ	0.5007	Skill premium = 1.9	CPS
<i>Common parameters</i>				
Skilled discount factor	β_s	0.9415	Capital to output ratio = 2	NIPA, FAT
Unskilled discount factor	β_u	0.9365	Relative skilled wealth = 2.78	US Census
Tax function parameter	λ	0.8142	Government budget balance	
Disutility of labor	ϕ	65.97	Labor supply = 1/3	

This table reports the internal calibration procedure. The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). The production function parameter κ controls the income shares of the skilled and unskilled labor in the Cobb-Douglas model (CD). 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.

equals 2. This number is calculated using the NIPA and Fixed Asset Tables as the average over the period 1967 – 2017. Krusell, Ohanian, Ríos-Rull, and Violante (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. (ii) The asset holdings of an average skilled agent are 2.78 times those of an average unskilled agent as we calculate using 2014 U.S. Census data. Finally, following Heathcote, Storesletten, and Violante (2017), we choose λ to clear the government budget constraint in equilibrium. Table 2 summarizes the internal calibration procedure.

4.1 Calibration of the Cobb-Douglas Economy

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

$$Y = AK^\theta(\kappa L_s^\varepsilon + (1 - \kappa)L_u^\varepsilon)^{\frac{1-\theta}{\varepsilon}}$$

where A is total factor productivity, θ is the usual Cobb-Douglas parameter that governs the income share of capital, κ is a share parameter that allows for skilled labor to be more effective than unskilled labor, and ε controls the degree of substitutability between skilled and unskilled labor. We set $\theta = 1/3$ as is common in the literature. This is also in line with the labor share target of the capital-skill complementarity economy. Following Katz and Murphy (1992), we set the elasticity of substitution between skilled and unskilled labor to 1.41, which implies $\varepsilon = 0.2908$. The depreciation rate of capital, δ , is assumed to equal the weighted average of depreciation rates of structure capital and that of equipment capital in the capital-skill complementarity economy. These exogenously calibrated technology parameters for the Cobb-Douglas economy are 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.

The internal calibration procedure in the Cobb-Douglas economy is identical to that in the capital-skill complementarity economy except that there are only six 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 capital-skill complementarity economy in the status-quo steady state. The calibrated value for A is reported in Table 2. The second parameter is κ , which is chosen to ensure that the skill premium equals 1.9. 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. As a result, the calibrated parameter values for these four are identical to those in the complementarity economy, and are given in the last four rows of Table 2.

It is worth emphasizing that the calibration procedures render the two economies completely identical. That is, the real interest rate, the skilled and unskilled wages, aggregate output, aggregate capital stock, aggregate labor and consumption, as well as the distribu-

tions of consumption, labor, assets, earnings and welfare across workers are identical in the initial steady states of the two economies. This synchronization of the capital-skill complementarity and Cobb-Douglas economies is important as it assures us that any differences in the optimal tax rates across the two economies cannot be coming from the differences in the initial conditions of the two economies.

5 Optimal Capital Taxation

This section reports the optimal capital tax rates for the two economies calibrated in Section 4. Table 3 displays the main findings. The first column of the table summarizes selected moments of the calibrated economy under the status-quo fiscal policy. The second and third columns report the steady states of the two economies under optimally chosen capital tax rates. In particular, the first two rows of these columns display optimal capital tax rate and the steady -state average labor taxes in the corresponding economies. The main finding 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. 48%. Accordingly, optimal average labor taxes in the steady state, represented by $1 - \lambda$, are relatively lower in the capital-skill complementarity economy.

In the standard Cobb-Douglas economy, increasing the tax rate on capital income has the benefit of decreasing consumption inequality since capital income is more unevenly distributed across the population than labor income. However, taxing capital also entails the usual cost of discouraging its accumulation, and hence, depressing output. That the optimal capital tax rate is positive and large, 48% in our calculation, arises from this trade off.¹² Similar large capital tax rates have been shown to be optimal previously in the literature, for instance by Domeij and Heathcote (2004) or Conesa, Kitao, and Krueger (2009).

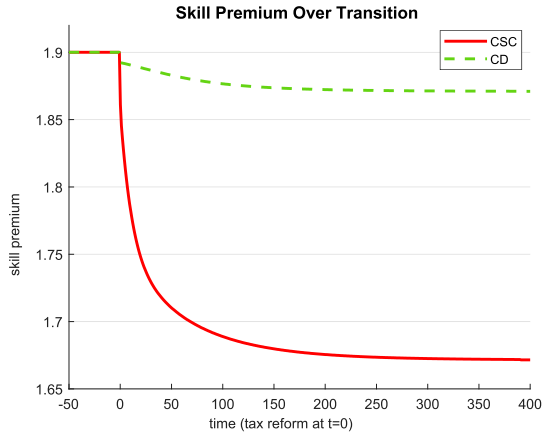
¹²Perhaps, a more standard version of the Cobb-Douglas model would have perfect substitutability between skilled and unskilled labor. In this case the optimal capital tax rate is 47%, very similar to what it is in the Cobb-Douglas benchmark model with imperfect substitutability.

Table 3: Optimal Capital Taxes

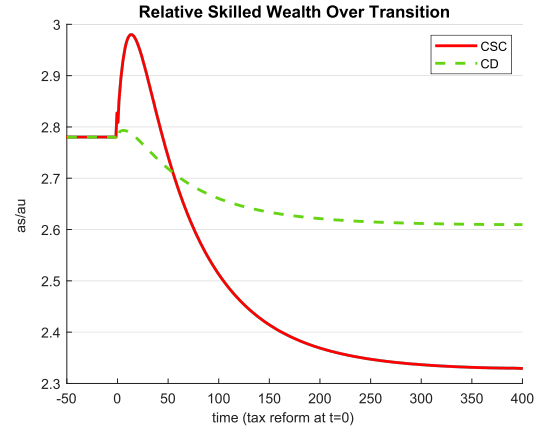
	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.48	0.60
Labor tax parameter, λ	0.81	0.84	0.88
Skill premium, w_s/w_u	1.90	1.87	1.67
Relative skilled wealth	2.78	2.61	2.33
Relative skilled consumption	1.55	1.52	1.41
Output, Y	100	96.5	92.8
Capital, K	100	86.5	73.2
Capital-output ratio, K/Y	2.00	1.79	1.58
Overall welfare gains	–	0.23%	0.78%
Skilled welfare gains	–	-0.70%	-3.42%
Unskilled welfare gains	–	0.58%	2.40%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system. The second and third columns report optimal capital taxes and the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models. Output and capital are normalized to 100 in the status-quo steady state.

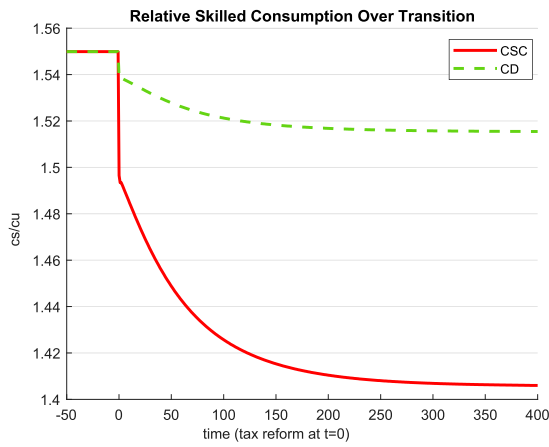
What is more interesting is the finding that with capital-skill complementarity, the capital tax rate should be set significantly, namely 12 percentage points, higher. The reason for this difference is that, in the capital-skill complementarity economy, in addition to the trade off explained above, increasing capital taxes has an additional redistributive benefit. Higher capital taxes slow down aggregate capital accumulation, and in particular the accumulation of equipment capital. When there is capital-skill complementarity, this decreases the relative demand for skilled labor, which then diminishes the skill premium. As a result, increasing capital taxes provides 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 social welfare from the perspective of a Utilitarian planner who cares equally about all agents. The striking nature of our finding is that this channel implies an optimal capital tax differential of as high as 12 percentage points.



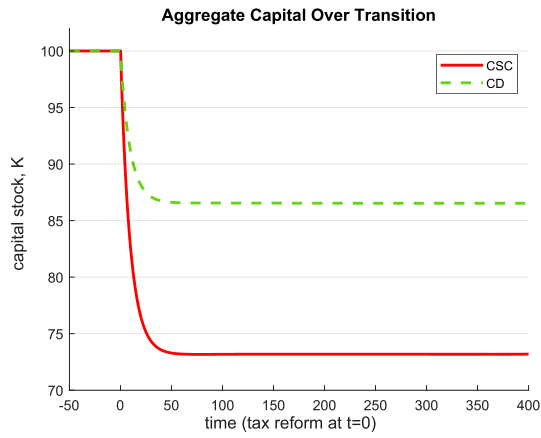
(a) Skill premium



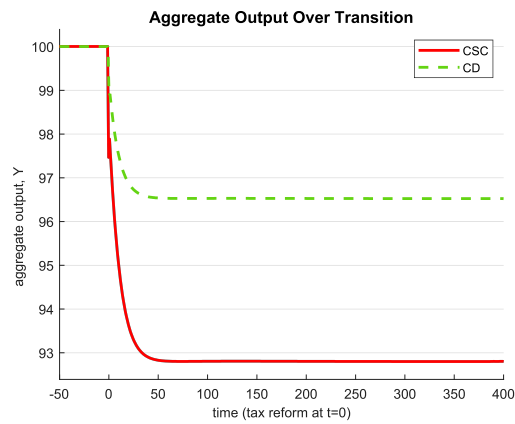
(b) Relative average skilled wealth



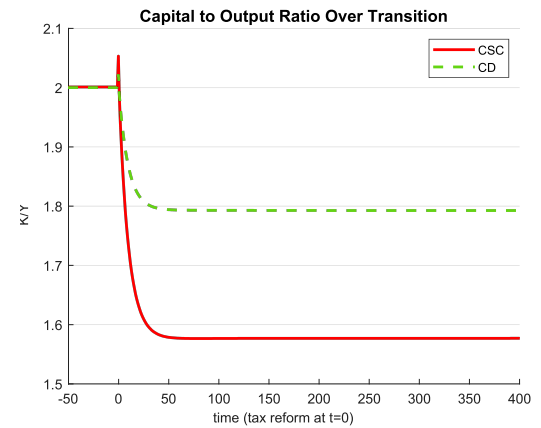
(c) Relative average skilled consumption



(d) Aggregate capital stock



(e) Aggregate output



(f) Capital-output ratio

Figure 1: Dynamics of key macroeconomic variables following the optimal reform

The six graphs report how the skill premium, relative skilled wealth, relative skilled consumption, aggregate capital stock, aggregate output, and the capital-output ratio change over the transition following the optimal tax reform. CSC and CD refer to capital-skill complementarity and Cobb-Douglas economies, respectively.

To get a better grasp of how taxing capital has an additional redistributive benefit under capital-skill complementarity, observe from Figure 1a that the reform from the status-quo tax rate of 36% to the optimal rate of 60% reduces the skill premium considerably, from 1.90 to 1.67 over the transition. This decline in wage inequality then reduces average wealth and consumption inequality between the two groups over the transition as can be seen from Figures 1b and 1c, respectively.¹³ Rising capital taxes have virtually no effect on the skill premium in the Cobb-Douglas case. Correspondingly, the decline in consumption and asset inequality in the Cobb-Douglas economy in response to increasing capital taxes is significantly less pronounced. It is this additional redistributive benefit of taxing capital that calls for higher capital taxes in the economy with capital-skill complementarity.

Of course, higher capital taxes come with higher efficiency costs in the capital-skill complementarity economy. As Figures 1d and 1e display, over the transition to the final steady state, both aggregate capital stock and output shrink more in the capital-skill complementarity economy relative to the Cobb-Douglas economy. As a result, as the capital-output ratio drawn in Figure 1f indicates, the production process in the capital-skill complementarity economy becomes much less capital intensive after the reform.

One interpretation of this finding is that a government that cares about equality, but does not have access to non-distortionary (type-dependent) transfers, finds it optimal to make the capital-skill complementarity economy substantially less capital intensive in order to take advantage of the indirect redistribution this generates. The decline in capital intensity leads to a decline in the skill premium to a level observed in the first half of the 1990s (see Heathcote, Perri, and Violante (2010)). In other words, the government finds it optimal to offset several decades of skill-biased technical change.

¹³Figure 1b shows that the relative skilled wealth increases for a short period of time following the rise in capital tax. This happens because, as a result of the reform, skilled wages decline until the new steady state, see Figure 2b. This decline in future wages induces skilled workers to save more, *ceteris paribus*, as workers have rational expectations. Unskilled wages decline over the transition too, but to a much lesser extent. Therefore, unskilled saving does not increase as much initially in response to a decline in future wages.

Welfare Gains. The welfare gains of the reform are equivalent to increasing the consumption of all agents at all dates and states by 0.78% in the economy with capital-skill complementarity, while the corresponding welfare gains number is 0.23% in the standard Cobb-Douglas economy. The first key takeaway of the welfare analysis is that carrying out the optimal capital tax reform is significantly more important in terms of its welfare effects when capital-skill complementarity in production is taken into account.

We also find that the distribution of welfare gains is more tilted toward the unskilled in the capital-skill complementarity model relative to the Cobb-Douglas model. The welfare of unskilled agents as a group increases by 2.40% in consumption equivalence units in the capital-skill complementarity economy. The skilled agents' welfare, on the other hand, decreases by 3.42%. The overall welfare increases since there are many more unskilled than skilled in the economy. Looking more closely, within both the skilled and the unskilled groups, it is the asset-poor agents who gain and the asset-rich agents who lose. In particular, while 78% of the unskilled gain, only 11% of the skilled do so. In the Cobb-Douglas economy, the unskilled welfare increases and the skilled welfare decreases to lesser extents, namely by 0.58% and 0.70%, respectively. Since the indirect redistribution channel is missing, the welfare implications are more symmetric across the two groups than in the capital-skill complementarity case: 69% of the unskilled and 45% of the skilled gain.

Transition vs. Steady State. It is instructive to discuss the timing of welfare gains. As Domeij and Heathcote (2004) demonstrate in a model similar to ours with a Cobb-Douglas production function, the optimal tax rate on capital is much lower, possibly negative, when the government maximizes steady-state welfare (see also Dyrda and Pedroni (2020) for a similar finding). This is also true in our framework, in both the Cobb-Douglas and the capital-skill complementarity models. This happens because low capital tax rates induce higher capital accumulation, which allows for higher consumption with lower labor effort in the long run. Building up the capital stock is costly in the short run, however, and the

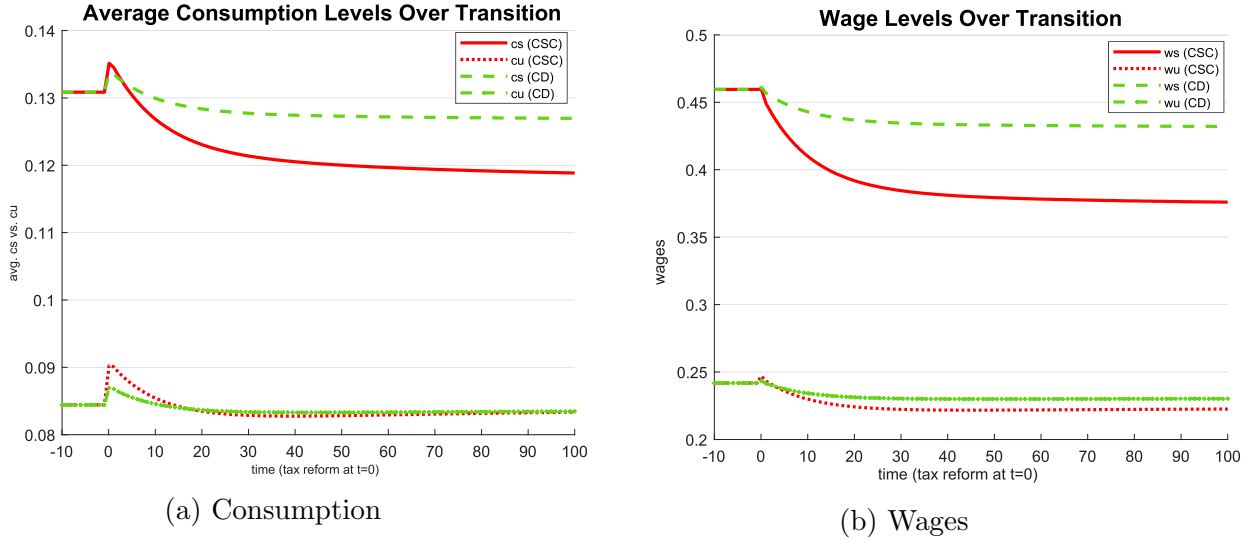


Figure 2: Dynamics of consumption and wages following the optimal reform

The two graphs report how consumption and wages change over the transition following the optimal tax reform. CSC and CD refer to capital-skill complementarity and Cobb-Douglas economies, respectively.

steady-state welfare criterion does not take these costs into account. This is why under this criterion, low capital taxes are optimal even in the presence of redistributive concerns.

Focusing on steady state comparisons may also shadow our understanding of why unskilled welfare increases as a result of the reform. The average unskilled consumption falls in the long run, much like the average skilled consumption, see Figure 2a. In the short run, however, consumption levels of both types are higher than they are in the initial steady state. This rise in unskilled consumption is large enough to imply welfare gains for them as the reform redistributes from the asset-rich skilled workers to the asset-poor unskilled workers. This happens both in the Cobb-Douglas and the capital-skill complementarity models. The difference between the two models is that, in the latter model, the indirect redistribution channel associated with taxing capital is also at work, both along the transition and in the long run, tilting the distribution of welfare gains further in the direction of unskilled agents.

Table 4: Optimal Capital Taxes under a Different Degree of Capital-Skill Complementarity

	Optimal KORV	Optimal MMT
Capital tax rate, τ	0.60	0.58
Labor tax parameter, λ	0.88	0.87
Skill premium, w_s/w_u	1.67	1.71
Relative skilled wealth	2.33	2.37
Relative skilled consumption	1.41	1.43
Output, Y	92.8	93.3
Capital, K	73.2	74.9
Capital-output ratio, K/Y	1.58	1.61
Overall welfare gains	0.78%	0.70%
Skilled welfare gains	-3.42%	-2.84%
Unskilled welfare gains	2.40%	2.06%

This table reports optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for two capital-skill complementarity economies. The first column reports these values for the benchmark capital-skill complementarity economy with Krusell, Ohanian, Ríos-Rull, and Violante (2000) estimates while the second column reports them for the economy with elasticities taken from Maliar, Maliar, and Tsener (2020) (MMT). Output and capital are normalized to 100 in the status-quo steady state.

5.1 Sensitivity Analysis

Capital-Skill Complementarity. The mechanism that calls for higher optimal taxes on capital income works through the presence of capital-skill complementarity in production. In this regard, we expect our results to be sensitive to the degree of relative substitutability of capital and skilled labor. In this section, we go beyond the elasticity estimates of Krusell, Ohanian, Ríos-Rull, and Violante (2000) and analyze the effect of capital-skill complementarity on optimal capital taxation using other elasticity estimates from the literature.

The estimation in Krusell, Ohanian, Ríos-Rull, and Violante (2000) uses data from the period 1963-1992. However, the world in general and the U.S. economy in particular has been going through an unprecetended technological change in the last three decades. A recent working paper by Maliar, Maliar, and Tsener (2020) estimates the same production function, given by (3), using more recent data, namely data from the period 1963-2017. They find that in the recent data, the elasticity of substitution between equipment and unskilled labor is

about 1.71, and the one between equipment and skilled labor of about 0.76. These values are higher than the corresponding numbers in Krusell, Ohanian, Ríos-Rull, and Violante (2000), which are 1.67 and 0.67, respectively. This implies that equipment capital has become more substitutable with both skilled and unskilled labor. Nonetheless, Maliar, Maliar, and Tsener (2020) conclude that the production function estimated by Krusell, Ohanian, Ríos-Rull, and Violante (2000) and the capital-skill complementarity mechanism remain remarkably successful in explaining the skill premium dynamics.

We take the parameters ρ and η from Maliar, Maliar, and Tsener (2020) and calibrate the capital-skill complementarity economy to make it comparable to both the Cobb-Douglas and the baseline capital-skill complementarity economies. The values of the rest of the externally assigned parameters are identical to the values reported in Table 1. The internal calibration procedure is summarized in Table 12 in Appendix D.1.

The second column of Table 4 reports optimal capital taxes along with the steady-state values of other key variables for this economy. For comparison purposes, we also report the same values for our baseline capital-skill complementarity model with Krusell, Ohanian, Ríos-Rull, and Violante (2000) elasticities in the first column. The optimal capital tax rate under the Maliar, Maliar, and Tsener (2020) calibration is 58%, which is two percentage points lower than the 60% in the baseline case. This might be due to the fact that although both elasticities increased in the new estimation, the elasticity of substitution between equipment and skilled workers increased more. This would imply that the resulting degree of capital-skill complementarity is lower. More importantly though, the difference in the optimal capital tax rates between the capital-skill complementarity model and Cobb-Douglas model is still significant at 10 percentage points.

A key assumption in the analysis of Krusell, Ohanian, Ríos-Rull, and Violante (2000) is their choice of the time series of the price of equipment capital. This choice determines the time series of real stock of equipment capital in the data, which affects the estimation of elasticities. Polgreen and Silos (2008) conduct two sensitivity checks to Krusell, Ohanian,

Ríos-Rull, and Violante (2000) by using two alternative series for the price of equipment capital. They estimate the production function given by (3) using these two alternative series. We conduct the optimal tax analysis for these two additional capital-skill complementarity economies. The optimal capital tax rates for these economies are also around 60%, providing further robustness to our baseline findings. The full discussion of the results of these exercises is relegated to Appendix C for brevity.

Lower Elasticity of Labor Supply. In our benchmark exercise, we take the parameter that controls the Frisch elasticity of labor supply to be $\gamma = 1$, which implies an elasticity of 1. As a sensitivity check, we conduct our main quantitative exercise for $\gamma = 2$, in other words, when Frisch elasticity equals 0.5. Before we calculate optimal taxes, we first calibrate the two economies under $\gamma = 2$.¹⁴ The results of this exercise are reported in Table 5. The optimal capital tax rate equals 64% in the economy with capital-skill complementarity while it is 52% in the Cobb-Douglas economy. The fact that the 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 in the region of empirically plausible elasticities.

Uniform Discount Factor. In the baseline analysis, we allow for differences in the discount factors of skilled and unskilled agents in order to match the ratio of average skilled wealth to average unskilled wealth. In this section, we show that the importance of capital-skill complementarity for the optimal capital tax rate does not depend on the assumption of heterogeneous discount factors. To do so, we first calibrate a version of our model in which all agents have the same discount factor. The calibration procedure is identical to that of the baseline model except that we drop the relative wealth of the skilled and unskilled agents as

¹⁴The values of parameters that are taken from the literature are identical to those in the baseline calibration, and hence are reported in Table 1. The values of internally calibrated parameters are reported in Table 13, which is relegated to Appendix D.2 for brevity.

Table 5: Optimal Capital Taxes under Lower Elasticity of Labor Supply

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.52	0.64
Labor tax parameter, λ	0.81	0.85	0.88
Skill premium, w_s/w_u	1.90	1.88	1.66
Relative skilled wealth	2.78	2.57	2.26
Relative skilled consumption	1.64	1.59	1.46
Output, Y	100	94.5	89.9
Capital, K	100	81.1	66.8
Capital-output ratio, K/Y	2.00	1.72	1.49
Overall welfare gains	–	0.37%	1.06%
Skilled welfare gains	–	-0.72%	-3.35%
Unskilled welfare gains	–	0.76%	2.70%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system for $\gamma = 2$. The second and the third columns report optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models. Output and capital are normalized to 100 in the status-quo steady state.

one of the calibration targets.¹⁵ The resulting model does a poor job in matching the wealth inequality between skill groups.

Table 6 reports our findings. The first row of the table shows that the optimal capital tax rate is 10 percentage points higher in the model with capital-skill complementarity, namely 47% relative to 37% in the Cobb-Douglas model. A comparison of these numbers with the optimal capital tax numbers in the baseline model, reported in the first row of Table 3, reveals that the main quantitative result – the fact that capital-skill complementarity calls for significantly higher optimal taxes on capital income – does not depend on the assumption of heterogeneous discount factors. The importance of capital-skill complementarity for optimal capital taxation is somewhat smaller in the model with homogeneous discount factors. This makes sense as the initial steady-state wealth and consumption inequality is lower in the

¹⁵The values of parameters that are taken from the literature are identical to those in the baseline calibration, and hence, are reported in Table 1. The values of internally calibrated parameters are reported in Table 14, which is relegated to Appendix D.3 for brevity.

Table 6: Optimal Capital Taxes with Uniform Discount Factor

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.37	0.47
Labor tax parameter, λ	0.816	0.818	0.843
Skill premium, w_s/w_u	1.90	1.90	1.84
Relative skilled wealth	1.07	1.07	1.02
Relative skilled consumption	1.39	1.39	1.36
Output, Y	100	99.7	97.3
Capital, K	100	99.0	88.8
Capital-output ratio, K/Y	2.00	1.98	1.83
Overall welfare gains	–	0.00%	0.15%
Skilled welfare gains	–	0.05%	-0.24%
Unskilled welfare gains	–	-0.02%	0.31%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system for the model where all agents have the same discount factor. The second and the third columns report optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models. Output and capital are normalized to 100 in the status-quo steady state.

homogeneous discount factor model. For the same reason, the optimal capital tax rates are lower in the models with homogeneous discount factors.

Two-Agent Model. This section analyzes the role of labor income risk for the main result. To that end, we analyze a two-agent version of our model in which labor income risk is shut down. There are two representative agents, a skilled and an unskilled with the corresponding population shares. The calibration procedure is similar to that of the baseline model.¹⁶

As Table 7 indicates, the optimal tax differential between the Cobb-Douglas and capital-skill complementarity models remains significant at 8 percentage points, but is lower than in the baseline model with wage shocks. The optimal capital tax rates are positive but lower

¹⁶In this model for a non-trivial steady state distribution of assets to exist, the discount factors must be uniform across the skilled and the unskilled. In fact, in that case any steady state asset distribution can arise in equilibrium. We focus on the steady state in which the ratio of skilled to unskilled asset holdings is 2.78, as in the baseline model. The parameters taken from the literature are the same as in the benchmark exercise, which are presented in Table 1. The internal calibration procedure is summarized in Table 15 in Appendix D.4. It is worth noting that given the calibrated initial steady state, the post-reform transition paths and the terminal steady states to which the economies converge are unique.

Table 7: Optimal Capital Taxes in a Two-Agent Model

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.30	0.38
Labor tax parameter, λ	0.533	0.524	0.536
Skill premium, w_s/w_u	1.90	1.91	1.84
Relative skilled wealth	2.78	2.76	2.80
Relative skilled consumption	1.55	1.56	1.55
Output, Y	100	101.4	99.6
Capital, K	100	106.2	98.0
Capital-output ratio, K/Y	2.00	2.09	1.97
Overall welfare gains	–	0.05%	0.01%
Skilled welfare gains	–	0.58%	-0.32%
Unskilled welfare gains	–	-0.13%	0.12%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system for the model in which there is no labor income risk, i.e., there are only two representative agents. The second and the third columns report optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models. Output and capital are normalized to 100 in the status-quo steady state.

relative to the baseline case with labor income risk for both the Cobb-Douglas and capital-skill complementarity models, which is in line with the findings of Domeij and Heathcote (2004). In addition, since the optimal capital tax rates are now closer to status quo, the welfare implications are smaller. These results imply that modelling risk, and the resulting rich heterogeneity, is important for optimal capital taxes in general and for evaluating the quantitative impact of capital-skill complementarity for optimal capital taxes in particular. The baseline model with labor income risk is also desirable as it allows us to analyze the distributional consequences of capital tax reforms in greater detail.

Role of Top Wealth Inequality. In our baseline analysis, we use Gaussian individual labor productivity processes that are estimated using panel data that follows workers over time. This approach is useful as it allows for a realistic modeling of individual labor income dynamics. However, it is well known that the class of models analyzed in this paper together with this type of modeling of labor productivity shocks fall short of matching the inequality at

Table 8: Optimal Capital Taxes With an Alternative Productivity Process

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.82	0.88
Labor tax parameter, λ	0.81	0.99	1.03
Skill premium, w_s/w_u	1.90	1.83	1.18
Relative skilled wealth	2.78	2.46	1.91
Relative skilled consumption	1.58	1.47	1.14
Output, Y	100	81.3	75.0
Capital, K	100	43.1	32.9
Capital-output ratio, K/Y	2.00	1.06	0.88
Overall welfare gains	–	4.47%	6.25%
Skilled welfare gains	–	0.60%	-5.16%
Unskilled welfare gains	–	5.90%	11.05%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system. The second and third columns report optimal capital taxes and the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models with the Castaneda, Díaz-Giménez, and Ríos-Rull (2003) productivity process. Output and capital are normalized to 100 in the status-quo steady state.

the top end of the wealth distribution. In an alternative approach to modeling productivity risk, proposed by Castaneda, Díaz-Giménez, and Ríos-Rull (2003), the productivity processes are parameterized so as to match the thick right tail of the wealth distribution. In this section, we analyze to what extent using this approach affects our main results. More specifically, we recalibrate the Cobb-Douglas and the capital-skill complementarity models using the Castaneda, Díaz-Giménez, and Ríos-Rull (2003) productivity process for both the skilled and the unskilled agents.¹⁷ In the two calibrated economies, individuals who constitute the top 1% of the wealth distribution hold 29% of aggregate wealth and the those at the top 5% hold approximately 50%. These numbers are in line with the data as documented by Castaneda, Díaz-Giménez, and Ríos-Rull (2003), 30% and 54%, respectively. The corresponding numbers are significantly lower in the baseline model: 8% and 27%, respectively.

¹⁷The parameters taken from the literature are the same as in the benchmark and the internal calibration procedure is summarized in Table 16 in Appendix D.5.

The results from these exercises are summarized in Table 8. We find that the optimal capital taxes are much larger in both the Cobb-Douglass and the capital-skill complementarity models relative to the baseline calibration. This is in line with Dyrda and Pedroni (2020) who also find high capital taxes to be optimal in the presence of Castaneda, Díaz-Giménez, and Ríos-Rull (2003) productivity processes.¹⁸ More importantly, there still is a significant difference between the optimal capital tax in the Cobb-Douglas and capital-skill complementarity models, namely 6 percentage points, confirming the robustness of our main result.

Rawlsian Social Welfare. In our baseline analysis, we assume that the government evaluates the consequences of the reform by aggregating agents' welfare using a Utilitarian social 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 fortunate 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 welfare for the agent with the lowest welfare. Formally, the government solves the following problem:

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

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

The results of this exercise are reported in Table 9. We find that the optimal capital tax rate is 85% in the economy with capital-skill complementarity while it is 66% in the Cobb-Douglas economy. Since redistributive considerations are more important 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 (in both economies) and to reduce the skill premium (in the capital-skill complementarity economy) to as low as 1.10.

¹⁸See Boar and Midrigan (2020) for an analysis of optimal taxation of total income - capital plus labor - in a model where labor productivity is modeled as in Castaneda, Díaz-Giménez, and Ríos-Rull (2003).

Table 9: Optimal Capital Taxes under Rawlsian Social Welfare

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.66	0.85
Labor tax parameter, λ	0.81	0.89	0.95
Skill premium, w_s/w_u	1.90	1.83	1.13
Relative skilled wealth	2.78	2.30	1.56
Relative skilled consumption	1.55	1.46	1.10
Output, Y	100	88.9	74.6
Capital, K	100	61.9	32.1
Capital-output ratio, K/Y	2.00	1.39	0.89
Overall welfare gains	–	-0.40%	-1.05%
Skilled welfare gains	–	-2.90%	-13.08%
Unskilled welfare gains	–	0.54%	4.23%

The first column of the table reports status-quo capital and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system. The second and the third columns report optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models when the government’s social welfare function is Rawlsian. Output and capital are normalized to 100 in the status-quo steady state.

6 Comprehensive Tax Reform

So far we have focused on the effect of capital-skill complementarity on the optimal capital tax rate in the context of a tax reform in which the government is only able to adjust the capital tax rate along with the parameter that controls the average labor income tax, λ . In particular, this reform does not involve setting the labor income tax progressivity parameter, τ_l , optimally. We pursued this route mainly because, perhaps due to political constraints, it is often quite difficult for the government to consider comprehensive reforms in which capital and labor tax codes are reformed together. This section aims to gauge the effect of capital-skill complementarity on the optimal capital tax rate in the context of such a comprehensive tax reform. To be precise, we consider a problem of a government which introduces a once and for all unannounced change in the capital tax rate, τ , and in labor tax progressivity, τ_l . As in the baseline, to ensure that its budget holds, the government adjusts the parameter that controls the average labor income tax, $\{\lambda_t\}_{t=0}^{\infty}$, along the transition to the new steady

Table 10: Optimal Capital Taxes under Comprehensive Tax Reform

	Status quo Calibrated	Optimal CD	Optimal CSC
Capital tax rate, τ	0.36	0.48	0.60
Labor tax progressivity, τ_l	0.18	0.14	0.12
Labor tax parameter, λ	0.81	0.85	0.89
Skill premium, w_s/w_u	1.90	1.86	1.67
Relative skilled wealth	2.78	2.45	2.13
Relative skilled consumption	1.55	1.52	1.41
Output, Y	100	98.2	95.3
Capital, K	100	88.7	76.1
Capital-output ratio, K/Y	2.00	1.81	1.60
Overall welfare gains	–	0.30%	0.88%
Skilled welfare gains	–	0.34%	-2.08%
Unskilled welfare gains	–	0.29%	2.01%

The first column of the table reports status-quo capital taxes, optimal degree of labor tax progressivity and average labor income taxes as well as the values of key variables in the steady state under the status-quo tax system. The second and the third columns report optimal capital taxes, optimal degree of labor tax progressivity and the steady-state values of average labor income taxes and other key variables under the optimal tax system for both the Cobb-Douglas (CD) and capital-skill complementarity (CSC) models. Output and capital are normalized to 100 in the status-quo steady state.

state. The welfare criterion puts equal weight on each agent and takes transition into account as in the baseline case.

Table 10 summarizes our findings. Looking at the first column of the table, we see that the optimal capital tax rate is much higher, namely 12 percentage points, in the model with capital-skill complementarity.¹⁹ This means that the quantitative significance of capital-skill complementarity for the optimal capital tax rate is not special to the partial reform considered in Section 5 and continues to apply when the government is able to impact the amount of redistribution and insurance via the progressivity of the labor tax system as well. In both economies, the government finds it optimal to decrease labor tax progressivity relative to the status-quo tax system. This mitigates the efficiency losses associated with higher capital taxes. The finding that in the full reform, the government increases capital

¹⁹The optimal capital taxes in the comprehensive reform are not exactly the same as in the partial reform. The fact that it appears that way in the table is a by-product of rounding.

taxes and decreases labor tax progressivity relative to the status-quo tax system is in line with the findings of Conesa, Kitao, and Krueger (2009).²⁰ However, in their paper the motive for capital taxation is different, namely that it mimics the missing age-dependent labor income taxation.

We also find that the decline in the progressivity of the labor income tax code benefits mostly the skilled: in both economies, the measure of skilled agents who gain from the reform is larger than in the partial reform. Specifically, the fraction of the skilled who gain from the comprehensive reform is 58% in the Cobb-Douglas case and 38% in the capital-skill complementarity case, while the corresponding numbers are 45% and 11% in the partial reform.

7 Conclusion

This paper shows that capital-skill complementarity provides a quantitatively significant rationale for taxing capital for redistributive governments. The paper finds that it is optimal to rely much more on capital income taxes and less on labor income taxes when capital-skill complementarity is taken into account. The welfare gains of an optimal tax reform are also significantly larger in the presence of capital-skill complementarity. Given the overwhelming empirical evidence on the presence of capital-skill complementarities in production, our analysis suggests that governments should take into account the presence of such complementarities when setting capital tax rates.

²⁰See Peterman (2013) for a thorough sensitivity analysis of Conesa, Kitao, and Krueger (2009).

References

- ACIKGOZ, O., M. HAGEDORN, H. HOLTER, AND Y. WANG (2018): “The Optimum Quantity of Capital and Debt,” Cepr discussion paper 12952.
- AIYAGARI, S. R. (1994): “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 109(3), 659–84.
- (1995): “Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting,” *Journal of Political Economy*, 103(6), 1158–75.
- ANGELOPOULOS, K., S. ASIMAKOPOULOS, AND J. MALLEY (2015): “Tax Smoothing in a Business Cycle Model with Capital-Skill Complementarity,” *Journal of Economic Dynamics and Control*, 51, 420–444.
- ATTANASIO, O. P., J. BANKS, C. MEGHIR, AND G. WEBER (1999): “Humps and Bumps in Lifetime Consumption,” *Journal of Business & Economic Statistics*, 17(1), 22–35.
- BAKIS, O., B. KAYMAK, AND M. POSCHKE (2015): “Transitional Dynamics and the Optimal Progressivity of Income Redistribution,” *Review of Economic Dynamics*, 18(3), 679–693.
- BEWLEY, T. F. (1986): “Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers,” in *Contributions to Mathematical Economics in Honor of Gerard Debreu*, ed. by W. Hildenbrand, and A. Mas-Colell, pp. 79–102. North Holland.
- BHATTARAI, S., J. W. LEE, W. Y. PARK, AND C. YANG (2020): “Macroeconomic Effects of Capital Tax Rate Changes,” Working paper.
- BOAR, C., AND V. MIDRIGAN (2020): “Efficient Redistribution,” Working Paper 25103, National Bureau of Economic Research, Inc.
- CASTANEDA, A., J. DÍAZ-GIMÉNEZ, AND J.-V. RÍOS-RULL (2003): “Accounting for the U.S. Earnings and Wealth Inequality,” *Journal of Political Economy*, 111(4), 818–857.

- CHAMLEY, C. (1986): “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, 54(3), 607–622.
- CHARI, V., J. P. NICOLINI, AND P. TELES (2019): “Optimal capital taxation revisited,” *Journal of Monetary Economics*.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): “Taxing Capital? Not a Bad Idea after All!,” *American Economic Review*, 99(1), 25–48.
- COSTINOT, A., AND I. WERNING (2018): “Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation,” NBER Working Papers 25103, National Bureau of Economic Research, Inc.
- DOLADO, J. J., G. MOTYOVSKI, AND E. PAPP (2020): “Monetary Policy and Inequality under Labor Market Frictions and Capital-Skill Complementarity,” *American Economic Journal: Macroeconomics*, forthcoming.
- DOMEIJ, D., AND J. HEATHCOTE (2004): “On the Distributional Effects of Reducing Capital Taxes,” *International Economic Review*, 45(2), 523–554.
- DUFFY, J., C. PAPAGEORGIOU, AND F. PEREZ-SEBASTIAN (2004): “Capital-Skill Complementarity? Evidence from a Panel of Countries,” *Review of Economics and Statistics*, 86(1), 327–344.
- DYRDA, S., AND M. PEDRONI (2020): “Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks,” Working paper.
- FALLON, P. R., AND P. R. G. LAYARD (1975): “Capital-Skill Complementarity, Income Distribution, and Output Accounting,” *Journal of Political Economy*, 83(2), 279–302.
- FLUG, K., AND Z. HERCOWITZ (2000): “Equipment Investment and the Relative Demand for Skilled Labor: International Evidence,” *Review of Economic Dynamics*, 3, 461–485.

- GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): “Optimal Indirect and Capital Taxation,” *Review of Economic Studies*, 70(3), 569–587.
- GREENWOOD, J., Z. HERCOWITZ, AND P. KRUSELL (1997): “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review*, 87(3), 342–362.
- GRILICHES, Z. (1969): “Capital-Skill Complementarity,” *Review of Economics and Statistics*, 51(4), 465–468.
- GUERREIRO, J., S. REBELO, AND P. TELES (2017): “Should Robots be Taxed?,” NBER Working Papers 23806, National Bureau of Economic Research, Inc.
- HEATHCOTE, J., F. PERRI, AND G. L. VIOLANTE (2010): “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967–2006,” *Review of Economic Dynamics*, 13(1), 15–51.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2017): “Optimal Tax Progressivity: An Analytical Framework,” *The Quarterly Journal of Economics*, 132(4), 1693–1754.
- HUGGETT, M. (1993): “The risk-free rate in heterogeneous-agent incomplete-insurance economies,” *Journal of Economic Dynamics and Control*, 17(5-6), 953–969.
- IMROHORGLU, A. (1989): “Cost of Business Cycles with Indivisibilities and Liquidity Constraints,” *Journal of Political Economy*, 97(6), 1364–1383.
- IMROHORGLU, S. (1998): “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): “Redistributive Taxation in a Simple Perfect Foresight Model,” *Journal of Public Economics*, 28(1), 59–83.

- KATZ, L. F., AND K. M. MURPHY (1992): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, 107(1), 35–78.
- KOPECKY, K., AND R. SUEN (2010): “Finite State Markov-chain Approximations to Highly Persistent Processes,” *Review of Economic Dynamics*, 13(3), 701–714.
- KRUEGER, D., AND A. LUDWIG (2016): “On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium,” *Journal of Monetary Economics*, 77, 72 – 98, Inequality, Institutions, and Redistribution held at the Stern School of Business, New York University, April 24-25, 2015.
- KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis,” *Econometrica*, 68(5), 1029–1054.
- MALIAR, L., S. MALIAR, AND I. TSENER (2020): “Capital-Skill Complementarity and Inequality: Twenty Years After,” Cepr discussion paper no: 15228.
- PETERMAN, W. B. (2013): “Determining the motives for a positive optimal tax on capital,” *Journal of Economic Dynamics and Control*, 37(1), 265–295.
- POLGREEN, L., AND P. SILOS (2008): “Capital-Skill Complementarity and Inequality: A Sensitivity Analysis,” *Review of Economic Dynamics*, 11(2), 302–313.
- SLAVÍK, C., AND H. YAZICI (2014): “Machines, Buildings, and Optimal Dynamic Taxes,” *Journal of Monetary Economics*, 66, 47 – 61.
- STRAUB, L., AND I. WERNING (2020): “Positive Long-Run Capital Taxation: Chamley-Judd Revisited,” *American Economic Review*, 110(1), 86–119.
- THUEMMEL, U. (2020): “Optimal Taxation of Robots,” Cesifo working papers.
- TRABANDT, M., AND H. UHLIG (2011): “The Laffer curve revisited,” *Journal of Monetary Economics*, 58(4), 305–327.

Appendix

A Equilibrium Definitions

A.1 Competitive Equilibrium in the 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 the consumer's problem, i.e.,

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

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,

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

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:

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

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

$$\begin{aligned}
K_t + D_t &= \sum_{i=u,s} \pi_i \int_{\mathcal{V}_i} \sum_{z^{t-1} \in Z_i^{t-1}} P_{i,t-1}(z^{t-1}|z_0) a_{i,t}(v_0, z^{t-1}) d\lambda_i^0(v_0), \\
L_{i,t} &= \pi_i \int_{\mathcal{V}_i} \sum_{z^t \in Z_i^t} P_{i,t}(z^t|z_0) l_{i,t}(v_0, z^t) z_t d\lambda_i^0(v_0), \text{ for } i = u, s, \\
G_t + C_t + K_{t+1} &= F(K_t, L_{s,t}, L_{u,t}) + (1 - \delta)K_t,
\end{aligned}$$

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 the 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 $i \in \{u, s\}$:

$$V_i(z, a) = \max_{(c_i, l_i, a'_i) \geq 0} (1 - \beta_i) [u(c_i) - v(l_i)] + \beta_i \sum_{z'} \Pi_i(z'|z) V_i(z', a'_i) \quad \text{s.t.}$$

$$c_i + a'_i \leq w_i z l_i - T(w_i z l_i) + R a_i,$$

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. $\forall i : \lambda'_i(z, a) = \lambda_i(z, a)$. Where

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

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.

B Data Construction

Fraction of skilled agents. The fraction of skilled agents is calculated using Current Population Survey ASEC (March) data administered by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics. We use data from the 2018 survey which includes information about 2017. We follow Krusell, Ohanian, Ríos-Rull, and Violante (2000) and define the fraction of skilled agents as the ratio of males aged 25 and older with earnings and a bachelor’s degree or more divided by the total number of males aged 25 and older with earnings in Table P-16.

Government consumption-to-GDP ratio. The government consumption-to-output ratio is recovered from the National Income and Product Accounts (NIPA) data. It is defined as the ratio of nominal government consumption expenditure (line 15 in NIPA Table 3.1) to nominal GDP (line 1 in NIPA Table 1.1.5).

Government debt-to-GDP ratio. The government debt to GDP ratio is taken from the St. Louis FED database FRED for year 2015. The data series is called “Federal Debt Held by Private Investors as Percent of Gross Domestic Product” (HBPIGDQ188S).

Share of equipments in total capital stock. The share of equipment capital in total capital stock is calculated using Fixed Asset Tables (FAT) data. It is defined as the ratio of private equipment capital (line 5 in FAT Table 1.1) to the sum of private equipment and structure capital (line 5 + line 6 in FAT Table 1.1). This calculation gives a value of 0.32 for the period 2010-2018, which we round to 1/3.

Capital-to-output ratio. Housing is excluded from both output and capital when calculating the capital-to-output ratio. For this calculation, 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 between the late 60s and the present. To abstract from short-term fluctuations, the capital-output ratio value of 2 is computed by taking the average of annual capital-output ratios over this period.

Relative skilled wealth. This statistic is calculated using “Wealth, Asset Ownership, & Debt of Households Detailed Tables: 2014” administered by US Census Bureau. Table 4 includes information about the number of respondents in a given education category and Table 5 includes information about average net worth by education category. Using these two tables, we calculate the average wealth of skilled workers as the weighted average net worth of people with a bachelor’s degree or higher. Similarly, we calculate average wealth of unskilled workers as the weighted average net worth of the rest. The relative skilled wealth is then calculated by dividing average skilled wealth with average unskilled wealth.

C Additional Sensitivity with Respect to Capital-Skill Complementarity

A key assumption in the analysis of Krusell, Ohanian, Ríos-Rull, and Violante (2000) is their choice of the time series of the price of equipment capital. This choice determines the time series of real stock of equipment capital in the data, which presumably affects the estimation of elasticities. Polgreen and Silos (2008) conduct two sensitivity checks to Krusell, Ohanian, Ríos-Rull, and Violante (2000) by using two alternative series for the price of equipment capital: (i) the price series from the official, revised National Income and Product Accounts data (NIPA) and (ii) the price series proposed by Greenwood, Hercowitz,

Table 11: Optimal Capital Taxes under Different Degrees of Capital-Skill Complementarity

	Optimal KORV	Optimal NIPA	Optimal GHK	Optimal MMT
Capital tax rate, τ	0.60	0.61	0.58	0.58
Labor tax parameter, λ	0.88	0.88	0.86	0.87
Skill premium, w_s/w_u	1.67	1.63	1.69	1.71
Relative skilled wealth	2.33	2.28	2.33	2.37
Relative skilled consumption	1.41	1.38	1.41	1.43
Output, Y	92.8	92.4	92.4	93.3
Capital, K	73.2	72.1	72.3	74.9
Capital-output ratio, K/Y	1.58	1.58	1.56	1.61
Overall welfare gains	0.78%	0.82%	0.68%	0.70%
Skilled welfare gains	-3.42%	-4.12%	-3.14%	-2.84%
Unskilled welfare gains	2.40%	2.76%	2.15%	2.06%

This table reports optimal capital taxes as well as the steady-state values of average labor income taxes and other key variables under the optimal tax system for a variety of capital-skill complementarity economies. The first column reports these values for the benchmark capital-skill complementarity economy with Krusell, Ohanian, Ríos-Rull, and Violante (2000) estimates while the second, the third and the fourth columns report them for the economies with elasticities taken from Polgreen and Silos (2008), for NIPA and Greenwood, Hercowitz, and Krusell (1997) cases, and Maliar, Maliar, and Tsener (2020). Output and capital are normalized to 100 in the status-quo steady state.

and Krusell (1997) (GHK). They estimate the two key parameters, ρ , which controls the elasticity of substitution between skilled labor and equipment, and η , which controls the elasticity of substitution between unskilled labor and equipment, for NIPA and GHK series.

The second and the third columns of Table 11 report optimal capital and average (steady-state) labor income taxes as well as the steady-state values of key variables under the optimal tax scheme in economies characterized by these alternative production function elasticities. The column entitled “NIPA” refers to the economy in which production function parameters are coming from the estimation where NIPA equipment capital price series are used, while the column entitled “GHK” refers to the one in which equipment capital price series of Greenwood, Hercowitz, and Krusell (1997) are used. We see that optimal capital tax rates in these economies are 61% and 58%, respectively, which is close to the optimal capital tax rate in the baseline capital-skill complementarity economy with Krusell, Ohanian, Ríos-Rull, and Violante (2000) elasticities, which is reported in the first column of the same table. For

comparison purposes, the last column of the table also reproduces optimal capital tax rates and the values of steady-state variables for the economy where production function elasticities are taken from Maliar, Maliar, and Tsener (2020).

D Details of Calibration for Sensitivity Analysis

D.1 Maliar, Maliar, and Tsener (2020) Elasticities Calibration

Table 12 below reports the values of internally calibrated parameters for the economy with the Maliar, Maliar, and Tsener (2020) elasticities.

Table 12: Maliar, Maliar, and Tsener (2020) Elasticities Calibration Procedure

Parameter	Symbol	Value	Target	Source
Technology				
Production parameter	ω	0.3111	Labor share = 2/3	NIPA
Production parameter	ν	0.6235	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K_e}{K} = 1/3$	FAT
Skilled discount factor	β_s	0.9415	Capital to output ratio = 2	NIPA, FAT
Unskilled discount factor	β_u	0.9365	Relative skilled wealth = 2.78	US Census
Tax function parameter	λ	0.814	Government budget balance	
Disutility of labor	ϕ	66.08	Labor supply = 1/3	

This table reports our calibration procedure for the economy with elasticities taken from Maliar, Maliar, and Tsener (2020). The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). 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.

D.2 $\gamma = 2$ Calibration

Table 13 below reports the values of internally calibrated parameters for the version of the model in which $\gamma = 2$.

Table 13: $\gamma = 2$ Calibration Procedure

Parameter	Symbol	Value	Target	Source
Technology				
Production parameter	ω	0.3160	Labor share = 2/3	NIPA
Production parameter	ν	0.6310	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K_e}{K} = 1/3$	FAT
Technology (Cobb-Douglas)				
Total factor productivity	A	0.7787	Output level of CSC economy	
Production parameter	κ	0.5254	Skill premium = 1.9	CPS
Skilled discount factor	β_s	0.9413	Capital to output ratio = 2	NIPA, FAT
Unskilled discount factor	β_u	0.9369	Relative skilled wealth = 2.78	US Census
Tax function parameter	λ	0.815	Government budget balance	
Disutility of labor	ϕ	195.74	Labor supply = 1/3	

This table reports our calibration procedure for the case where $\gamma = 2$. The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). The production function parameter κ controls the income shares of the skilled and unskilled labor in the Cobb-Douglas model (CD). 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.

D.3 Uniform Discount Factor Calibration

Table 14 below reports the values of internally calibrated parameters for the version of the model in which discount factors are uniform across the skilled and unskilled agents.

D.4 Two-Agent Model Calibration

Table 15 below reports the values of internally calibrated parameters for the two-agent version of the model with no labor income risk.

Table 14: Uniform Discount Factor Calibration Procedure

Parameter	Symbol	Value	Target	Source
Technology				
Production parameter	ω	0.3065	Labor share = 2/3	NIPA
Production parameter	ν	0.6371	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K^e}{K} = 1/3$	FAT
Technology (Cobb-Douglas)				
Total factor productivity	A	0.7846	Output level of CSC economy	
Production parameter	κ	0.5254	Skill premium = 1.9	CPS
Discount factor	β	0.9395	Capital to output ratio = 2	NIPA, FAT
Tax function parameter	λ	0.816	Government budget balance	
Disutility of labor	ϕ	62.80	Labor supply = 1/3	

This table reports our calibration procedure for the case in which all agents have the same discount factor. The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). The production function parameter κ controls the income shares of the skilled and unskilled labor in the Cobb-Douglas model (CD). 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.

Table 15: Two-Agent Model Calibration Procedure

Parameter	Symbol	Value	Target	Source
Technology				
Production parameter	ω	0.3329	Labor share = 2/3	NIPA
Production parameter	ν	0.6207	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K^e}{K} = 1/3$	FAT
Technology (Cobb-Douglas)				
Total factor productivity	A	0.7830	Output level of CSC economy	
Production parameter	κ	0.5010	Skill premium = 1.9	CPS
Discount factor	β	0.9467	Capital to output ratio = 2	NIPA, FAT
Tax function parameter	λ	0.533	Government budget balance	
Disutility of labor	ϕ	58.90	Labor supply = 1/3	

This table reports our calibration procedure for the two-agent model. The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). The production function parameter κ controls the income shares of the skilled and unskilled labor in the Cobb-Douglas model (CD). 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.

D.5 Top Wealth Inequality Model Calibration

Table 16 below reports the values of internally calibrated parameters for the economy with the Castaneda, Díaz-Giménez, and Ríos-Rull (2003) productivity process.

Table 16: Alternative Productivity Process Calibration Procedure

Parameter	Symbol	Value	Target	Source
Technology				
Production parameter	ω	0.3247	Labor share = 2/3	NIPA
Production parameter	ν	0.6256	Skill premium = 1.9	CPS
Production parameter	α	0.1920	Share of equipments, $\frac{K_e}{K} = 1/3$	FAT
Technology (Cobb-Douglas)				
Total factor productivity	A	0.7835	Output level of CSC economy	
Production parameter	κ	0.5083	Skill premium = 1.9	CPS
Skilled discount factor	β_s	0.9226	Capital to output ratio = 2	NIPA, FAT
Unskilled discount factor	β_u	0.9094	Relative skilled wealth = 2.78	US Census
Tax function parameter	λ	0.860	Government budget balance	
Disutility of labor	ϕ	119.72	Labor supply = 1/3	

This table reports our calibration procedure for the case in which productivity process taken from Castaneda, Díaz-Giménez, and Ríos-Rull (2003). The production function parameters α , ν and ω control the income shares of structure capital, equipment capital, skilled and unskilled labor in the capital-skill complementarity model (CSC). The production function parameter κ controls the income shares of the skilled and unskilled labor in the Cobb-Douglas model (CD). 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.