On the optimal capital tax rate in overlapping generations models with capital–skill complementarity

Burkhard Heer

Angaben zur Veröffentlichung / Publication details:

Heer, Burkhard. 2025. “On the optimal capital tax rate in overlapping generations models with capital–skill complementarity.” International Tax and Public Finance. https://doi.org/10.1007/s10797-025-09938-0.

International Tax and Public Finance
https://doi.org/10.1007/s10797-025-09938-0

On the optimal capital tax rate in overlapping generations models with capital–skill complementarity

Burkhard Heer[1,2,3]

Received: 22 April 2025 / Accepted: 25 November 2025

Abstract

We show in a medium-scale overlapping-generations model that the optimal capital income tax rate is highly sensitive to the assumption of capital–skill complementarity in production technology. The imposition of the production function of (Krusell et al. in Econometrica, 2000, 68, 1029–1053) rather than the standard Cobb–Douglas function increases the optimal capital tax from 9.2% to 27.3% in our benchmark model. We also study the sensitivity of these results in the context of an aging economy and find that the optimal capital income tax increases over the subsequent decades. With respect to redistributive policies, higher capital taxes are more welfare-improving than are progressive labor income taxes if capital is strongly complementary to skilled labor.

Keywords Capital income taxes · Chamley-Judd result · Skill-biased technological change · Demographic change · Redistributive policies

JEL classification E13 · H21 · H24 · H25

We would like to thank Thiess Büttner, Andreas Irmen, Vito Polito, Ctirad Slavík and three anonymous referees for helpful comments. All remaining errors are ours.

✉ Burkhard Heer
burkhard.heer@wiwi.uni-augsburg.de

1 University of Augsburg, Department of Economics, Universitatsstrasse 16, 86159 Augsburg, Germany

2 CESifo, Munich, Germany

3 Netspar, Tilburg, Netherlands

Published online: 09 December 2025

1 Introduction

Chamley (1986) and Judd (1985) show that the optimal long-term capital tax rate is zero in the standard neoclassical growth model with infinitely lived agents. The intuition for the result is straightforward. The distortion in the intertemporal consumption allocation introduced by a capital income tax is compounded over the lifetime and outweighs the intratemporal distortion from a labor income tax. The literature has identified many modeling choices that modify the Chamley-Judd result and imply optimal nonzero capital taxes. The present study focuses on two prominent modeling choices, 1) capital-skill complementarity and 2) life-cycle aspects, and combines them to evaluate quantitatively the optimal income tax rate. Our motivation for the joint analysis of these two elements is twofold: First, we show that the quantitative effect of a production technology with capital-skill complementarity vis-à-vis a standard Cobb-Douglas technology is much stronger in a model with finite than infinite lifetime. In our numerical example, the optimal tax rate increases by approximately 21 percentage points (finite lifetime) rather than 8 percentage points (infinite lifetime). Second, the combination of these two model elements allows us to study the effects of various fiscal and pension policies as well as demographic factors on welfare and distribution in a medium-scale overlapping generations (OLG) model. As a natural application, we consider the effects of the demographic transition on optimal capital taxes.

The study most closely related to the present one is Slavík and Yazici (2014). As a common model element, production is characterized by skill complementarity such that capital taxes redistribute income indirectly as the skill premium becomes endogenous. Higher capital taxes reduce capital, which is complementary to skilled labor, and consequently, the wage of the skilled worker decreases relative to that of the unskilled worker. In our analysis, a more equitable distribution of after-tax income results in higher welfare ceteris paribus. Whereas Slavík and Yazici (2014) study a model with individuals who live for infinitely many periods, we introduce life-cycle elements into our model, including a finite lifetime, stochastic survival, pensions and age-dependent productivities. Life-cycle aspects have been demonstrated to imply a significantly positive optimal capital tax. For example, Conesa and Krueger (2009) find optimal capital tax rates in excess of 30% in their medium-sized overlapping generations (OLG) model.[1] As one of our main results, we show that the two mechanisms, 1) capital-skill complementarity and 2) life-cycle elements, reinforce each other and that the welfare effects of the combined mechanisms are greater than the sum of their individual welfare effects, thus implying significantly greater optimal capital taxes.

[1] The theoretical result that the Chamley-Judd results do not hold in models with finite lifetimes has been shown by Alvarez et al. (1992) and Erosa and Gervais (2002). As a necessary assumption for this result, the government cannot use age-specific labor income taxes. Workers at different ages are characterized by different labor supply elasticities such that a uniform labor income tax schedule is not optimal and a capital income tax might help substitute for age-specific labor income taxes.

Our welfare analysis follows Conesa and Krueger (2009).[2] We study a competitive equilibrium. Firms and individuals react to changes in fiscal policy, so the allocation of goods and factors, along with prices, reflects the optimizing behavior of the agents. The government chooses from a set of fiscal policies that we restrict to a proportional tax on capital income and the corresponding level of labor income taxes that balance the fiscal budget for exogenously given government expenditures. As our main welfare criterion, we use the expected lifetime utility of the newborn generation. However, unlike in the analysis of Conesa and Krueger (2009) we use the technology with capital-skill complementarity specified in Krusell et al. (2000) rather than the standard Cobb-Douglas technology. Our assumption of capital-skill complementarity in production is motivated mainly by empirical evidence, as the 'Krusell et al.' technology helps explain the behavior of the US skill premium and its increase from 140% in the early 1980s to 200% in the present.[3] Various studies consider the empirical validity of the underlying mechanism in the model of Krusell et al. (2000). In particular, Flug and Hercowitz (2000) provide cross-country empirical evidence for skill complementarity in production. Using panel data from a wide range of countries, they show that the relative wages of skilled and unskilled workers depend positively and strongly on equipment investment. Maliar et al. (2022) extend the sample period 1963-1992 in Krusell et al. (2000) to a more recent period (1992-2017) during which the skill premium increased monotonically and confirm that capital skill complementarity helps explain this trend.

As the main contribution of this paper, we consider the effects of production technology on optimal capital taxes. We set up a medium-scale OLG model with progressive income taxes and public pay-as-you-go pensions and find, in our benchmark calibrated with respect to the characteristics of the U.S. economy during 1980-2020, an optimal capital income tax rate of 27%, whereas the optimal tax rate in the corresponding model with Cobb–Douglas production equals only 9%. We identify two main channels through which production technology affects the welfare effects of high capital taxes. 1) The marginal product of capital is more sensitive to aggregate capital in the case of capital-skill complementary production technology than in the case of Cobb-Douglas technology. Therefore, higher capital taxes imply a stronger decrease in net real interest rates and a stronger adjustment of the allocation of labor over the life cycle. The flatter labor-supply profile that results is welfare increasing. 2) The skill premium is constant for the Cobb–Douglas technology, whereas it decreases with higher capital taxes in the case of the technology with capital–skill complementarity. Through this channel, higher capital taxes support redistribution from the rich to the poor and, given the concave functional form of utility, increase

[2] Unlike our study, the welfare analysis of Slavík and Yazici (2014) follows Mirrless (1971). In the canonical Mirrless framework, the planner can tailor the tax system to the underlying skill (or productivity) type inferred via income and labor effort. However, in practice, governments must apply a single tax schedule. The application of the Mirrless framework becomes even less feasible, both theoretically and computationally, if we consider an overlapping generations model with age-dependent productivities in which the heterogeneity of individuals implies several hundred incentive compatibility constraints for the welfare-maximizing planner. Instead, we apply a single tax function over income that is identical for all individuals.

[3] See, for example, Maliar et al. (2022).

the average lifetime utility. We subject our analysis to a series of sensitivity analyses and show that the optimal capital income tax rate is even greater in the case of a flat-rate income tax or with low economic growth and, in the presence of a progressive income tax, is rather insensitive to the introduction of stochastic idiosyncratic uninsurable income risk.

As one interesting application, we analyze the effects of demographic change on optimal capital taxes. We assume that the population characteristics of the U.S. economy are as predicted by United Nations (2022) for 2050. We demonstrate that the optimal capital income tax increases in an aging economy and reaches 35% in 2050. In addition, we show that the transitional effects from a capital tax increase deviate significantly from its long-term effects and emphasize that current generations would suffer from an increase in the capital tax rate.

1.1 Related literature

This paper focuses on the contribution of capital-skill complementarity and life-cycle elements to the optimal capital tax rate. A vast literature considers other reasons for nonzero optimal capital income taxation. Hubbard and Judd (1987) and Aiyagari (1995) have identified 1) borrowing constraints and 2) uninsurable income risk as potential reasons why the Chamley–Judd result does not hold. The intuition for their results is very straightforward. 1) If a borrowing constraint on the household is binding, the Euler condition for the intertemporal allocation of consumption is violated, and consumption growth deviates from the Pareto-optimal condition in which the marginal rate of substitution between consumption in two periods is equal to (one plus) the rate of return on capital. If fiscal revenue stems from capital rather than labor income taxes, borrowing constraints, particularly in the early periods of life, become less binding as net labor income increases. 2) In the case of uninsurable stochastic income shocks, capital income taxes help redistribute income from rich to poor agents and increase aggregate welfare if society values a more equal distribution. To derive a quantitative estimate of the optimal capital income tax in life-cycle models with idiosyncratic income uncertainty, Conesa and Krueger (2009) consider a quantitative medium-scale overlapping-generations (OLG) model. They find that the optimal capital income tax amounts to 36% but is rather insensitive to borrowing constraints and idiosyncratic income risks.

Another early strand of the literature, e.g., Jones et al. (1997), has noted that the result of a zero optimal tax rate on capital income need not hold in models with human capital accumulation and innovation. Gross and Klein (2022) consider endogenous growth through innovation in a model based on Romer (1990) and find that the optimal capital tax rate is negative and that capital taxes should not be used to increase government revenues. The optimal tax rate on labor versus capital income is also sensitive to the accumulation technology of human capital. If the opportunity costs of accumulating human capital accrue in the form of time rather than goods, Grüner and Heer (2000) find an optimal tax rate on physical capital of 32% in the supply-side model of Lucas (1990).[4] Tsai et al. (2022) introduce equipment-specific

[4] For a survey on optimal capital income taxes, see Bastani and Waldenström (2020).

technological progress in the heterogeneous-agent Ramsey model with infinite lives and find that capital taxes should increase or decrease over time depending on the change in the technological growth rate. For instance, faster progress increases the depreciation of old equipment, so the real interest rate must increase. To internalize this pecuniary externality, the capital tax rate should be adjusted downward. In a model of semi-endogenous growth, Jones (2022) argues that the top income recipients provide a large part of the pool for new ideas as a result of their innovative activity. Higher tax rates on the top incomes slow the creation of new ideas and, hence, decrease economic growth and welfare. In his model, the optimal top income tax rate decreases from 65% (exogenous growth) to 9% (endogenous growth) if new ideas are created with skilled labor. Although his model abstracts from the use of capital in the research sector, his results clearly imply that capital gains that reflect a return to innovation should be taxed at a much lower rate.

Given that robots are part of capital inputs, the paper is also related to the recent surge in studies on the optimal tax on robots, including Guerreiro et al. (2022), Thuemmel (2022) and Prettner and Strulik (2020), among others. Guerreiro et al. (2022) find that it is optimal to decrease the tax on robots to zero in the long term. With technological progress that replaces routine workers with robots, all households should eventually switch to nonroutine jobs. However, in the short term, taxes should be used to redistribute money to older routine workers who can no longer build the necessary human capital to move into a nonroutine occupation. Thuemmel (2022) distinguishes three types of occupations: manual nonroutine, routine and cognitive nonroutine; however, unlike Guerreiro et al. (2022), households cannot change their occupational type. Robots substitute for routine labor, which is located in the medium range of the wage distribution. As automation is a cognitive-biased technology, robots compress the wage distribution and increase inequality at the bottom but benefit high income individuals. Thuemmel (2022) find that capital redistribution is optimal with the help of more progressive labor taxes rather than with the help of robot taxes. Only if labor-tax reform is infeasible do robot taxes help to improve welfare substantially. In this case, it is optimal to tax robots in the long term.

Although Guerreiro et al. (2022) and Thuemmel (2022) use a normative approach to capital (or robot) taxation, Prettner and Strulik (2020) use a positive approach, closer in spirit to the present study. Following Jones (1995), machine inputs into production are interpreted as robots that are invented by scientists in a research sector. Like the economic mechanism in Guerreiro et al. (2022), higher robot taxes lead to lower incentives to invest in education and human capital development. As a consequence, education and growth worsen. The general equilibrium effects of higher robot taxes may even be detrimental to wage inequality in the short and medium term, as a larger share of low-skilled individuals may depress their wages.

In sum, the literature has provided arguments to tax capital stock at differential rates depending on its nature (equipment versus structure as in Slavík and Yazici (2014)) and its effect on growth (research capital). In addition, the welfare results are sensitive with regard to the substitutability of capital. Although robots and automation have essentially substituted for routine and low-skill jobs in recent decades, AI might substitute for high-skilled workers in the coming decades. As a consequence, optimal capital tax rates will be sensitive with respect to the direction of technological change in the coming decades; therefore, our results should be interpreted and applied carefully, as we assume capital-skill complementarity and do not consider the replacement of skilled workers by AI.

1.2 Structure of the paper

The paper is structured as follows. Section 2 presents the medium-scale overlapping generations model with skilled and unskilled workers and a public social security system. Section 3 calibrates the model with the help of observations from the U.S. economy from 1980-2020. Section 4 analyzes a stripped-down version of the benchmark model to focus on the main economic mechanism and compares the optimal capital tax rate in models with finite and infinite lifetime. Section 5 presents our results. First, we show that the optimal steady-state capital income tax rate is 27% in the benchmark model and explain the underlying mechanisms. Second, we study the optimal redistribution policy and find that, in the presence of strong capital-skill complementarity, it is welfare-improving to use capital taxes rather than progressive labor income taxes to decrease income and wealth inequality. In Sect. 6, we discuss various extensions of the benchmark model. First, we demonstrate that our results are rather insensitive regarding the assumption of income uncertainty and borrowing constraints. Second, we show that the results of Slavík and Yazici (2014) to tax capital structures at a lower rate than capital equipment also extend to the model with a finite lifetime. Third, we consider the effect of aging. In this case, optimal capital taxes increase. A numerical analysis of the demographic transition demonstrates that the short- and medium-term effects of higher taxes imply significant losses for the generations currently alive. Section 7 concludes. Additional sensitivity and policy analyses together with more details on the modeling are delegated to the Appendix.

2 OLG model

In this section, we introduce the model. We consider an overlapping generations model and distinguish three sectors: households, a representative firm and the government. In addition, we compare two different types of production. As our benchmark, we consider the production technology of Krusell et al. (2000) that displays skill complementarity of capital. As an alternative scenario, we analyse the standard Cobb-Douglas technology where skilled and unskilled workers are perfect substitutes.

Our focus on the overlapping generations (OLG) model with finite lifetime, rather than the neoclassical growth model with infinite lifetime, is discussed in Sect. 4. There, we show that the sensitivity of the optimal capital tax rate to the production technology is substantially stronger in the former.[5]

[5] Specifically, assuming linear taxes, we find that accounting for capital-skill complementarity in production increases the optimal capital tax rate by 8.2 percentage points in the neoclassical growth model and by 21.4 percentage points in the OLG model (see Table 2).

2.1 Demographics

In every period $t \geq 0$, a new generation of households is born. The age of the newborns is $j = 1$. All of the generations retire at the end of age $j = T^W > 1$ and live up to a maximum age $j = T > T^W$. Population growth varies over time and is denoted as n_t. The total population in a given period *t* is N_t, and the number of households of age *j* is $N_t(j)$. Consequently, the share of the *j*-year-old cohort in the total population in period *t* is $N_t(j)/N_t$. All agents of age *j* survive until age $j+1$ with probability ϕ_t^j. Thus, $\phi_t^0 = 1$ and $\phi_t^T = 0$.

2.2 Households

Each household comprises one individual. Households are heterogeneous with respect to labor productivity. We distinguish two types of workers, skilled (S) and unskilled (U). The shares of the household types are denoted by ψ^i, $i \in \{S, U\}$, are constant across cohorts and sum to one: $\psi^S + \psi^U = 1$.

In every period *t*, newborn households with skills $i \in \{S, U\}$ maximize the expected intertemporal lifetime utility:

$$U_t^i = \sum_{j=1}^{T} \beta^{j-1} \left(\prod_{m=0}^{j-1} \phi_{t+m-1}^m \right) u(c_{t+j-1}^{i,j}, l_{t+j-1}^{i,j}), \tag{2.1}$$

where β, $c_t^{i,j}$ and $l_t^{i,j}$ denote the household's discount factor and consumption and labor supply at age *j* in period *t* for workers with skill $i \in \{S, U\}$.

Instantaneous utility *u*(.) is specified as in King et al. (1988):

$$u(c, l) = \frac{\left(c^{\gamma} (1-l)^{(1-\gamma)}\right)^{1-\eta} - 1}{1-\eta}, \tag{2.2}$$

where η is the inverse elasticity of intertemporal substitution, and γ denotes the weight of consumption in terms of utility.

The wage rate of the worker at age *j* and with skill type *i* in period *t*, $w_t^i \bar{y}^j A_t$, depends on household type $i \in \{S, U\}$, age *j* and aggregate labor productivity A_t. In particular, we assume that the age efficiency, $\bar{y}^j$, of the household is described by a hump-shaped function, as estimated by Hansen (1993). Accordingly, total labor income at age *j* for efficiency type *i* in period *t*, $y_t^{i,j}$, is the product of the wage rate per efficiency unit, w_t^i; the age-efficiency factor, $\bar{y}^j$; the aggregate labor productivity, A_t; and working hours, $l_t^{i,j}$:

$$y_t^{i,j} = w_t^i \bar{y}^j A_t l_t^{i,j}. \tag{2.3}$$

To model a progressive tax system for labor income, we follow Holter et al. (2019) and define a household's total tax burden as

$$T(y_t^{i,j}) = y_t^{i,j} - \theta_{0,t}(y_t^{i,j})^{1-\theta_{1,t}}, \tag{2.4}$$

which implies a marginal income tax of

$$\tau(y_t^{i,j}) = 1 - (1-\theta_{1,t})\,\theta_{0,t}(y_t^{i,j})^{-\theta_{1,t}}. \tag{2.5}$$

Parameter $\theta_{1,t}$ measures the degree of tax progressivity, whereas $\theta_{0,t}$ defines the tax level in period *t*. The tax system is progressive for $\theta_{1,t} \in (0,1)$, regressive for $\theta_{1,t} < 0$ and linear for $\theta_{1,t} = 0$.

Labor income is also subject to a social security contribution levied at rate τ_t^p.[6] Accordingly, net labor income $\hat{y}_t^{i,j}$ of *j*-year-old worker of type *i* in period *t* is presented as follows:

$$\hat{y}_t^{i,j} = \theta_{0,t}(y_t^{i,j})^{1-\theta_{1,t}} - \tau_t^p y_t^{i,j}. \tag{2.6}$$

The pension income received by the household upon retirement is denoted as pen_t^i and depends on skill type $i \in \{S, U\}$. Pension income is not subject to income taxation.

In any period *t*, the budget constraint of a household of age $j = 1, \ldots, T$ with skill type *i* is given by:

$$(1+\tau_t^c)c_t^{i,j} = \begin{cases} \hat{y}_t^{i,j} + [1+(1-\tau_t^k)r_t]a_t^{i,j} + tr_t - a_{t+1}^{i,j+1}, & j = 1, \ldots T^W \\ pen_t^i + [1+(1-\tau_t^k)r_t]a_t^{i,j} + tr_t - a_{t+1}^{i,j+1}, & j = T^W+1, \ldots, T \end{cases} \tag{2.7}$$

where τ_t^c is the tax rate on consumption, τ_t^k is the tax rate on capital income, $a_t^{i,j}$ denotes the stock of assets held by the *j*-year-old household with skills *i* at the beginning of period *t*, r_t is the rate of return on assets, and tr_t denotes nonpension-related transfers from the government received by the household in period *t*. As the household does not work during retirement, $l_t^{i,j} = 0$ for $j = T^W + 1, \ldots, T$ and $i \in \{S, U\}$. Furthermore, the household starts his life without assets and the households that survive until maximum age *T* do not leave bequests; thus, $a_t^{i,1} = a_t^{i,T+1} = 0$. Accidental bequests of those households that die prior to age *T* are confiscated by the government.

The household can hold two asset types, bonds $b_t^{i,j}$ and physical capital $k_t^{i,j}$, $a_t^{i,j} = b_t^{i,j} + k_t^{i,j}$. In equilibrium, the household is indifferent between holding assets in the form of either physical capital or government bonds since both yield the same (certain) after-tax return. With a single household living for two periods, the proportion of asset holdings would be the same at the household and aggregate levels, but with many periods, the portfolio allocation is indeterminate. Consequently, without loss of generality, we assume that each household holds the two assets in the same fixed proportions. This is determined in the aggregate as the share of capital in total

[6] We assume a flat-rate for all labor income, whereas the U.S. Social Security system specifies an earnings cap. De Nardi and Yang (2016) estimate the cap to amount to 2.47 times the average earnings.

assets, $K_t/(K_t + B_t)$, where K_t and B_t denote aggregate capital and government bonds, respectively.

In every period *t*, the maximization of lifetime utility in (2.1) subject to (2.3)-(2.7) yields equilibrium conditions for the optimal allocation of consumption, labor and assets given by:

$$u_{c,t}^{i,j} = \lambda_t^{i,j}(1 + \tau_t^c), \qquad j = 1, \ldots, T, \tag{2.8}$$

$$-u_{l,t}^{i,j} = \lambda_t^{i,j}\left(1 - \tau(y_t^{i,j}) - \tau_t^p\right) w_t^i A_t \bar{y}^j, \qquad j = 1, \ldots, T^W, \tag{2.9}$$

$$\lambda_t^{i,j} = \beta\phi_t^j \lambda_{t+1}^{i,j+1}[1 + (1 - \tau_{t+1}^k)r_{t+1}], \qquad j = 1, \ldots, T - 1, \tag{2.10}$$

where $u_{c,t}^{i,j}$ and $u_{l,t}^{i,j}$ denote the derivatives in period *t* of the household's utility at age *j* with skill level $i \in \{S, U\}$ with respect to consumption and labor, respectively, whereas $\lambda_t^{i,j}$ is the Lagrange multiplier of the budget constraint in equation (2.7).

2.3 Production

We consider two different production technologies. As our benchmark, we analyze capital–skill complementarities. As an alternative scenario, we consider the model with the standard Cobb–Douglas production technology.

2.3.1 Benchmark case: capital–skill complementarity

As our benchmark case, we use the nested CES production function suggested by Krusell et al. (2000):

$$Y_t = \left[\mu\left(A_t L_t^U\right)^\sigma + (1 - \mu)\left(\alpha K_t^\rho + (1 - \alpha)\left(A_t L_t^S\right)^\rho\right)^{\frac{\sigma}{\rho}}\right]^{1/\sigma}, \tag{2.11}$$

where $\frac{1}{1-\sigma}$ denotes the elasticity of substitution between capital K_t and unskilled labor L_t^U, and $\frac{1}{1-\rho}$ equals the elasticity of substitution between capital K_t and skilled labor L_t^S. If $\sigma > \rho$, capital is more complementary to skilled labor than it is to unskilled labor, and the technology features capital-skill complementarity. Technological progress A_t is labor-augmenting and grows over time at exogenous rate $g \geq 0$, which is equal to the balanced growth rate of the economy.

Aggregate high- and low-skilled labor are simply the sums of the efficient labor supplies of the two household types:

$$L_t^i = \sum_{j=1}^{T^w} \psi^i N_t(j) \bar{y}^j l_t^{i,j}, \quad i \in \{S, U\}. \tag{2.12}$$

Capital depreciates at the rate δ.

Firms maximize profits

$$\Pi_t = Y_t - r_t K_t - w_t^U A_t L_t^U - w_t^S A_t L_t^S - \delta K_t \tag{2.13}$$

resulting in the first-order conditions

$$\begin{aligned} r_t + \delta = \alpha(1-\mu) & \left[\mu \left(A_t L_t^U\right)^{\sigma} + (1-\mu)\left(\alpha K_t^{\rho} + (1-\alpha)\left(A_t L_t^S\right)^{\rho}\right)^{\frac{\sigma}{\rho}}\right]^{1/\sigma-1} \\ & \times \left(\alpha K_t^{\rho} + (1-\alpha)\left(A_t L_t^S\right)^{\rho}\right)^{\frac{\sigma}{\rho}-1} K_t^{\rho-1}, \end{aligned} \tag{2.14a}$$

$$w_t^U = \mu \left[\mu \left(A_t L_t^U\right)^{\sigma} + (1-\mu)\left(\alpha K_t^{\rho} + (1-\alpha)\left(A_t L_t^S\right)^{\rho}\right)^{\frac{\sigma}{\rho}}\right]^{1/\sigma-1} \left(A_t L_t^U\right)^{\sigma-1}, \tag{2.14b}$$

$$\begin{aligned} w_t^S = (1-\alpha)(1-\mu) & \left[\mu \left(A_t L_t^U\right)^{\sigma} + (1-\mu)\left(\alpha K_t^{\rho} + (1-\alpha)\left(A_t L_t^S\right)^{\rho}\right)^{\frac{\sigma}{\rho}}\right]^{1/\sigma-1} \\ & \times \left(\alpha K_t^{\rho} + (1-\alpha)\left(A_t L_t^S\right)^{\rho}\right)^{\frac{\sigma}{\rho}-1} \left(A_t L_t^S\right)^{\rho-1}, \end{aligned} \tag{2.14c}$$

Profits are equal to zero in equilibrium.

2.3.2 Comparison case: Cobb–Douglas production function

As a comparison case, we consider the standard Cobb–Douglas production function with labor-augmenting technological progress A_t:

$$Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}. \tag{2.15}$$

In this case, we assume that skilled and unskilled workers have permanent productivity ϵ^i, $i \in \{S, U\}$, so the effective labor of the worker at age *j* with skill *i* in period *t* is given by $\bar{y}^j \epsilon^i l_t^{i,j}$. This modeling of individual productivity allows us to reproduce the skill premium because the relative wages of the skilled and unskilled workers at age *j* is equal to ϵ^S/ϵ^U. Accordingly, the aggregate labor in efficiency units L_t is given by:

$$L_t = \sum_{i \in \{S,U\}} \sum_{j=1}^{T^w} \psi^i N_t(j) \bar{y}^j \epsilon^i l_t^{i,j}. \tag{2.16}$$

Since production is perfectly competitive, in equilibrium, labor and capital are remunerated at their marginal products, that is:

$$w_t = (1-\alpha)K_t^{\alpha}(A_t L_t)^{-\alpha}, \tag{2.17a}$$

$$r_t = \alpha K_t^{\alpha-1}(A_t L_t)^{1-\alpha} - \delta, \tag{2.17b}$$

and the wages of skilled and unskilled workers, w_t^i, $i \in \{S, U\}$, in (2.3) are presented by $w_t^i = w_t \epsilon^i$.

2.4 Government

Government expenditure is allocated to public consumption G_t, aggregate transfers Tr_t and interest payments on public debt $r_t^b B_t$, with r_t^b denoting the return on government bonds. Government revenue is increased through taxation Tax_t, new debt issuance $B_{t+1} - B_t$ and the confiscation of accidental bequests Beq_t.[7] In total, the government budget constraint is as follows:

$$G_t + Tr_t + r_t^b B_t = Tax_t + B_{t+1} - B_t + Beq_t, \tag{2.18}$$

where tax revenue from the taxation of aggregate consumption C_t, capital K_t and labor L_t is given by:

$$Tax_t = \tau_t^c C_t + \tau_t^k r_t K_t + \sum_{j=1}^{T^W} \sum_{i \in \{S,U\}} \psi^i T(y_t^{i,j}) N_t(j). \tag{2.19}$$

The government sector also includes a pay-as-you-go pension system. In the aggregate, the pension expenditure is the sum of the pension payments made to retired households:

$$Pen_t = \sum_{j=T^W+1}^{T} \sum_{i \in \{S,U\}} \psi^i pen_t^i N_t(j), \tag{2.20}$$

and the social security budget is formulated as:

$$Pen_t = \tau_t^p \left(w_t^S A_t L_t^S + w_t^U A_t L_t^U \right). \tag{2.21}$$

2.5 Equilibrium and aggregate conditions

At the aggregate level, consumption, assets, transfers and accidental bequests are determined as the sum of the corresponding individual variables; thus:

[7] As we neglected any bequest motive in the specification of the utility function, the assumption of the confiscation of Beq_t by the government is rather innocuous. Accidental bequests amount to only 0.1% of total wealth in our benchmark economy. Alternatively, one could introduce a parent–child link as in Heer (2001) or De Nardi and Yang (2016).

$$C_t = \sum_{j=1}^{T} \sum_{i \in \{S,U\}} \psi^i N_t(j) c_t^{i,j}, \tag{2.22}$$

$$\mathcal{A}_t = \sum_{j=1}^{T} \sum_{i \in \{S,U\}} \psi^i N_t(j) a_t^{i,j}, \tag{2.23}$$

$$Tr_t = \sum_{j=1}^{T} \sum_{i \in \{S,U\}} \psi^i N_t(j) tr_t, \tag{2.24}$$

$$Beq_{t+1} = \sum_{j=1}^{T-1} \sum_{i \in \{S,U\}} \psi^i N_t(j)(1-\phi_t^j)[1+(1-\tau_{t+1}^k)r_{t+1}]a_{t+1}^{i,j+1}. \tag{2.25}$$

Equilibrium in the goods market requires that aggregate output is equal to aggregate demand:

$$Y_t = C_t + G_t + K_{t+1} - (1-\delta)K_t, \tag{2.26}$$

whereas equilibrium in the capital market requires that the aggregate assets purchased by the households are equal to the sum of the aggregate capital and government bonds demanded by firms and the government, respectively:

$$\mathcal{A}_t = K_t + B_t. \tag{2.27}$$

The no-arbitrage condition implies that, in equilibrium, all assets have the same after-tax rate of return:

$$r_t^b = (1-\tau_t^k)r_t. \tag{2.28}$$

3 Calibration

The calibration is designed to approximate a steady state for the U.S. economy on the basis of data averages from 1980-2020. Table 1 summarizes our parameter choices.[8]

Demographics Period t in the model corresponds to one year. Newborns are assumed to have a real-life age of 20, corresponding to $j=1$, and to live up to a real-life age of 99; thus, T=80. Retirement ages $T^W+1=48$ (corresponding to the real-life age of 67) are calculated as the average of the effective age of retirement for 2014, 2018 and 2020 obtained from OECD (2021). Population growth rates n are

[8] The GAUSS computer code for the calibration and all applications is available from the author upon request.

Table 1 Calibration of parameters for the U.S. economy during 1980–2020

Parameter	Value	Description
Household type		
(ψ^S, ψ^U)	(0.5, 0.5)	share of household type $i \in \{S, U\}$
Preference		
$1/\eta$	1/2	intertemporal elasticity of substitution
γ	0.226	utility weight of consumption
β	1.010	discount factor
Production		
$1/(1-\sigma)$	1.67	elasticity of substitution between capital and unskilled labor
$1/(1-\rho)$	0.67	elasticity of substitution between capital and skilled labor
(α, μ)	(0.657, 0.261)	weight of production factors in (2.11)
δ	8.3%	depreciation rate of capital
g	1.529%	labor productivity growth rate
Government		
$(\tilde{\theta}_0, \theta_1)$	(0.757, 0.137)	income tax parameters
τ^k	22.8%	tax on capital income
τ^c	6.2%	tax on consumption
G/Y	19.4%	share of government spending in steady-state production
B/Y	72.8%	debt-output ratio
$(repl^S, repl^U)$	(32.6%,40.6%)	gross pension replacement rate of (un)skilled workers

calculated as the averages of the annual population growth rates from 1980-2020, obtained from United Nations (2022). Survival probabilities ϕ_t^j are computed from annual data on life expectancy for both sexes combined (from age 20 to age 99), which are also obtained from United Nations (2022). The calibration employs the average survival probabilities from 1980-2020.

Household Types We set the shares of skilled and unskilled households equal to $\psi^S = \psi^U = 0.5$, corresponding to the average number of college graduates during the period 1980-2020 reported by the U.S. Bureau of Labor Statistics.[9] Age–productivity file $\{\bar{y}^j\}_{j=1}^{47}$ is taken from Hansen (1993) and extrapolated linearly to the real-life age of 66.

Preferences We choose an intertemporal elasticity of substitution, $1/\eta$, equal to 1/2. We set the weight of utility from consumption, $\gamma = 0.226$, to imply an average number of working hours of 0.25 in this case. Discount factor $\beta = 1.010$ implies a real interest rate of 4.0%.[10]

[9] The data are taken from Table A-1 of the Historical Time Series Tables. The average share of those with a college education relative to the total number of households for 1980–2020 was 0.4863.

[10] Hurd (1989) provides an empirical estimate of 1.011 for the discount factor after accounting for stochastic survival.

Production The average growth rate of per capita GDP from 1980–2020 is $g = 1.529\%$. This is calculated using data obtained from OECD (2022) by scaling the annual nominal value of the gross domestic product at market prices by the corresponding deflator and then dividing by the total population. Capital depreciates at the annual rate $\delta = 8.3\%$.

In the calibration of the parameters in production function (2.11), we follow Angelopoulos et al. (2015) and set $\sigma = 0.401$ and $\rho = -0.495$ as estimated by Krusell et al. (2000).[11] The values of σ and ρ imply a substitution elasticity between capital and unskilled labor equal to $1/(1-\sigma) = 1.67$ and a substitution elasticity between capital and skilled labor equal to $1/(1-\rho) = 0.67$. The remaining production parameters, α and μ, are calibrated to satisfy the following conditions simultaneously:

1. The skill premium is set so the skilled workers receive a skill premium of 150% above the wage of the unskilled workers[12]:

$$\frac{w^S}{w^U} = 2.50.$$

2. Total wages are equal to 65% of GDP:

$$\frac{w_t^U L_t^U + w_t^S L_t^S}{Y_t} = 0.65.$$

We find that $\alpha = 0.657$, $\mu = 0.261$.

In the case of the Cobb–Douglas production function (2.15), we set $\alpha = 0.35$ and $(\epsilon^S, \epsilon^U) = (1.43, 0.57)$ to imply the same skill premium and aggregate labor share as in the case of production function (2.11).[13]

Government The government parameters are also chosen as averages from empirical data from 1980–2020. In particular, we follow Brinca et al. (2016) and set tax progressivity $\theta_1 = 0.137$ to replicate the average tax progressivity in the United States.[14] The average labor income taxes (both income tax and social security) are 21.7%, implying that $\tilde{\theta}_0 = \theta_{0,t}/A_t = 0.757$, whereas the capital income tax and the consumption tax are equal to $\tau^k = 22.8\%$ and $\tau^c = 6.2\%$, respectively.[15]

[11] Krusell et al. (2000) estimated the parameters σ and ρ using data for the period 1963-1992. Maliar et al. (2022) extended the observation period to 1963-2017 and estimated nearly identical values for the parameters σ and ρ. For example, the authors find $\sigma = 0.415$.

[12] A skill premium of 150% is in accordance with data from the U.S. Bureau of Labor Statistics on the median weekly earnings of workers with a bachelor's degree or higher (LEU0252918500) relative to those with less than a high school diploma (LEU0252916700). Krueger and Ludwig (2007) apply the same skill premium.

[13] We must also recalibrate the parameters $(\beta, \gamma, \tilde{\theta}_0) = (1.0095, 0.211, 0.804)$ that are set with the help of the steady-state conditions of the model (not presented in Table 1).

[14] Heathcote et al. (2017) apply a slightly higher value of $\theta_1 = 0.18$ because they include capital income as part of taxable income.

[15] We calculate the average tax rates on income from labor, capital and consumption using the revision of the method of Mendoza et al. (1994) proposed by Carey and Rabesona (2003).

Government consumption G is 19.4% of GDP, whereas the debt–GDP ratio during this period averaged 72.8%. To calibrate the replacement rate of pensions pen_t^i with respect to the average gross income of worker type $i \in \{S, U\}$, $\overline{(w_t^i A_t \bar{y}^s l^{i,j})}$, we use linear interpolation of the replacement rate with respect to gross income provided by OECD (2021) for men with 50% and 200% of the average wage, 49.4% and 27.8%, respectively, implying replacement rates $repl^S = 32.6\%$ and $repl^U = 40.6\%$.

In equilibrium, the social security tax τ^p must be set equal to 7.51% to balance the budget of the social security authority. The share of capital (government bonds) in total assets amounts to 79.6% (20.4%).

4 Optimal capital taxes with finite and infinite lifetimes

In this section, we study the welfare effects of capital income taxes in the two prototype models, the neoclassical growth model and the OLG model, and demonstrate that the type of technology, skill-complementarity versus Cobb-Douglas production, has a much stronger effect in the finite than in the infinite lifetime model.

We derive the optimal capital taxes following Conesa and Krueger (2009). In particular, we evaluate numerically how a change in the capital income tax rate affects the steady-state utility of households whereas government expenditures (transfers and consumption) are kept constant. As our welfare criterion, we use average lifetime utility U^i in steady state so that our measure of welfare W is presented by

$$W = \sum_{i \in \{S,U\}} \psi^i U^i. \tag{4.1}$$

In the case of the OLG model, lifetime utility U^i is presented by (2.1); in the case of the neoclassical growth model with infinite horizon, lifetime utility is also presented by (2.1) with $T \to \infty$ and $\phi^j = 1$ for all $j \in \mathbb{N}_0$.

Notice that our welfare analysis differs from the consideration of the Ramsey policy in Chamley (1986) and Judd (1985). One imminent advantage of our optimal tax analysis is the avoidance of one of the unappealing properties of the optimal Ramsey tax policy. In particular, the Ramsey policy assumes a balanced intertemporal but not a balanced budget in every period. Therefore, the optimal policy might be characterized by the accumulation of large assets by the government during the initial phase of the transition. Given the empirical evidence on debt levels in most OECD countries and, in particular, the USA, the accumulation of public assets (negative debt) may be politically infeasible. In addition, some recent theoretical arguments against the results of Chamley-Judd are provided by Straub and Werning (2020). These authors have proven that the assumption of an interior solution does not hold in Judd (1985) for the realistic case of an intertemporal elasticity of substitution below one and that, in this case, the optimal long-term tax on capital is positive. For elasticities above unity, convergence to the steady state with optimal zero capital taxes may take centuries.[16]

[16] We find a similar behavior of the dynamics in Sect. 6.

We start our analysis from the most simple neoclassical growth model with a representative household developed by Ramsey (1928); Cass (1965) and Koopmans (1965) but augmented by a government sector.[17] The household chooses its optimal labor supply and consumption intertemporally and is subject to linear taxes on consumption and labor and capital income. Competitive firms hire labor and rent capital to produce output with a Cobb-Douglas production technology. The government uses the tax revenues to finance exogenous government consumption and transfers so that the fiscal budget is balanced. The production and utility parameters of the model are calibrated in accordance with those presented in Table 1 with the exception of the linear labor income tax rate that is set to $\tau^l = 21.7\%$.

The optimal positive capital income tax τ^k in this basic neoclassical growth model, augmented by a government sector and linear taxes, is equal to 0% and it is not optimal to tax capital income.[18] The results are presented as 'Case 1a' in Table 2. For $\tau^k = 0$, the labor income tax rate τ^l must be increased from 21.7% to 25.0% to balance the fiscal budget.

In the first step, we add the heterogeneity element in the form of two worker types, skilled and unskilled, to the infinite-lifetime model. The resulting model is labeled 'Case 4a' in Table 2. We choose a constant skill premium of 150% and equal size for the two group types. In our calibration with respect to the U.S. distribution of wealth, we identify the wealth of the unskilled (skilled) workers with that of the lower (upper) half of the wealth distribution. Higher capital taxes are predominantly imposed on the returns of the wealth of the skilled workers, whereas the unskilled workers pay barely any capital income taxes. In this environment, a welfare-increasing effect of capital taxation arises because income is redistributed from the rich skilled workers to the poor unskilled workers, and the optimal capital tax rate τ^k increases to 18.0%.[19]

In a second step, we assume capital-skill complementarity of production rather than a Cobb-Douglas technology. Again, the optimal capital tax rate increases significantly, from 18.0% to 26.1% ('Case 5a' in Table 2). The consideration of capital-skill complementarity introduces an additional welfare-enhancing effect of capital taxation. With higher capital taxes, savings and, hence, capital decrease. As a conse-

Table 2 Model types and optimal capital income taxes $\tau^k \geq 0$

Case	a) Infinite life	b) OLG model
1. Representative agent/cohort	0.0%	9.1%
2. + Age-dependent productivities		7.5%
3. + PAYG pensions		12.7%
4. Two worker types	18.0%	27.7%
5. + Capital-skill complementarity	26.1%	49.1%

[17] The model and the extended variant that includes different skill types of the households is described in more detail in Appendix A.

[18] We restrict the range of optimal capital taxes to the interval $[0, 100\%]$. The solution for Model 1a is a corner solution, and it would even be optimal to subsidize capital in this economy.

[19] The optimal capital tax of 18.0% is sensitive with respect to the distribution of wealth. If we exclude the top percentile of the wealth distribution (which holds approximately one third of total wealth) from our calibration of the respective wealth levels of the skilled and unskilled workers, the optimal capital tax rate decreases to 16.5%.

quence, the wages of the skilled workers decrease relative to those of the unskilled workers, redistributing income from the high-income individuals to the low-income individuals.[20] As a consequence, the skill premium decreases, as does inequality in the economy, which increases average lifetime utility.

In the far-right column of Table 2, we report the optimal capital tax rates in the corresponding models with finite lifetimes. In the most basic version of the life-cycle model ('Case 1b' in Table 2), all of the agents in a cohort are identical and do not receive a pension in old age. Thus, the model is a simplified version of our benchmark model in Sect. 2, assuming 1) linear labor income taxes, $\theta_1 = 0$, 2) no pensions, 3) no government debt, $B/Y = 0$, 4) no growth, $g = 0$, 5) a representative cohort and 6) a Cobb-Douglas production technology.

The introduction of a finite lifetime is sufficient to establish a welfare effect of capital taxation, as noted by Erosa and Gervais (2002), and the optimal capital tax amounts to 9.1%. The addition of life-cycle elements in the form of age-dependent productivities (Case 2b) decreases optimal capital taxes to 7.5%. Age productivity is hump shaped and much greater in old age than it is in young age. Therefore, the labor supply decreases less strongly over the lifetime, and, as argued by Erosa and Gervais (2002), the welfare effect of capital taxes stemming from a higher discounted price of leisure in old age decreases. A public pay-as-you-go pension (Case 3b) increases the optimal capital tax rate to 12.7%. In our model, social security taxes are imposed on labor income. Therefore, the effect of the tax wedge on labor income increases in the life-cycle economy with PAYG pensions. Because distortions from labor income taxes increase nonlinearly with the tax rate, a decrease in the labor income tax rate has a stronger effect in the economy with PAYG pensions (from a high tax rate) than it does in an economy without pensions (from a low tax rate).[21]

The introduction of skill heterogeneity increases optimal capital tax rates by 15 percentage points, from 12.7% (Case 3b) to 27.7% (Case 4b). In comparison with the infinite-lifetime model, the quantitative effects of heterogeneity are comparable even though we consider a more equal wealth distribution in the OLG model.[22] When we switch from a Cobb-Douglas production function to one with capital-skill complementarity (Case 5b), optimal capital taxes increase by more than 20 percentage points, from 27.7% to 49.1%. Evidently, the increase in optimal capital taxes is much stronger quantitatively in the life cycle model than it is in the infinite-lifetime model. Therefore, we concentrate our welfare analysis on the study of the life-cycle model and use the more elaborate medium-scale OLG model described in Sect. 2.

[20] The lower labor supply of the skilled workers relative to that of the unskilled workers exhibits an opposite effect on relative wages, which, in our case, is significantly smaller than the effect stemming from the decrease in capital stock.

[21] Public pensions also increase the magnitude of the opposing negative welfare effect on capital accumulation that results from higher capital taxes. In a model with PAYG pensions, savings are lower, causing a stronger underaccumulation of capital stock relative to the optimal (golden rule) level. Again, distortions increase nonlinearly with the distance to the optimal level of the capital stock. (The model is calibrated to be dynamically efficient for all tax policies considered.)

[22] In the OLG model, the distribution of wealth is endogenous and cannot be imposed. The ratio of average wealth among the skilled relative to that among the unskilled amounts to 2.83.

5 Results

In this section, we present our main results. First, we show that the optimal capital income tax rate is substantially higher when production exhibits capital–skill complementarity in production than under the standard Cobb–Douglas production function. Second, we examine the optimal progressivity of labor income taxes and find that redistribution is optimally achieved through capital income taxation rather than through progressive labor income taxes. Extensions of the model to incorporate income uncertainty, capital heterogeneity and transition dynamics are discussed in the next section, while the sensitivity analysis with respect to the model parameterization is presented in Appendix C.

5.1 Optimal capital taxes

Figure 1 displays the effects of capital income taxes on welfare for the benchmark calibration with capital–skill complementarity in production (solid blue line 'Krusell et al.') and the standard Cobb–Douglas technology (broken black line 'Cobb–Douglas'). As our welfare criterion, we employ the ex ante expected lifetime utility of the newborn in the steady state as presented in (4.1). Thus, the unborn individual does not know her ability type $i \in \{S, U\}$ and forms expectations with respect to her stochastic survival. Welfare changes relative to the benchmark with $\tau^k = 22.8\%$ are expressed in consumption equivalent changes. The optimal capital income tax rate is 27.3% in the model with capital–skill complementarity in production technology, as in Krusell et al. (2000), whereas it decreases to 9.2% in the case of the standard Cobb–Douglas technology.[23] The associated welfare gains needed to move from the present capital tax rate of $\tau^k = 22.8\%$ to the new optimal tax rates amount to 0.01%

[23] Conesa and Krueger (2009) find an optimal capital income tax rate of 36% for the Cobb–Douglas technology (as presented in the 'Bench' scenario in Table 4). We replicate their result in a modified version and calibration of our benchmark model, which is described in the Sect. 6.1. The main reason why these authors derive a much higher optimal capital income tax rate in comparison with our benchmark calibration is their assumption of an economic growth rate of zero.

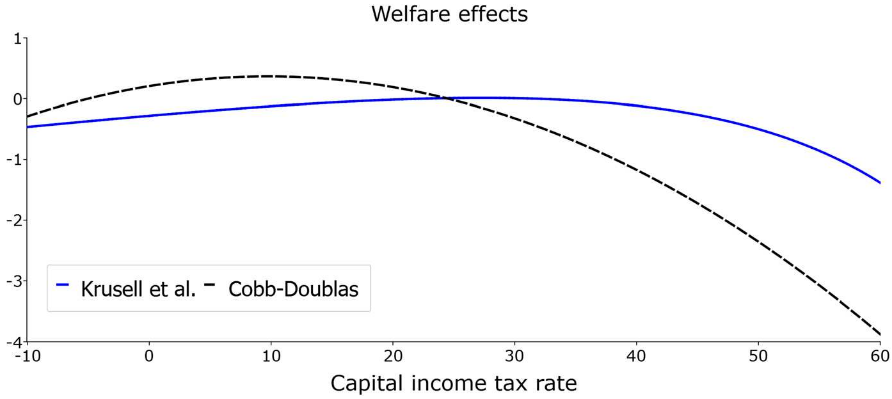


Fig. 1 Welfare effects of capital income taxes τ^k (in percentage points)

Table 3 Optimal tax rate $\tau^{k,*}$

Case	Optimal	Welfare effects	
	$\tau^{k,*}$	$\tau^{k,*}$	$\tau^{k}=0$
Benchmark	27.3%	0.01%	−0.28%
Cobb–Douglas technology	9.2%	0.25%	0.13%
Optimal income tax progressivity			
– $\theta_1 = -0.006\%$	37.9%	1.69%	1.01%
Income uncertainty			
– Krusell et al. technology, $g = 1.529\%$, $\eta = 1$	23.1%	0.12%	−1.01%
– Krusell et al. technology, $g = 0$, $\eta = 4$	52.0%	0.90%	−1.03%
– Cobb–Douglas, $g = 0$, $\eta = 4$	34.7%	0.65%	−3.31%
Two types of capital			
(τ^{ks}, τ^{ke})	(59.3%,9.2%)	0.70%	−0.60%
Population in 2050	35.2%	0.10%	−0.43%

Table 4 The effects of abolishing capital income taxes τ^k on factor supplies (percentage changes)

Variable	Krusell et al	Cobb–Douglas
K	7.02	11.39
$\bar{l}^S$	0.20	−0.33
$\bar{l}^U$	−0.14	−0.33
L^S	0.04	−0.42
L^U	−0.40	−0.41

(Krusell et al.) and 0.25% (Cobb-Douglas technology) of total consumption, respectively. If capital income taxes are reduced to zero, as suggested by the results of Chamley (1986) and Judd (1985), welfare decreases by 0.28% in the case of capital–skill complementarity and increases by 0.13% in the Cobb–Douglas case. Table 3 summarizes the optimal capital income taxes and the welfare gains from changing the present capital income tax of 22.8% to the optimal and zero tax rates (expressed in consumption equivalent changes) for various scenarios.

To understand these welfare effects, we must consider the different behaviors of the factor prices for the two production technologies. In accordance with Conesa and Krueger (2009), we find that a change in the capital income tax, τ^k, which is financed by an adjustment of the labor income tax, has a strong effect on the aggregate capital stock but a negligible effect on average working hours and the aggregate labor supply. Table 4 presents the effects of abolishing capital income tax τ^k on aggregate factor supplies. In both technology cases, the changes in the capital stock are 10–30 times greater than are the changes in the aggregate labor supplies. For example, in the case of skill-capital complementarity, the capital stock increases by 7.02% if capital taxes τ^k decrease from 22.8% to 0%, whereas the average working hours of the skilled and unskilled workers change by 0.20% and −0.14%, respectively.[24] There-

[24] The small quantitative response of labor is not surprising. The Frisch labor supply elasticity amounts to $\frac{1-l}{l}\frac{1-\gamma(1-\eta)}{\eta}$ and averages 3.68 in our model. However, Frisch elasticity isolates the substitution effect. The uncompensated labor supply elasticity for the Cobb–Douglas utility in the simple one-period model is zero if nonlabor income is zero and becomes negative in the presence of positive nonlabor income,

fore, to understand the effects of technology, we first need to analyze how a change in the capital stock affects factor prices and, subsequently, study how these factor price changes affect individual behavior.

Figure 2 presents the effects of capital stock on the real interest rate for the constant labor inputs. The capital stock in the steady state with capital income tax rate $\tau^k = 22.8\%$ is normalized to one in both the case with skill–capital complementarity ('Krusell et al.') and the case with Cobb–Douglas technology ('Cobb–Douglas'). In the steady state, the real interest rate before taxes (corresponding to $r = \partial Y/\partial K - \delta$ in the model) amounts to 4.0% with both technologies. We consider a change in the capital stock in the range $\pm 20\%$ and note that the associated changes in the real interest rate are approximately 40% greater under capital–skill complementarity than they are under Cobb–Douglas technology. For example, for a decrease in the capital stock of 20%, interest rates increase to 6.66% and 5.92%. This result reflects our calibration of the production function (2.11) with a substitution elasticity between capital K_t and the labor supply of unskilled labor, L_t^U, above one, $1/(1-\sigma) = 1.67$, which is amplified by a substitution elasticity between capital stock and skilled labor below one, $1/(1-\rho) = 0.67$.

Figure 3 presents the life-cycle profiles of labor supply for the skilled and unskilled workers for technology with capital–skill complementarity and for Cobb–Douglas technology. The figure compares the case of zero capital taxation (broken black line) with the benchmark, $\tau^k = 22.8\%$ (solid blue line). In all of the cases, the labor-supply profiles become flatter as the capital tax increases from 0% to 22.8%; therefore, the after-tax interest rate decreases.[25] The quantitative response of the labor supply is much more pronounced in the case of the skill complementarity of capital than it

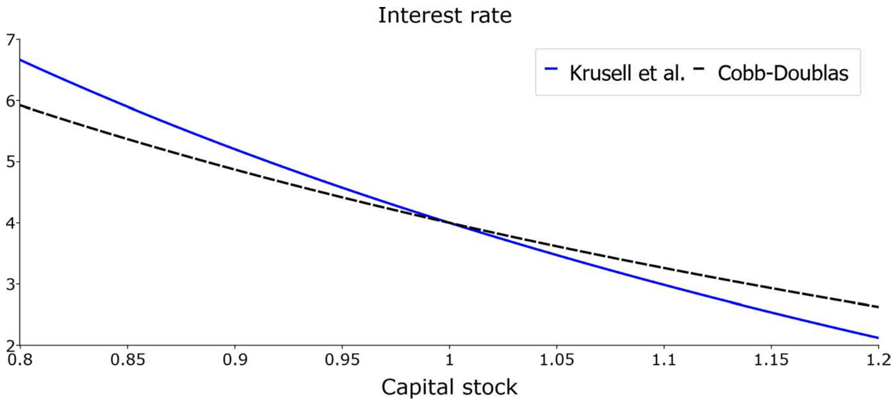
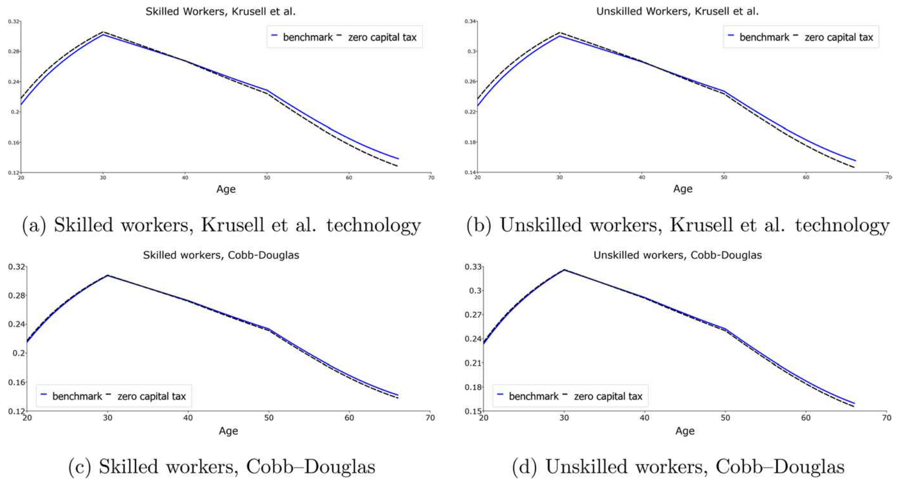


Fig. 2 Capital stock and interest rates

such as transfers or capital income. As the income effect almost offsets or even overcompensates for the substitution effect in our model, aggregate and average labor do not change much in response to a change in net wages.

[25] The labor-supply age profile in the competitive economy differs from that in the first-best allocation. Unlike progressive labor income taxes, capital income taxes may help to bring these two labor-supply age profiles in closer accordance. Appendix B shows that, in the steady state of the first-best allocation, the growth factor of leisure, $(1-l^{i,j+1})/(1-l^{i,j})$, is a function of the relative age-dependent productivi-



(a) Skilled workers, Krusell et al. technology (b) Unskilled workers, Krusell et al. technology

(c) Skilled workers, Cobb–Douglas (d) Unskilled workers, Cobb–Douglas

Fig. 3 Life-cycle profiles of working hours and capital income taxes

is in the case of the Cobb–Douglas technology, by a factor of approximately 3. For example, the working hours of the 20-year-olds decrease by 3.82% for higher taxes in the 'Krusell et al.' case but by only 1.01% in the 'Cobb–Douglas' case. Similarly, the working hours of the 65-year-olds increase by 7.40% and 2.76% in the two cases. The stronger response in the 'Krusell et al.' case follows directly from the stronger decrease in the interest rate depicted in Fig. 2. For a lower interest rate, the household reallocates labor intertemporally and increases labor in old age.[26]

As argued in detail by Erosa and Gervais (2002) in a theoretical analysis and by Conesa and Krueger (2009) in a quantitative analysis, the intertemporal reallocation of labor and leisure is the main explanation for the welfare-increasing effects of a (strictly) positive optimal capital income tax rate. The formal argument has also been advocated by Erosa and Gervais (2002). If the government could use age-dependent labor income taxes, it would be efficient in the present model with a decreasing labor supply over the life cycle to impose labor taxes that also decrease with age.[27] If, as in

ties, $\bar{y}^{j+1}/\bar{y}^j$, and factor $\beta\phi^s(1+F_K-\delta)$, where F_K denotes the marginal product of capital. Therefore, leisure growth in the first-best steady state does not depend on skill type $i \in \{S, U\}$. As j increases and, hence, ϕ^j decreases, the growth factor of leisure decreases and even falls below one for a certain threshold of ϕ^j. The resulting labor-supply profile in the first-best allocation decreases with age (even at a young age) but might increase moderately during the final years of the working life depending on the model calibration.

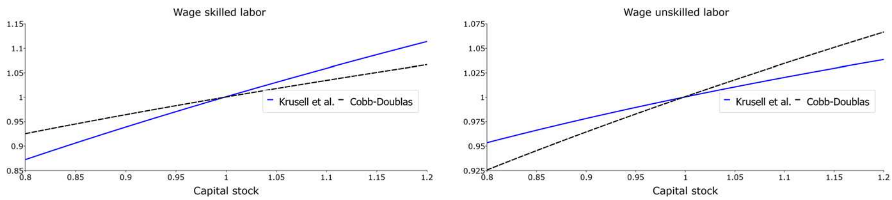
[26] İmrohoroğlu and Kitao (2009) consider a related phenomenon. In a life-cycle model similar to ours, they consider different social security reforms and their effects on the intertemporal allocation of labor. They find that the privatization of social security has a strong effect on capital but a negligible effect on aggregate labor; importantly, individuals shift their labor supplies from younger to older years as the real interest rate falls.

[27] To see this point intuitively, consider the *Ramsey* rule emanating from the classic work of Ramsey (1927). In our model, the Frisch labor supply elasticity initially decreases with age until the age of 30 and subsequently increases over the life cycle, so, ceteris paribus, the optimal labor income tax should decrease with age (beyond the threshold of 30 years). In a quantitative study on optimal fiscal policy, Guerreiro et

the present model, labor income rates cannot be conditioned on age, the government can imitate this age-dependent tax policy by taxing capital income and increasing the net present value of older workers' labor income so the opportunity costs of leisure in old age increase.

Next, we analyze how a change in the capital stock (while keeping the labor supplies constant) affects the wages of unskilled and skilled workers for the two different technologies. Figure 4 presents the wages of skilled and unskilled workers for the 'Krusell et al.' technology (blue solid line) and the 'Cobb–Douglas' technology (black broken line). All of the wages are normalized to one in the steady state for ease of comparison. Evidently, the wages of unskilled labor react less strongly to a change in the aggregate capital stock K in the case of skill–capital complementarity than they do in the case of the Cobb–Douglas technology, and vice versa for the wages of skilled workers. This observation follows directly from the production technology of Krusell et al. (2000), where higher σ and lower ρ amplify the increase in inequality between the skilled and the unskilled workers as capital deepening increases. Therefore, the skill premium decreases with lower capital stock in the case of skill–capital complementarity, whereas it remains constant in the Cobb–Douglas case. For example, the skill premium decreases from 150% to 128% with 'Krusell et al.' technology if the capital stock decreases by 20%.

In the general equilibrium model with endogenous capital and labor, high capital income taxes decrease savings and, hence, aggregate capital stock, whereas labor remains rather constant; therefore, wage inequality decreases in the case of 'Krusell et al.' technology but not in the case of Cobb–Douglas technology.[28] As a consequence, the decreasing skill premium helps improve the lifetime utility of the unskilled workers compared with the skilled workers in the case of 'Krusell et al.' technology. Figure 5 presents the lifetime utilities U_t^i of the skilled and the unskilled



(a) Skilled labor wage w^S (b) Unskilled labor wage w^U

Fig. 4 Aggregate capital stock K_t and wages w^i, $i \in \{S, U\}$

workers as a function of the capital income tax rate, τ^k, for both technology types.

al. (2022) derive a hump-shaped labor-tax-age profile in the standard overlapping generations model with a representative cohort and Cobb-Douglas utility, as in our model. Optimal labor income taxes decrease from 58% at age 28 to 32% at the end of the working life.

[28] In addition to the decrease in wage inequality, higher capital taxes τ^k decrease the labor income tax level, θ_0. As a consequence, the net wage rate of the unskilled workers (not presented) increases over the whole working life even though gross wages decrease with higher capital taxes and, hence, lower capital stock, as illustrated in Fig. 4.

For the Cobb–Douglas technology, the lifetime utilities of both the unskilled workers and the skilled workers are concave functions of the capital income tax rate and peak around a tax rate of 10%. The lifetime utilities of workers behave fundamentally differently in the case of technology with skill–capital complementarity. In this case, the lifetime utility of the unskilled worker increases with τ^k over the entire range considered, $\tau^k \in [-10\%, 60\%]$, whereas it decreases in the case of the skilled workers. Because utility is a concave function of consumption, a redistribution of income from rich (skilled) workers to poor (unskilled) workers increases average utility and, hence, welfare as measured by the ex ante expected lifetime utility.

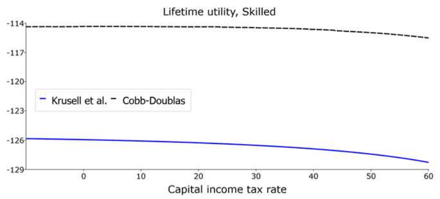
In summary, we find that the skill complementarity of capital provides a rationale for redistribution via capital income taxation through both its redistribution between worker-skill types and its strong effect on interest rates.

5.2 Optimal income tax progressivity

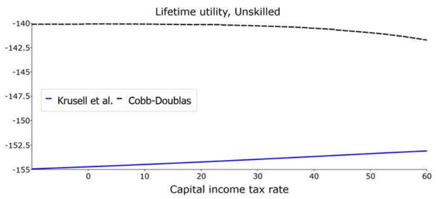
In the previous section, we kept the labor income tax progressivity constant, $\theta_1 = 0.137$. In the following, we consider the simultaneous optimization of the two tax parameters θ_1 and τ^k. The optimal capital income tax rate, τ^k, is presented in Fig. 6 (solid blue line, left-hand scale) as a function of $\theta_1 \in [-0.05, 0.256]$, where $\theta_1 < 0$ characterizes a regressive labor income tax system and $\theta_1 = 0.256$ corresponds to the highest progessivity parameter estimated by Holter et al. (2019) for a sample of 22 OECD countries during 2000–2007.[29] Evidently, the optimal capital income tax rate, τ^k, decreases with the tax progressivity index, θ_1, as the additional welfare gain from redistribution via the capital income tax rate, τ^k, also decreases with θ_1. The broken black line ('Welfare' on the right-hand scale) presents the welfare gain compared to the benchmark case if the progressivity parameter changes to θ_1 and τ^k is chosen optimally. Welfare is measured by the increase in total consumption and is a hump-shaped function of θ_1, with a maximum at $\theta_1 = -0.006$. The optimal joint tax policy is a combination of an approximately flat income tax rate ($\theta_1 \approx 0$) and a high capital income tax rate, $\tau^k = 37.9\%$.[30] The welfare gain is considerable

[29] The value $\theta_1 = 0.256$ is taken from their Table 1 for the country 'Denmark'.

[30] Our result differs to some extent from the one in Conesa and Krueger (2009). They find an optimal flat labor income tax schedule but with a large lump-sum transfer component. Our tax function (2.4) differs



(a) Skilled workers



(b) Unskilled workers

Fig. 5 Lifetime utilities U_t^i of skilled and unskilled workers

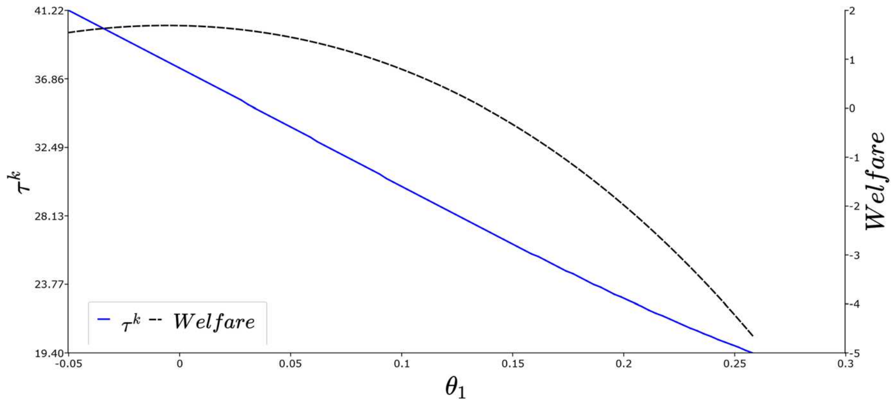


Fig. 6 Tax progressivity θ_1, capital tax τ^k and welfare

Table 5 The effects of optimal tax policy $(\theta_1, \tau_k) = (-0.006, 37.9\%)$ on factor supplies and prices (percentage changes, relative to benchmark)

Variable	Optimal policy
K	4.91
L^S	12.04
L^U	10.36
w^S	−4.11
w^U	−0.64

and amounts to 1.69% of total consumption. Even if we decrease the capital income tax rate to zero, $\tau^k = 0$, so the level of labor income taxes must be increased significantly (as reflected in a decline of θ_0), welfare gains from a flat-rate labor income tax would still accrue at 1.0% of total consumption. Our results clearly indicate that, in the present model with capital-skill technology, 1) redistribution functions through capital income rather than progressive labor income taxation, and 2) the dominant part of the quantitative welfare gain results from the decrease in labor income progressivity to a flat-rate regime.

Table 5 presents the effects of the optimal tax policy (θ_1, τ^k) on factor supplies and factor prices. The strong decrease in tax progressivity to an almost flat rate labor income tax, $\theta_1 = -0.006$, and the simultaneous increase in τ^k from 22.8% to 37.9% increase the labor supplies of both the skilled and the unskilled workers strongly as net labor income increases. The increase in skilled labor L_t^S of 12.0% is slightly stronger than the increase in unskilled labor L_t^U of 10.3%.[31] As a consequence of greater working hours, aggregate savings increase, so capital stock K_t increases by 4.9%, even though capital income taxation more than doubles. As the increase

from that in Conesa and Krueger (2009) based on more recent empirical evidence on the tax curve. In addition, we assume a technology with capital skill complementarity.

[31] In addition, both labor-supply-age profiles become flatter. The average age of the unskilled (skilled) workers increases from 38.65 (38.75) years to 39.14 (39.25) years. As argued in detail by Erosa and Gervais (2001), a flatter labor-supply-age profile increases welfare.

in labor is stronger than is the increase in savings, wages decrease. However, the decrease in gross wages is rather disproportionate. The skilled workers suffer from a wage decrease of 4.1%, whereas the decrease in the gross wages of the unskilled workers is more moderate, amounting to 0.64%. In fact, the decrease in the wage for the unskilled workers, w^U, and the effects from a flatter labor income tax schedule, $\theta_1 = -0.006$, are more than offset by the decrease in the level of labor income taxation (as measured by the variable θ_0) that is necessary to balance the fiscal budget for the higher capital income tax rate, τ^k.[32] Therefore, we observe two main drivers for the welfare results: 1) the net labor income of unskilled workers increases, so (net and gross) income inequality decreases, and 2) aggregate factor supplies and, hence, total production and consumption increase.[33]

6 Extensions

In this section, we consider the robustness of our results in Sect. 5. First, we integrate the effects of income uncertainty in the form of idiosyncratic uninsurable productivity shocks into our analysis and find that the optimal capital tax rate is rather insensitive to it. Second, we study the effects of different types of capital, structures and equipment, and find that these two capital types should be taxed at different rates. Finally, we describe effects of the aging US economy on capital taxation and study the dynamics of an instant once-and-for-all change in the capital income tax rates on the economic dynamics during the (ongoing) demographic transition.

6.1 Income uncertainty

The various and, to some extent, opposing welfare effects of capital taxes are magnified in the presence of income uncertainty. On the one hand, stochastic income increases the need to build up precautionary savings. Therefore, a capital tax that introduces a wedge in the optimality condition on savings inhibits the individual from insuring him- or herself against the bad luck of negative idiosyncratic productivity shocks by means of savings. On the other hand, income uncertainty increases the welfare-improving effects of redistribution and insurance. In dynamic general equilibrium models with infinite and finite lifetimes, Hubbard and Judd (1986) and İmrohoroğlu (1998), respectively, show that the optimal capital income tax is positive if households are subject to uninsurable idiosyncratic income risk or borrowing constraints.

In this section, we introduce income uncertainty, which is standard in the quantitative analysis of medium-scale OLG models, e.g., as in Huggett (1996), Conesa

[32] Owing to our assumption of capital-skill complementarity in production and the simultaneous increase in the capital income tax rate τ^k, our result differs from that in early studies on the optimality of flat-rate taxes in overlapping generations models. For example, Ventura (1999) reports that the labor supply of the less productive workers decreases in the case of a flat-rate labor income tax. In his model, the net wage of the low-productivity workers decreases after a flat-rate tax reform.

[33] The consumption increase also implies greater revenues from consumption taxation for our constant consumption tax rate, τ^c, so the decrease in the level of labor income taxes is reinforced.

and Krueger (2009) or Kitao (2014). In particular, we add multiplicative stochastic component ζ to individual productivity, so the gross labor income of the j-year-old worker with idiosyncratic productivity ζ and permanent productivity $i \in \{S, U\}$ in period t, $y_t(i, j, \zeta) = \zeta \bar{y}^j A_t w_t^i l_t$, consists of the product of his or her idiosyncratic productivity ζ; age component $\bar{y}^j$; aggregate labor productivity A_t; the wage of skill type i, w_t^i; and his or her working time $l_t(i, j, \zeta, a)$. Labor supply l_t also depends on individual assets a_t. We assume that the logarithm of the idiosyncratic component ζ follows a simple AR(1) process with persistence parameter ρ_ζ and unconditional variance σ_ζ^2. We select $\rho_\zeta = 0.97$ and $\sigma_\zeta^2 = 0.02$ in accordance with Kitao (2014).[34] In the numerical procedure, we discretize the state space for stochastic component ζ using 7 values and choose the grid size, so the Gini coefficient of hourly wages is 0.374, as estimated by Heer and Maußner (2024) using PSID data.[35] In addition, we impose borrowing constraint $a_t(i, j, \zeta) \geq 0$. We find that 23% of the households are credit constrained in the economy, which matches the empirical value reported by Budría Rodríguez et al. (2002) and Krueger et al. (2016).

In accordance with Conesa and Krueger (2009), we find that the various effects of income uncertainty and borrowing constraints on welfare approximately cancel each other out in the presence of progressive labor income taxation. The optimal capital income tax rate is 23.1% in our model, and the welfare effects of implementing the optimal tax policy or abolishing the capital income tax equal 0.12% and -1.01% of total consumption, respectively (as presented in Table 3).[36]

For comparison, we recalculate the model to replicate the basic elements of the model in Conesa and Krueger (2009). To do so, we set parameter $\eta = 4.0$ and economic growth rate g equal to zero; we assume a Cobb–Douglas production technology and recalibrate parameters α, γ and β to imply a wage share of 65%, a real interest rate of 4% and an average number of working hours of 0.25, respectively. The results are presented in the third entry row of Table 3 under the header 'Income uncertainty'. We find an optimal capital income tax rate of 34.7% for this case, whereas Conesa and Krueger (2009) estimate an optimal tax rate of 36%. The very small remaining difference can be explained by the use of a different calibration period and slightly different pension and income tax schedules. If we assume the 'Krusell et al.' technology (2.11) with capital skill complementarity rather than the Cobb-Douglas technology for the cases $\eta = 4.0$ and $g = 0\%$, the optimal capital income tax rate equals 52.0%, as presented in Table 3.

[34] The calibration $\rho_\zeta = 0.97$ constitutes the median value in the three cited studies. For example, Huggett (1996) and Conesa and Krueger (2009) apply values of 0.96 and 0.98, respectively, for ρ_ζ.

[35] The Gini coefficients of gross income, wealth and consumption amount to 0.470, 0.666 and 0.426, respectively. Quadrini and Ríos-Rull (2015) present empirical evidence that the Gini coefficients of US earnings and wealth in 2010 were 0.65 and 0.85, respectively, according to data from the Survey of Consumer Finance (SCF), whereas Krueger et al. (2016) find a Gini coefficient of earnings of 0.43 when using PSID data. Heathcote et al. (2010) presents empirical evidence for the U.S. economy that the Gini coefficient of nondurable consumption was substantially smaller than was the Gini coefficient of income from 1990–2006, with a Gini coefficient of consumption ranging between 0.32 and 0.40.

[36] A detailed decomposition analysis on how the various model elements (heterogeneity types, income uncertainty, borrowing limit, general equilibrium effects) affect optimal capital taxes is presented in Appendix D.

6.2 Two types of capital: structures and equipment

In our model in Sect. 2, we consider only aggregate capital as a production factor in (2.11) following Angelopoulos et al. (2015). In this sensitivity analysis, we distinguish structures and equipment capital – K_t^S and K_t^E, respectively – as in the original article by Krusell et al. (2000):

$$Y_t = (K_t^s)^\alpha \left[\mu \left(A_t L_t^U\right)^\sigma + (1-\mu)\left(\lambda (K_t^e)^\rho + (1-\lambda)\left(A_t L_t^S\right)^\rho\right)^{\frac{\sigma}{\rho}}\right]^{(1-\alpha)/\sigma}. \quad (6.1)$$

Only equipment capital is complementary to production skills. The two types of capital also depreciate at the different rates, δ^s and δ^e.

Households can hold their savings in the form of three assets: government bonds, equipment capital and structures. In our deterministic model, the no-arbitrage condition that all assets have the same after-tax returns must hold:

$$r^b = (1-\tau_t^{ks})r_t^s = (1-\tau_t^{ke})r_t^e, \quad (6.2)$$

where r_t^s and r_t^e denote the rental returns on the structures and the equipment in period t with corresponding tax rates τ_t^{ks} and τ_t^{ke}, respectively. The rest of the model is identical to the benchmark model in Sect. 2 and is described in more detail in Appendix E.

With two different types of capital, the two tax rates (τ^{ks}, τ^{ke}) that maximize steady-state expected lifetime utility diverge substantially, as only capital equipment is complementary to skills in production. Specifically, we find $\tau^{ke} = 59.3\%$ and $\tau^{ks} = 9.0\%$.[37] Our results are in close accordance with those reported by Slavík and Yazici (2014) for a model with an infinite lifetime. For the same intertemporal elasticity of substitution, $1/\sigma = 1/2$, the authors find optimal tax rates on the rental rates of equipment and structures of 39.5% and 0%, respectively. Our optimal tax rates on the two types of capital are again higher than are those found in the infinite-lifetime model by 9 and 20 percentage points, respectively, because of life-cycle effects. The associated steady-state welfare change from imposing the optimal tax rates equals 0.70% of total consumption, as presented in the last row of Table 3. Decreasing both capital taxes τ^{ke} and τ^{ks} to zero implies steady-state welfare losses equal to 0.60% of total consumption. In summary, our welfare effects of optimal capital taxation are quantitatively more significant in comparison with our benchmark model if we distinguish two types of capital.

6.3 Demographic change and transition dynamics

The demographic transition is likely to affect the capital stock per worker over the subsequent years. Starting with the seminal paper by De Nardi et al. (1999), studies such as Krueger and Ludwig (2007), Kitao (2014) or Heer et al. (2020) have exam-

[37] The average tax rate on capital income, $r^{ks}K^s + r^{ke}K^e$, amounts to 42.3% and is significantly higher than is that in the benchmark model with only one type of capital.

ined the effects of aging on factor supplies and prices depending on possible pension reforms. In essence, the projected increase in capital stock relative to working hours implies a lower real interest rate and a higher wage rate, with the quantitative effect depending on pension policies. Applying our arguments based on the production technology described above — that is, higher taxes are associated with lower after-tax real interest rates and a more equal wage distribution — we find that the higher capital intensity in the upcoming years increases the optimal capital income tax rate. However, the optimal policy is sensitive with respect to the transitional effects.[38]

Optimal Capital Taxes in the Year 2050 Considering our model in the steady state for the benchmark calibration but changing the population parameters, survival probabilities and population growth rate to the values projected for 2050 by United Nations (2022) and adjusting social security tax τ^p from 9.49% to 10.97%, we find that the optimal capital income tax rate, $\tau^{k,*}$, increases strongly, from 27.3% in 2020 to 35.2% in 2050. In addition, the quantitative welfare effects measured by consumption equivalent changes approximately double in size, so a decrease in the capital tax rate to 0% implies a welfare loss of 0.43% of total consumption, as presented in the last row of Table 3.

Transition Dynamics Thus far, we have analyzed only the steady-state welfare effects. Assuming that additional pension expenditures due to the demographic transition are financed with the help of social security taxes on labor income, the optimal steady-state capital tax rate amounts to 35.2% in 2050. However, during the transition to the new steady state, generations may either lose or gain from the optimal tax policy. In addition, the transitional time can be quite long and span decades or even centuries. For example, Straub and Werning (2020) report that the convergence to the steady state of the optimal Ramsey policy with a zero capital income tax rate can take centuries in the Chamley-Judd economy. The convergence speed is often very slow in overlapping generations models as well. Typically, in models that study pension policies during the demographic transition, a convergence period of 100−150 years is added to the period when the population dynamics become stationary. Therefore, we study the transitional dynamics as follows. Prior to 1966, we assume the economy to be in the steady state at the calibrated capital income tax rate, $\tau^k = 22.8\%$, and the population parameters observed in 1965. During the subsequent years 1966-2050, the population parameters are as projected by United Nations (2022). After 2050, the population parameters remain constant. We extend the time horizon to 2260 and assume that the economy is stationary thereafter. In 1966, the government announces the change in the capital income tax rate, τ^k, to the optimal steady-state tax rate $\tau^k = 35.2\%$, which becomes effective once-and-for-all in 2021. Therefore, households adjusted their labor supply and savings to the new capital tax rate in 1966.

Figure 7 presents the lifetime utilities of those generations alive in period 2020 and those born over the next 240 years. The oldest generation in period 2020 is 99

[38] Appendix F presents the effects of various pension policies (later retirement, lower pensions), higher public debt and a higher share of skilled workers in the labor force on optimal capital taxes. While the optimal capital tax rate is relatively insensitive to pension and debt policies, an increasing share of skilled workers leads to a substantially lower optimal capital tax rate. As the number of skilled workers rises relative to unskilled workers, the wage distribution becomes more equal, reducing the redistributive impact of capital taxation.

years old and was born in 1941.[39] The graphs depict the consumption equivalent change in the average lifetime utility of a generation born in year *t* relative to the case with a constant capital income tax rate $\tau^k = 22.8\%$ (solid blue line). All generations alive in 2020 would suffer from an increase in the capital income tax rate, τ^k, to the long-term optimal rate of 35.2% in 2021 (broken black line); the first generation that benefits from a higher capital tax rate enters the economy in 2097.[40] The maximum welfare loss accrues to the generation aged 57 years in 2021 and amounts to 0.42% of total consumption.

To understand the behavior of generational welfare, we need to study the dynamics of the factor supplies, factor prices and tax rates, as presented in Figs. 8 and 9. Because of the demographic transition, the number of retirees relative to workers increases from 12.2% in 1965 to 30.9% in the long run for our calibration (upper left panel in Fig. 8).[41] As the relative number of workers decreases during the transition, the aggregate labor supplies of both the skilled and the unskilled workers, L_t^S and L_t^U, decrease, as presented in the second panel row in Fig. 8. The change in the skilled and unskilled labor supply is more pronounced in the case with $\tau^k = 22.8\%$ (solid blue line) than it is in the case of a higher capital tax rate $\tau^k = 35.2\%$ (broken black line) due to the disincentive effects from higher labor income taxation. In addition, retirees dissave in old age such that aggregate savings and, hence, capital decrease in the long run because of the higher share of retirees. In the case with $\tau^k = 22.8\%$ (solid blue line), the long-term decrease in capital stock is less pronounced than it is in

[39] We assume that individuals are born at age 20 and live a maximum life of 80 years, corresponding to a real-life age of 99. We continue to use real-life age rather than age index *j* when we refer to individual generations.

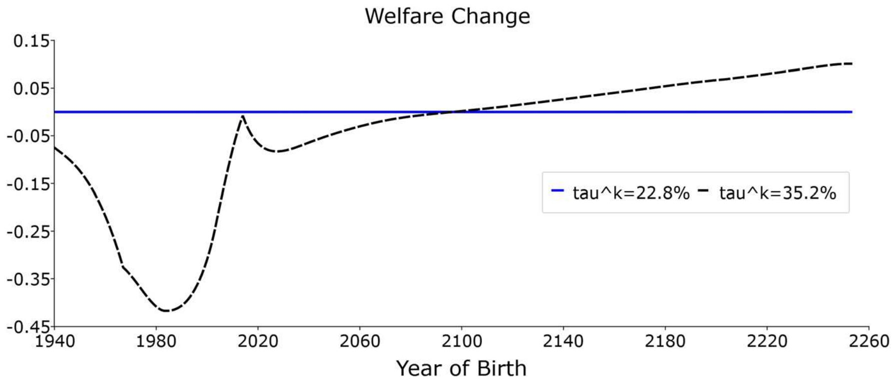
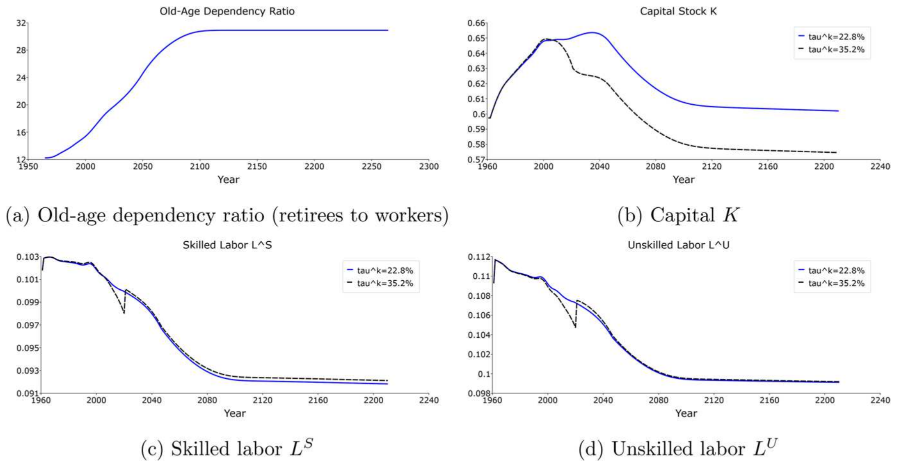


Fig. 7 Transitional welfare effects

[40] Our result is in close accordance with those of Grüner and Heer (2000), who consider the welfare effects of a once-and-for-all change in the capital income tax rate in a model with infinite lifetime. The authors find that a change to the optimal steady-state capital tax rate (32%) implies lower instantaneous utility for the first 46 years after the policy change.

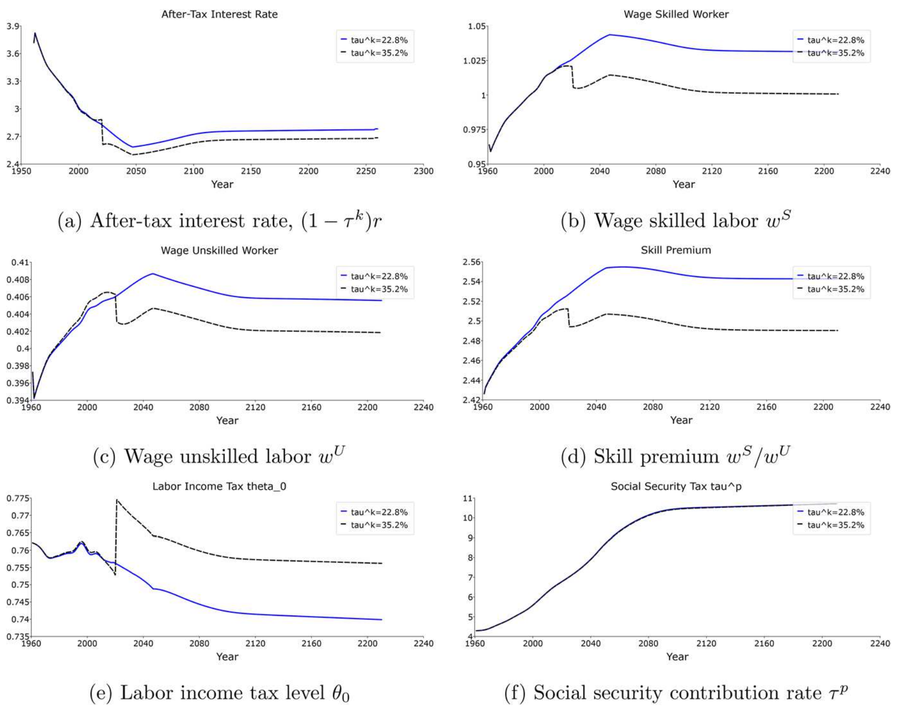
[41] The population and, hence, the dependency ratio become stationary only 80 years after the survival probabilities and the population growth are set constant in 2050.



(a) Old-age dependency ratio (retirees to workers) (b) Capital K

(c) Skilled labor L^S (d) Unskilled labor L^U

Fig. 8 Dynamics of the dependency ratio and the production factors K_t, L_t^S and L_t^U



(a) After-tax interest rate, $(1-\tau^k)r$ (b) Wage skilled labor w^S

(c) Wage unskilled labor w^U (d) Skill premium w^S/w^U

(e) Labor income tax level θ_0 (f) Social security contribution rate τ^p

Fig. 9 Dynamics of factor prices and taxes

the case of a higher capital income tax rate, $\tau^k = 35.2\%$, and the dynamics of the capital stock are hump-shaped — as typically found in the overlapping generations models of the demographic transition[42] — because the share of the households with large wealth holdings around age 50–65 initially increases (as presently observed in the U.S. economy for the case of the baby boomers). Since the decrease in labor is more pronounced than is the decrease in capital for the case $\tau^k = 22.8\%$, the wages of both the skilled and unskilled workers increase, as presented in Fig. 9. In the case of a higher capital income tax rate τ^k, the relative changes in the production factors are smaller such that, compared with the values in 2020, gross wages even decrease in the long run. However, after-tax factor prices are substantially different for these two cases. As the response of the capital stock is sluggish, the after-tax interest rate decreases by almost 0.5 percentage points on impact in 2021 in the case of the higher capital tax rate, $\tau^k = 35.2\%$ (upper left panel in Fig. 9). The higher capital tax rate, $\tau^k = 35.2\%$, also allows for a smaller burden on labor income taxes, which is reflected in an increase in parameter θ_0 (lower left panel in Fig. 9), whereas the social security tax rate, τ^p, increases from 7.1% to 10.5% between 2020 and 2100 in both cases (lower right panel) to finance a higher public pension load with fewer workers.

The welfare effects displayed in Fig. 7 can be explained with the help of the dynamics of after-tax factor prices. Explaining the effects of higher capital taxes on the lifetime utility for retirees alive in 2020 is straightforward. Interest income constitutes a large component of the retirees' total income and decreases due to the higher capital tax. Therefore, the retirees' lifetime utility decreases. The welfare loss resulting from a higher capital tax rate is strongest for the middle-aged individuals in 2021. They have already supplied their working hours at a young age and do not benefit from lower labor income taxation until they reach their late working life, when their working hours have approximately halved (compare Fig. 3). In our steady-state welfare analysis in Sect. 5, we also noted that one of the most important welfare-increasing effects of the capital tax rate is derived from the lower inequality, as reflected in a lower skill premium. However, as illustrated in the middle right panel of Fig. 9, the skill premium adjusts only slowly to its new long-term value in the case with $\tau^k = 22.8\%$ due to the sluggish response of the capital stock. Therefore, the welfare-increasing effect of a relatively lower skill premium in the case of $\tau^k = 35.2\%$ (vis-à-vis the case with $\tau^k = 22.8\%$) sets in only during the later phase of the transition.[43]

In summary, our results for the transitional dynamics after a change in the capital income tax rate suggest that we should be careful to draw firm policy recommendations from the steady-state results in Sect. 5. Furthermore, our transitional welfare analysis is sensitive with respect to the two assumptions that 1) capital tax rate changes are preannounced in 1966 and 2) implemented in 2021 once-and-for-all. If, for example, the change in capital taxes in 2021 is unexpected, the welfare losses for the generations alive in 2020 are even greater than they are in the case of a preannounced change (not presented) because such generations cannot adjust their behavior at a young age prior to 2021 and reallocate their labor supply and consumption intertemporally in optimal response to the changes in net factor prices after 2020. We also do not consider a gradual implementation of tax rate increases or allow the

[42] See, e.g., Fig. 13 in Kitao (2014) or Figs. 19 and 20 in De Nardi et al. (1999).

[43] The dynamic behavior of the skill premium also holds over to that of the relative pensions of the skilled and unskilled workers (not presented), so inequality in pensions also decreases gradually after 2020.

government to run fiscal imbalances and increase debt during the demographic transition, as, for example, analyzed by Braun and Joines (2015), which could cushion the welfare losses for the present generations.

7 Conclusion

The optimal tax rate on capital income is sensitive to the underlying production technology. We compare the standard Cobb–Douglas technology with the technology of Krusell et al. (2000), in which capital is complementary to skilled work. We find that the optimal capital tax rate is significantly higher in the case of capital–skill complementarity and reaches 27% (versus 9% in the Cobb–Douglas case) in our economy calibrated with regard to the characteristics the U.S. economy. Importantly, the optimal capital tax rate increases in an aging economy and is equal to 35% in the US by 2050. Technology with skill-capital complementarity has two properties that account for these results. First, high capital taxes imply a significantly greater decrease in the after-tax interest rate under such technology than in the case of a Cobb–Douglas technology; therefore, such taxes provoke a flatter life-cycle profile of labor. Second, with high capital taxes and low capital in production, the skill premium decreases in the case of skill–capital complementarity. As a consequence, the distribution of income becomes more equal. Both effects are welfare-enhancing.

We also find that the optimal redistributive fiscal policy is sensitive to the presence of capital skill complementarity in production. For our calibration, it is even beneficial to use capital taxes rather than progressive labor income taxes to redistribute wealth from high- to low-income workers and retirees. However, our distribution and welfare results need to be interpreted cautiously because they hold only at the steady state. Importantly, the transition to the new steady state could easily amount to many decades or even centuries. In our transitional welfare analysis, we find that the welfare-increasing effect of higher capital income taxes does not manifest itself until the end of this century. Therefore, one should be careful to draw firm policy conclusions from our steady-state results.

In conclusion, we highlight three directions for future research to extend and improve upon our model assumptions. First, we consider a variation in labor in response to fiscal policies along the intensive margin. Technological progress is very likely to affect labor along the extensive margin as well. Robots and automation may be perfect substitutes for workers. Second, we assume that capital deepening benefits mainly skilled workers by increasing the skill premium. However, one of the most dramatic underlying current trends in production technology is the shift from unskilled to skilled worker replacement because of AI. As emphasized by Acemoglu and Restrepo (2018), the net impact of AI depends on the speed and magnitude of the displacement of old tasks versus the creation of new tasks. We expect the optimal capital tax rate to be sensitive to the nature and direction of technological change. Third, we neglect any effect that the capital tax rate may have on the growth rate. However, high capital taxes decrease the incentives to invest in research and human capital and may harm technological progress.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10797-025-09938-0.

Author Contributions The paper has been written by a single author. All contributions and errors are mine (Burkhard Heer).

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no Conflict of interest.

References

Acemoglu, D., & Restrepo, P. (2018). Artificial intelligence, automation, and work. In *The Economics of Artificial Intelligence: An Agenda*, pp. 197–236. National Bureau of Economic Research, Inc.

Aiyagari, R. S. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy, 103*(6), 1158–1175.

Alvarez, Y., Burbidge, J., Farell, T., & Palmer, L. (1992). Optimal taxation in a life-cycle model. *Canadian Journal of Economics, 25*(1), 111–122.

Angelopoulos, K., Asimakopoulos, S., & Malley, J. (2015). Tax smoothing in a business cycle model with capital-skill complementarity. *Journal of Economic Dynamics and Control, 51*, 420–444.

Bastani, S., & Waldenström, D. (2020). How should capital be taxed? *Journal of Economic Surveys, 34*(4), 812–846.

Braun, R. A., & Joines, D. H. (2015). The implications of a Graying Japan for Government Policy. *Journal of Economic Dynamics and Control, 57*, 1–23.

Brinca, P., Holter, H. A., Krusell, P., & Malafry, L. (2016). Fiscal multipliers in the 21st century. *Journal of Monetary Economics, 77*, 53–69.

Budría Rodríguez, S., Díaz-Giménez, J., Quadrini, V., & Ríos-Rull, J.-V. (2002). Updated facts on the U.S. distributions of earnings, income, and wealth. *Quarterly Review, 26*(3), 2–35.

Carey, D., & Rabesona, J. (2003). Tax ratios on labour and capital income and on consumption. *OECD Economic Studies, 2002*(2), 129–174.

Cass, D. (1965). Optimum growth in an aggregate model of capital accumulation. *Review of Economic Studies, 32*, 233–240.

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

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

De Nardi, M., & Yang, F. (2016). Wealth inequality, family background, and estate taxation. *Journal of Monetary Economics, 77*, 130–145.

De Nardi, M., Imrohoroğlu, S., & Sargent, T. J. (1999). Projected US demographics and Social Security. *Review of Economic Dynamics, 2*, 575–615.

Erosa, A., & Gervais, M. (2001). Optimal taxation in infinitely-lived agent and overlapping generations models: A review. *Federal Reserve Bank of Richmond Economic Quarterly, 87*(2), 23–44.

Erosa, A., & Gervais, M. (2002). Optimal taxation in life-cycle economies. *Journal of Economic Theory, 105*(2), 338–369.

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

Gross, T., & Klein, P. (2022). Optimal tax policy and endogenous growth through innovation. *Journal of Public Economics, 209*, Article 104645.

Grüner, H.-P., & Heer, B. (2000). Optimal flat-rate taxes on capital - A re-examination of Lucas' supply-side model. *Oxford Economic Papers, 52*, 289–305.

Guerreiro, J., Rebelo, S., & Teles, P. (2022). Should robots be taxed? *Review of Economic Studies, 89*, 279–311.

Hansen, G. D. (1993). The cyclical and secular behavior of the labor input: Comparing efficiency units and hours worked. *Journal of Applied Econometrics, 8*(1), 71–80.

Heathcote, J., Perri, F., & Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics, 13*, 15–51.

Heathcote, J., Storesletten, K., & Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics, 132*(4), 1693–1754.

Heer, B. (2001). Wealth distribution and optimal inheritance taxation in life-cycle economies. *Scandinavian Journal of Economic, 103*, 445–465.

Heer, B., & Maußner, A. (2024). *Dynamic general equilibrium modeling: Computational methods and applications* (3rd ed). Springer.

Heer, B., Polito, V., & Wickens, M. R. (2020). Population aging, social security and fiscal limits. *Journal of Economic Dynamics & Control, 116*, Article 104645.

Holter, H. A., Krueger, D., & Stepanchuk, S. (2019). How do tax progressivity and household heterogeneity affect Laffer curves? *Quantitative Economics, 10*(4), 1317–1356.

Hubbard, R. G., & Judd, K. I. (1986). Liquidity constraints fiscal policies, and consumption. *Brookings Papers on Economic Activity, 1*, 601–641.

Hubbard, R. G., & Judd, K. L. (1987). Social security and individual welfare: Precautionary saving, borrowing constraints, and the payroll tax. *American Economic Review, 77*, 630–646.

Huggett, M. (1996). Wealth distribution in life-cycle economies. *Journal of Monetary Economics, 38*(3), 469–494.

Hurd, M. D. (1989). Mortality risk and bequests. *Econometrica, 57*(4), 799–813.

İmrohoroğlu, S. (1998). A quantitative analysis of capital income taxation. *International Economic Review, 39*(2), 307–328.

İmrohoroğlu, S., & Kitao, S. (2009). Labor supply elasticity and social security reform. *Journal of Public Economics, 93*, 867–878.

Jones, C. I. (1995). R & D-based models of economic growth. *Journal of Political Economy, 103*(4), 759–783.

Jones, C. I. (2022). Taxing top incomes in a world of ideas. *Journal of Political Economy, 130*, 2227–2274.

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

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

King, R. G., Plosser, C. I., & Rebelo, S. T. (1988). Production, growth and business cycles I. The basic neoclassical model. *Journal of Monetary Economics, 21*(2–3), 195–232.

Kitao, S. (2014). Sustainable social security: Four options. *Review of Economic Dynamics, 17*, 756–779.

Koopmans, T. C. (1965). *On the concept of optimal economic growth. In The economic approach to development planning*. Elsevier.

Krueger, D., & Ludwig, A. (2007). On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare. *Journal of Monetary Economics, 54*, 49–87.

Krueger, D., Mitman, K., & Perri, F. (2016). Macroeconomics and household heterogeneity. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of the Macroeconomics* (pp. 843–921). Elsevier.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., & Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica, 68*(5), 1029–1053.

Lucas, R. E. (1990). Supply-side economics: An analytical review. *Oxford Economic Papers, 42*(2), 293–316.

Maliar, L., Maliar, S., & Tsener, I. (2022). Capital-skill complementarity and inequality: Twenty years after. *Economics Letters, 220*, Article 110844.

Mendoza, E. G., Razin, A., & Tesar, L. L. (1994). Effective tax rates in macroeconomics: Cross country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics, 34*, 297–323.

Mirrless, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies, 38*(114), 175–208.

OECD (2022). *OECD Economic Outlook database No 111*. Paris: OECD Publishing.

OECD (2021). *Pensions at a glance 2021: OECD and G20 indicators*. OECD Publishing.

Prettner, K., & Strulik, H. (2020). Innovation, automation and inequality: Policy challenges in the race against the machine. *Journal of Monetary Economics, 116*, 249–265.

Quadrini, V., & Ríos-Rull, J.-V. (2015). Inequality in macroeconomics. In A. Atkinson & F. Bourguignon (Eds.), *Handbook of Income Distribution* (pp. 1229–1302). North Holland.

Ramsey, F. (1927). A contribution to the theory of taxation. *Economic Journal, 37*, 47–61.

Ramsey, F. (1928). A mathematical theory of saving. *Economic Journal, 38*, 543–559.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5), 71–102.

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

Straub, L., & Werning, I. (2020). Positive long-run capital taxation: Chamley-Judd revisited. *American Economic Review, 110*(1), 86–119.

Thuemmel, U. (2022). Optimal taxation of robots. *Journal of the European Economic Association, 21*(3), 1154–1190.

Tsai, Y.-C., Yang, C., & Yu, H.-J. (2022). Rising skill premium and the dynamics of optimal capital and labor taxation. *Quantitative Economics, 13*, 1061–1099.

United Nations, D. (2022). *World Population Prospects 2022* online edition. New York: United Nations.

Ventura, G. (1999). Flat-rate tax reform: A quantitative exploration. *Journal of Economic Dynamics & Control, 23*, 1425–1458.