

Machines, Buildings, and Optimal Dynamic Taxes

Ctirad Slavík and Hakki Yazici*

October 20, 2013

Abstract

The effective taxes on capital returns differ depending on capital type in the U.S. tax code. This paper uncovers a novel reason for the optimality of differential capital taxation. We set up a model with two types of capital - equipments and structures - and equipment-skill complementarity. We show that it is optimal to tax equipments at a higher rate than structures. In a calibrated model, the optimal tax differential rises from 27 to 40 percentage points over the transition to steady state. The welfare gains of optimal differential capital taxation can be as high as 0.4% of lifetime consumption.

JEL classification: E62, H21.

Keywords: Differential capital asset taxation, equipment capital, structure capital, capital-skill complementarity.

*Ctirad Slavík, Goethe University Frankfurt, email: slavik@econ.uni-frankfurt.de, Hakki Yazici, Sabanci University in Istanbul, email: hakkiyazici@sabanciuniv.edu. We thank Laurence Ales, Alex Bick, Nicola Fuchs-Schündeln, Kenichi Fukushima, Mike Golosov, Larry Jones, Fabrizio Perri, Florian Scheuer, Yuichiro Waki, participants at the 2013 Annual Meeting of the Society for Economic Dynamics, the 2013 Annual Meeting of the Royal Economic Society, the 2012 and 2013 Cologne Workshop on Macroeconomics, the 2013 Annual Congress of the European Economic Association, the 2013 REDg Workshop in Barcelona, the Conference in Honor of Larry Jones at the University of Minnesota, and seminar participants at Bilkent University, Central Bank of Turkey, Max Planck Institute in Bonn, Oxford University, and the University of Zürich for their comments. All remaining errors are ours.

1 Introduction

In the U.S. corporate tax code, the effective marginal tax rates on returns to capital assets show a considerable amount of variation depending on the capital type. For instance, according to Gravelle (2011), the effective marginal tax rate on the returns to communications equipment is 19%, whereas it is above 35% for nonresidential buildings.¹ This feature of the tax code has been the subject of numerous reform proposals since the 1980s. Recently, President Obama called for a reform to abolish the tax rules that create differential taxation of capital assets in order to “level the playing field” across companies.² Many economists have argued in favor of the proposals to abolish tax differentials following an efficiency argument first raised by Diamond and Mirrlees (1971): taxing different types of capital at different rates distorts firms’ production decisions, thereby creating production inefficiencies.

In this paper, we take a step back and reassess whether differential taxation of capital income is a desirable feature of the tax code. Theoretically, we uncover a novel economic mechanism that calls for optimality of differential capital asset taxation, but with an important caveat. In the current U.S. tax code, the effective tax rate on equipment capital (i.e., mostly machines) is on average 5% below the effective tax rate on structure capital (i.e., mostly nonresidential buildings). In contrast, our theory suggests that capital equipments should be taxed at a higher rate than capital structures. We conduct a quantitative exercise to assess the quantitative importance of optimal differential capital taxation. In our baseline calibration, we find that the tax rate on capital equipments should be at least 27 percentage points higher than the tax rate on capital structures in the transition and at the steady state. Furthermore, we find that the welfare gains of optimal differential capital taxation can be as high as 0.4% of lifetime consumption for reasonable parameter values.

We study dynamic optimal taxes in an economy in which people face idiosyncratic shocks to their labor skills and the government uses capital and labor income taxes to insure people against this risk. The key feature of our environment is capital-skill complementarity in the production technology. More specifically, following Gravelle (2011), we group capital assets into two categories: structure capital and equipment capital. We further assume that there are two types of labor: skilled and unskilled. Following the empirical evidence in Krusell, Ohanian, Ríos-Rull, and Violante (2000), we assume that the degree of complementarity between equipment capital and skilled labor is higher than the degree of complementarity between equipment capital and unskilled labor.³ Structure capital is neutral in terms of its

¹In Section 2, we explain in detail how these differentials are created and provide further information on the historical evolution of capital tax differentials in the U.S. tax code.

²The 2011 U.S. President’s State of the Union Address. Retrieved from <http://www.whitehouse.gov/the-press-office/2011/01/25/remarks-president-state-union-address>.

³Krusell, Ohanian, Ríos-Rull, and Violante (2000) document equipment-skill complementarity for the U.S.

complementarity with skilled and unskilled labor.

Capital-skill complementarity implies that skilled and unskilled labor are not perfect substitutes and that the skill premium – defined as the ratio of the skilled wage to the unskilled wage – is endogenous. In particular, a decrease in the stock of equipment capital decreases the skill premium, thereby creating an indirect transfer from the skilled agents to the unskilled ones. Therefore, depressing the level of equipment capital creates an extra channel of redistribution or social insurance. In order to depress equipment capital accumulation, the government taxes returns to equipment capital at a higher rate than it taxes returns to structure capital. This implies the optimality of differential asset taxation.

Our approach to optimal dynamic taxation follows the recent New Dynamic Public Finance literature, which analyzes optimal capital and labor income taxation à la Mirrlees (1971) in dynamic settings. We first use mechanism design theory to transform the optimal tax problem into a social planning problem and characterize its solution – the constrained efficient allocation – analytically. Then, we provide an implementation of the constrained efficient allocation in an incomplete markets environment in which people trade risk-free bonds. We show that in this implementation, returns to equipment and structure capital are taxed differentially at the corporate level as in the U.S. tax code.

Our theory has interesting implications for labor income taxes as well. A well-known result in Mirrleesian optimal taxation problems is that the marginal labor income tax on agents with the highest skill level should be zero. In our environment in which skilled and unskilled labor are not perfect substitutes, this result is no longer valid: it is optimal to subsidize skilled agents.⁴ Because of imperfect substitutability between skilled and unskilled labor, increasing the amount of skilled labor decreases the skill premium, which creates indirect redistribution. In order to encourage the supply of skilled labor, the government finds it optimal to subsidize skilled labor at the margin.

We build a quantitative version of our model to assess the importance of differential capital taxation. We do so in an environment with permanent skill types in which the only role of taxation is redistribution. In our benchmark calibration, the optimal equipment capital income tax is 27.6 percentage points higher than the tax on structure capital in the first period. The tax differential rises along the transition path to 39.6 percentage points at the steady state. Optimal labor wedges do not change much over time. At the steady state, the skilled labor wedge is -11.1% and the unskilled labor wedge is 26.6%.

economy. Flug and Hercowitz (2000) provide empirical evidence supporting equipment-skill complementarity in a large panel of countries.

⁴Stiglitz (1982) is the first paper to point out that the no distortion at the top result does not apply in environments with imperfect substitutability between skilled and unskilled labor. Our result is a generalization of his result to dynamic economies with production.

Next, we assess the welfare gains of optimal differential capital taxation. We compare welfare in the optimal tax system with welfare in a tax system in which the government is unrestricted in its choice of labor income taxes, but capital taxes on both types of capital are restricted to be equal. The welfare gains of optimal differential capital taxation are 0.05% in terms of lifetime consumption in the benchmark and can be as high as 0.35% within the set of reasonable parameter values. We find that the welfare gains of optimal differential capital taxation are even larger if we compare the optimal tax system with one in which structure and equipment capital are taxed as in the U.S. tax code (and labor taxes are chosen optimally). In this case, the welfare gains are 0.20% in the benchmark and can be as high as 0.40%.

We focus on the redistribution and insurance provision role of differential capital taxation and find that it is optimal to tax equipment capital at a higher rate than structure capital. Our mechanism relies on equipment-skill complementarity which has been documented empirically by Krusell, Ohanian, Ríos-Rull, and Violante (2000), among others. There could be other reasons for differential taxation of capital. For instance, some authors have argued that investment in equipment capital might create positive externalities. Other things being equal, positive externalities would be a reason to tax equipment capital at a lower rate than structure capital. Auerbach, Hassett, and Oliner (1994) point out, however, that it is hard to support the existence of such positive externalities on empirical grounds. In this paper, we abstract from all other possible reasons for differential capital taxation in order to isolate the redistributive and insurance provision role of differential taxation of capital.

Related Literature. Our paper relates to three different strands of literature.

First, the seminal paper of Diamond and Mirrlees (1971) provides an influential backdrop to the literature on differential taxation of capital assets: if the government has to raise revenue using linear taxes on income from different factors of production, then all factors should be taxed at the same rate. The implication of this result for capital taxation is clear: all capital should be taxed at the same rate.⁵ Several papers have analyzed the optimality of differential capital taxation, assuming that the government is exogenously restricted to a narrower set of fiscal instruments than in Diamond and Mirrlees (1971). For instance, Auerbach (1979) shows in an overlapping generations environment that it is optimal to tax capital differentially if the government cannot set labor income taxes optimally. In contrast, the government's inability to use debt to finance its spending does not typically call for differential capital taxation. Feldstein (1990) proves the optimality of differential capital

⁵Naturally, most papers in the literature on differential capital taxation have focused on estimating the welfare costs of taxing capital differently. See, for example, Auerbach (1983) and Gravelle (1994).

taxation in a static model in which the government is restricted to set the tax rate on one type of capital equal to zero. A problem of this approach to differential capital taxation is that its optimality depends, unappealingly, on the specific ad hoc restriction on the set of fiscal policy tools available to the government. Our paper provides a novel mechanism that implies the optimality of differential capital taxation in an environment in which the government is unrestricted in terms of the set of taxes it can employ. Therefore, our results do not depend on any ad hoc assumptions.

In this regard, our paper follows the New Dynamic Public Finance (NDPF) tradition. This literature studies optimal capital and labor income taxation in dynamic settings in which agents' labor skills change stochastically over time.⁶ The distinguishing feature of the papers in this literature is that they do not make any ad hoc assumptions about the nature of the tax system that is available to the government: the optimal tax system can be arbitrarily nonlinear in the history of capital and labor income. No paper in this literature, however, has studied differential taxation of capital assets prior to this paper. In addition, most papers in the NDPF literature only provide theoretical characterizations of the optimal dynamic tax structure. In contrast, this paper provides a set of theoretical results as well as a carefully calibrated quantitative analysis. In this sense, we contribute to the NDPF literature by adding to a set of recent papers that aim to provide practical policy recommendations by quantifying the theoretical implications of the NDPF literature.⁷

Finally, our paper is related to a set of theoretical studies on optimal static Mirrleesian taxation with endogenous wages. Stiglitz (1982) assumes that the labor supplies of agents with different skills are imperfect substitutes and shows that the agent with the highest income should be subsidized. In a similar setup with endogenous wages, Naito (1999) shows that the uniform commodity taxation result of Atkinson and Stiglitz (1976) is no longer valid under imperfect labor substitutability. We differ from this literature by focusing on a dynamic environment with different types of capital, which we use to analyze optimal differential taxation of capital assets both theoretically and quantitatively.

The rest of the paper is structured as follows. The next section discusses differential taxation of capital assets in the U.S. tax code. In Section 3, we lay out the model. In Section 4 we show that differential capital taxation is optimal. In Section 5 we show how to implement the constrained efficient allocation in a competitive market environment. Section 6 discusses our quantitative results, and Section 7 concludes.

⁶Golosov, Kocherlakota, and Tsyvinski (2003) is the seminal paper in this literature. For an excellent review of the NDPF literature, see the work of Kocherlakota (2010).

⁷See, for instance, Fukushima (2010), Golosov, Troshkin, and Tsyvinski (2011), and Farhi and Werning (2013).

2 Differential Taxation of Capital in the U.S. Tax Code

In this section, we first explain how the current U.S. tax system taxes returns to capital assets differentially. Then we provide a brief summary of the historical evolution of capital tax differentials.

According to the current U.S. corporate tax code, all capital income net of depreciation is taxed at the statutory rate of 35%. However, effective marginal taxes on net returns to capital might differ from the statutory rate if tax depreciation allowances differ from actual economic depreciation. To see this, let ρ_i be the return to capital type i , δ_i be its economic depreciation rate, and $\bar{\delta}_i$ be the depreciation rate allowed by the tax code. Suppose for simplicity that at the end of a period, the firm liquidates and there is no further production. Letting τ be the statutory tax rate, the effective tax rate, call it τ_i , is given by

$$\tau_i = \frac{(\rho_i - \bar{\delta}_i)\tau}{(\rho_i - \delta_i)}.$$

As first pointed out by Samuelson (1964), if $\bar{\delta}_i = \delta_i$, then $\tau_i = \tau$. In words, when tax depreciation equals economic depreciation, then the effective tax on the return to capital i equals the statutory rate. If this is true for all types of capital, there are no tax differentials. If, instead, for capital of type i the tax depreciation allowance is higher (lower) than the actual depreciation rate, then its return is effectively taxed at a lower (higher) rate (i.e., $\tau_i < (>)\tau$). As argued by Gravelle (2003), inter alia, such discrepancies between tax depreciation allowances and economic depreciation rates across capital assets are the main cause of differential capital taxation in the United States.⁸

Gravelle (2003) concisely summarizes the historical evolution of tax differentials.⁹ She reports that before the 1986 Tax Reform Act, the statutory corporate income tax rate was 46%. The effective tax rates on returns to most equipment assets were less than 10%, whereas they were around 35% to 40% for buildings. With the 1986 tax reform, the statutory rate was reduced to 35%, and the depreciation rules were altered to eliminate the tax differentials between equipment capital and structure capital. These policy changes resulted in effective

⁸Of course, in reality many firms continue operating for many periods and hence deduct the whole cost of a capital investment over time via depreciation deductions. The differences in effective tax rates are created by tax depreciation rules that make firms depreciate their capital assets either faster or slower than their economic depreciation over time. As long as the real interest rate is positive, tax depreciation rules that allow firms to deduct depreciation faster relative to actual economic depreciation decrease the effective tax rate on capital. Inflation also affects the real value of depreciation deductions because these deductions are based on the historical acquisition cost. For a detailed description of how effective tax rates are calculated, see Gravelle (1994).

⁹For a more detailed description of how tax differentials between capital equipments and structures have changed between 1950 and 1983, see Auerbach (1983).

equipment capital tax rates that were about 32% and structure capital tax rates that were still about 35% to 40%. As documented by Gravelle (2011), in the current U.S. corporate tax code (which went through a minor reform in 1993), equipments are taxed at 26% on average and structures are taxed at 32%. To sum up, capital equipments have historically been favored relative to capital structures, but the difference in effective tax rates has been declining over time.

3 Model

There is a continuum of measure one of agents who live for infinitely many periods.¹⁰ They differ in their skill levels: they are born either skilled or unskilled, $j \in H = \{u, s\}$. The skilled agent has productivity z_s and the unskilled agent z_u . Their skill type evolves stochastically over time according to a stochastic process π . History in period t is denoted by h^t and lies in the set of all possible period t histories H^t . Let $\pi_t(h^t)$ be the unconditional probability of history h^t .

Production Technology. An agent of skill level j in period t produces $l \cdot z_j$ units of effective j type labor when he works l units of labor. Aggregate amounts of skilled labor L_s and unskilled labor L_u in period t are given by, for $j = u, s$,

$$L_{j,t} = \sum_{\{h^t \in H^t | h_t = j\}} \pi_t(h^t) l_t(h^t) z_j. \quad (1)$$

We assume that there are two different occupational sectors in this economy: a skilled sector in which only skilled agents are allowed to work and an unskilled sector in which only unskilled agents are allowed to work. The first assumption reflects the fact that unskilled people do not have the skills to work in the skilled sector. The second assumption can be rationalized as follows. We assume that agents get the same disutility from working in the two sectors. Therefore, a skilled agent will choose to work in the skilled sector as long as he gets a higher wage in the skilled sector. This reasoning holds in the presence of taxes under our assumption that taxes are functions of income histories only. We discuss the nature of the tax system in more detail below.

There are two different types of capital stock in the economy: structure capital K_s and equipment capital K_e . The economy is initially endowed with $K_{s,1}^*$ and $K_{e,1}^*$ units of these capital goods, respectively. There is a representative firm that produces the output in the

¹⁰None of the results presented here rely on the fact that time lasts forever.

economy according to the following production function:

$$Y = F(K_s, K_e, L_s, L_u),$$

where K_s, K_e refer to structure capital and equipment capital, respectively. We also define a function \tilde{F} that gives the total wealth of the economy:

$$\tilde{F} = F + (1 - \delta_s)K_s + (1 - \delta_e)K_e.$$

Wages. Given the production function F , agents of skill type $j \in \{s, u\}$ in period t receive a wage rate $w_{j,t}$ for one unit of their labor, where

$$w_{j,t} = \frac{\partial F(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})}{\partial L_{j,t}} \cdot z_j. \quad (2)$$

We denote the wage rate that an agent with skill history h^t receives as $w_t(h_t)$, where $w_t(h_t) = w_{j,t}$ if $h_t = j$. We use $w_t(h_t)$ rather than $w_t(h^t)$ because an agent's wage rate in history h^t depends only on the current skill level h_t .

Capital-Skill Complementarity. The key feature of the technology is equipment capital-skill complementarity, which means that the degree of complementarity between equipment capital and skilled labor is higher than that between equipment capital and unskilled labor. This assumption has two important implications that make our model different from the standard model in the NDPF literature. First, an increase in the stock of equipment capital decreases the ratio of the marginal product of unskilled labor to the marginal product of skilled labor. In other words, the ratio of skilled to unskilled wages is endogenous, and this ratio is increasing in equipment capital. Structure capital, on the other hand, is assumed to be neutral in terms of its complementarity with skilled and unskilled labor. Second, skilled and unskilled labor are no longer perfect substitutes which implies that the skill premium is decreasing in the total amount of skilled labor and increasing in the total amount of unskilled labor. We formalize these assumptions on technology as follows.

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

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

Assumption 3. $\frac{\partial F/\partial L_s}{\partial F/\partial L_u}$ is strictly decreasing in L_s and strictly increasing in L_u .

Preferences. All agents have the same ex ante expected discounted utility function,

$$\sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} [u(c_t(h^t)) - v(l_t(h^t))],$$

where $\beta \in (0, 1)$ is a discount factor, $u, v : \mathbb{R}_+ \rightarrow \mathbb{R}$, and $c_t(h^t)$ and $l_t(h^t)$ represent consumption and labor at history h^t . We assume that u' , $-u''$, v' , and v'' all exist and are positive.

Allocation. An allocation is $x = (c_t(h^t), l_t(h^t), K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^{\infty}$, where $c_t, l_t : H^t \rightarrow \mathbb{R}_+$.

Feasibility. An allocation is feasible if in any period $t \geq 1$,

$$\sum_{h^t \in H^t} \pi_t(h^t) c_t(h^t) + K_{s,t+1} + K_{e,t+1} + G_t \leq \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \quad (3)$$

$K_{s,1} \leq K_{s,1}^*$, $K_{e,1} \leq K_{e,1}^*$, and (1) holds.¹¹ Here, $\{G_t\}_{t=0}^{\infty}$ is a sequence of exogenously given wasteful government consumption.

Optimal Tax Problem. We allow the government to choose from a very rich set of tax instruments. Specifically, taxes are allowed to be arbitrary nonlinear functions of income histories. Following Mirrlees (1971), we do not allow taxes to depend directly on agents' skill levels and their labor supplies. We also do not allow taxes to depend on the sectors people work in. One could motivate the assumption that taxes only depend on income by arguing that the government cannot condition taxes on skills, labor supplies, or people's sectors because they are unobservable, but none of what follows hinges on this interpretation.

However, for technical reasons, it is convenient to use the private information assumption explicitly. Formally, we assume that an agent's skill history, h^t , and hence his wage rate, $w_t(h^t)$, along with his labor supply $l_t(h^t)$ are private information. Income $y_t(h^t) = w_t(h_t)l_t(h^t)$ and consumption $c_t(h^t)$ are public information. Whether the agent works in the skilled or unskilled sector is also private information. This implies that when an agent wants to mimic the other type, he only has to mimic his income level. Since we assume that agents cannot switch sectors, an agent can only mimic the other type's income level by adjusting his labor hours.

Observe that any market arrangement with taxes is a particular mechanism. By revelation principle, no such mechanism can do better than the optimal direct truth-telling mechanism. Conversely, in section 5, we prove that there is a tax system that implements

¹¹One explanation for the increasing use of equipment capital observed in the data is that its relative price is decreasing. It is straightforward to extend our analysis to account for this possibility.

the allocation that arises from the optimal direct truth-telling mechanism. Therefore, finding the optimal tax system reduces to finding the optimal direct truth-telling mechanism, which is the problem of a social planner who assigns allocations as functions of agents' types subject to incentive compatibility constraints. One can think of this problem as one in which the agents report their types to the planner. The planner then assigns allocations as functions of the reports. The incentive compatibility constraints make sure that agents do not misreport their type. We now formally define the notion of incentive compatibility.

Incentive Compatibility. Define $\sigma_t : H^t \rightarrow \{u, s\}$. A reporting strategy is $\sigma = (\sigma_t)_{t=1}^\infty$. Let $\sigma^t(h^t) = (\sigma_1(h^1), \dots, \sigma_t(h^t))$ denote the history of reports along history h^t . Let Σ denote the set of all reporting strategies. The truth-telling strategy, which we denote by σ^* , is the one that prescribes reporting the true type at each and every node: for all h^t , $\sigma_t^*(h^t) = h_t$. Define

$$W(\sigma|x) = \sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} \left[u(c_t(\sigma^t(h^t))) - v \left(\frac{l_t(\sigma^t(h^t)) w_t(\sigma_t(h^t))}{w_t(h_t)} \right) \right].$$

as the expected discounted value of using reporting strategy σ given an allocation x .

An allocation x is incentive compatible if and only if for all $\sigma \in \Sigma$,

$$W(\sigma^*|x) \geq W(\sigma|x). \tag{4}$$

Simplifying the Incentive Constraints. Following Fernandes and Phelan (2000), without loss of generality, we restrict attention to the set of reporting strategies that has lying only at a single node.¹² This allows us to replace (4) with a sequence of *temporary incentive constraints*, one for each node.

An allocation x is incentive compatible if and only if in any period t and at any node

¹²Temporary incentive constraints were first shown to be necessary and sufficient for incentive compatibility by Green (1987) in the context of a dynamic moral hazard problem where people face i.i.d. endowment shocks that are private information. Fernandes and Phelan (2000) generalized this result to environments with persistent private shocks. To be precise, we need to use two more assumptions to be able to use temporary incentive constraints. First, we need each possible skill history to be reached with strictly positive probability. Second, we need a transversality condition that is automatically satisfied if we assume that instantaneous utility is bounded.

h^{t-1} and for all $h_t \in H$,

$$\begin{aligned}
& u(c_t(h^{t-1}, h_t)) - v(l_t(h^{t-1}, h_t)) + \sum_{m=t+1}^{\infty} \sum_{\{h^m \in H^m | h^m \succ h^t\}} \pi_m(h^m) \beta^{m-t} [u(c_m(h^m)) - v(l_m(h^m))] \\
& \geq u(c_t(h^{t-1}, h_t^o)) - v\left(\frac{l_t(h^{t-1}, h_t^o) w_t(h_t^o)}{w_t(h_t)}\right) \\
& + \sum_{m=t+1}^{\infty} \sum_{\{h^m \in H^m | h^m \succ h^t\}} \pi_m(h^m) \beta^{m-t} [u(c_m(\tilde{h}^m)) - v(l_m(\tilde{h}^m))], \tag{5}
\end{aligned}$$

where h_t^o is the complement of h_t in the set H and $\tilde{h}^m = (h^{t-1}, h_t^o, h_{t+1}, \dots, h_m)$.

In the rest of the paper, we use (5) to represent incentive compatibility.

Social Planning Problem. The social planning problem that defines the constrained efficient allocation is

$$\max_x \sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} [u(c_t(h^t)) - v(l_t(h^t))]$$

subject to

$$(1), (2), (3), \text{ and } (5).$$

The allocation that solves the social planning problem is called the *constrained efficient* allocation and is denoted with an asterisk throughout the paper.

4 Optimality of Differential Capital Taxation

In this section, we lay out the economic mechanism that calls for differential capital taxation.

4.1 Capital Return Wedge

In the standard two-sector growth model, a given level of aggregate savings is allocated between the two types of capital to equate their physical marginal returns. The proposition below shows that at the constrained efficient allocation, this is not true, meaning it is optimal to create a wedge between the physical returns to the two types of capital. This result is the basis for differential taxation of capital: to create the wedge in the market equilibrium, the two types of capital should be taxed differently.

Proposition 1. *At the constrained efficient allocation, in any period $t \geq 2$,*

$$\tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) = \tilde{F}_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) + \frac{1}{\lambda_t^*} X_t^*,$$

where

$$X_t^* = \sum_{\{h^t \in H^t\}} \mu_t^*(h^t) v' \left(\frac{l_t^*(h^{t-1}, h_t^o) w_t^*(h_t^o)}{w_t^*(h_t)} \right) l_t^*(h^{t-1}, h_t^o) \frac{\partial \frac{w_t^*(h_t^o)}{w_t^*(h_t)}}{\partial K_{e,t}^*}$$

and λ_t and $\mu_t(h^t)$ are Lagrange multipliers on period t feasibility constraint and the incentive constraint at history h^t .

Proof. Under Assumptions 1 and 2, the result follows from the first-order conditions with respect to two capital types:

$$\begin{aligned} (K_{s,t}) : & -\lambda_{t-1}^* + \lambda_t^* \tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) = 0, \\ (K_{e,t}) : & -\lambda_{t-1}^* + \lambda_t^* \tilde{F}_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) \\ & + \sum_{\{h^t \in H^t\}} \mu_t^*(h^t) v' \left(\frac{l_t^*(h^{t-1}, h_t^o) w_t^*(h_t^o)}{w_t^*(h_t)} \right) l_t^*(h^{t-1}, h_t^o) \frac{\partial \frac{w_t^*(h_t^o)}{w_t^*(h_t)}}{\partial K_{e,t}^*} = 0. \end{aligned}$$

□

Because of capital-skill complementarity, increasing the level of equipment capital in a given period increases the wage ratio in that period and hence affects the temporary incentive constraints in that period. From a planning perspective, this means that increasing equipment capital has an extra marginal return, X_t^*/λ_t , in addition to the physical return, $\tilde{F}_{2,t}$. Since structure capital is neutral, changing its level does not affect the incentive constraints, and hence its only return is its physical return, $\tilde{F}_{1,t}$. The proposition then just collects the total return to equipment capital and equates it to the return to structure capital at the constrained efficient allocation.

Observe that two assumptions are key for the capital return wedge result. First, if there was no complementarity between equipment capital and skilled labor, it would be efficient to equate the physical marginal returns to two types of capital. Second, if skills were not private information but publicly known, then there would be no incentive constraints, and hence, $X_t^* = 0$, and the optimal capital return wedge would be zero. Intuitively, if skill types were known, the government could condition taxes on skill types, which is nondistortionary. In that case, the government could redistribute via type-specific lump-sum taxes at zero efficiency cost and would not need the additional distortionary channel of redistribution, which works through the capital return wedge.

4.2 Optimal Differential Capital Taxes

In this section, we provide a link between the optimality of the capital return wedge and the optimality of differential capital taxation. We do so in two steps. First, in Proposition 2 we provide intertemporal characterizations of the constrained efficient allocation. Second, in Corollary 1 we characterize the properties of the optimal wedges (distortions) that a planner would have to create in order to implement the constrained efficient allocation in a competitive markets environment in which people were allowed to save through both types of capital.

Proposition 2. *At the constrained efficient allocation, in any period t and following any history h^t ,*

$$\frac{1}{u'(c_t^*(h^t))} = \frac{1}{\beta \tilde{F}_{1,t+1}^*} E_t \left\{ \frac{1}{u'(c_{t+1}^*(h^{t+1}))} | h^t \right\} \quad (6)$$

and

$$\frac{1}{u'(c_t^*(h^t))} = \frac{1}{\beta (\tilde{F}_{2,t+1}^* + X_{t+1}^*/\lambda_{t+1}^*)} E_t \left\{ \frac{1}{u'(c_{t+1}^*(h^{t+1}))} | h^t \right\}. \quad (7)$$

Proof. Under Assumption 1, equation (6) follows directly from Golosov, Kocherlakota, and Tsyvinski (2003). Equation (7) is obtained by using Proposition 1 in equation (6). \square

Equation (6) describes the efficient way of allocating consumption across two consecutive periods using structure capital. It is the standard inverse Euler equation first derived by Rogerson (1985) and then generalized to dynamic economies with a general stochastic process governing the evolution of skills by Golosov, Kocherlakota, and Tsyvinski (2003). Intuitively, equation (6) holds because its left-hand side defines how much resources are saved by decreasing period t utility by one unit at node h^t . The right-hand side is the period t resource cost of increasing utility by $1/\beta$ units in both contingencies in period $t + 1$. These values have to be equal at the constrained efficient allocation. Otherwise, by increasing or decreasing (depending on the direction of the inequality) savings at node h^t , the planner could achieve the same welfare level in an incentive-compatible way at a lower total resource cost. That would contradict the (constrained) efficiency of the original allocation.

Observe that we use $\tilde{F}_{1,t+1}^*$ in equation (6) when discounting the period $t + 1$ cost of increasing period $t + 1$ utility to period t . We do so because here we assume that the planner is using $K_{s,t+1}$ to transfer resources between periods. In equation (7), the planner is again comparing period t and $t + 1$ costs of increasing agents' utility by one unit. However, now resources are transferred between periods using $K_{e,t+1}$, and there is an additional incentive cost (or benefit depending on which incentive constraints bind) of changing the level of $K_{e,t+1}$. This additional cost is equal to $X_{t+1}^*/\lambda_{t+1}^*$, which is reflected in the overall marginal return of equipment capital, $\tilde{F}_{2,t+1}^* + X_{t+1}^*/\lambda_{t+1}^*$.

The corollary below establishes the implications of Proposition 2 for the optimal taxation of different types of capital. First, we define intertemporal wedges in each period t for both types of capital $i \in \{s, e\}$ as

$$\tau_{i,t}(h^t) = 1 - \frac{u'(c_t(h^t))}{\beta \tilde{F}_{i,t+1} E_t \{u'(c_{t+1}(h^{t+1})) | h^t\}}. \quad (8)$$

The intertemporal wedge for capital of type i at history h^t measures the efficient distortion that the planner needs to create in the intertemporal allocation of resources between current consumption and saving through capital asset i at that history. The efficient (i.e., optimal) distortion is computed by comparing the intertemporal marginal rate of substitution at h^t with the marginal rate of transformation through capital of type i in period t . We denote the optimal wedges by an asterisk.

Corollary 1. *The optimal wedge on structure capital in any period t and history h^t satisfies*

$$\tau_{s,t}^*(h^t) \geq 0. \quad (9)$$

The inequality is strict if and only if there is no $j \in H$ such that $\pi_{t+1}(h^t, j | h^t) = 1$.

The optimal wedge on equipment capital in any period t and history h^t is given by

$$[1 - \tau_{e,t}^*(h^t)] = [1 - \tau_{s,t}^*(h^t)] \cdot \frac{\left(\tilde{F}_{2,t+1}^* + X_{t+1}^*/\lambda_{t+1}^*\right)}{\tilde{F}_{2,t+1}^*}. \quad (10)$$

Proof. Inequality (9) follows from equation (6) and the conditional version of Jensen's inequality. Equation (10) is a direct consequence of Proposition 2 and the definitions of intertemporal wedges in equation (8). \square

The first part of the corollary is standard in the NDPF literature and says that the intertemporal wedge on structure capital is positive if there is skill risk. The intuition for this result is the following. If we allow an agent to borrow and save at the marginal rate of return to structure capital, he will save above the efficient level in a given period. In the next period, he will work less than socially optimal if he turns out to be of the skilled type. To prevent this double deviation, it is optimal to discourage savings (i.e. implement a positive intertemporal wedge). The government can achieve a positive wedge with a positive tax on structure capital. Naturally, there is no reason to tax structure capital if there is no skill risk.

The second part of the corollary shows that the intertemporal wedge on equipment capital can be decomposed into two parts. The government wants to discourage savings in equipment

capital for the same reasons as it wants to discourage savings in structure capital, which is captured in the first term on the right-hand side of equation (10). The second term on the right-hand side of equation (10) is present in order to create a wedge between the returns to the two types of capital, which is optimal as shown in Proposition 1. The presence of this term implies that generically the optimal wedges on the two types of capital are *different* in any period and history, which establishes the optimality of differential taxation of capital.

Observe that the second term on the right-hand side of equation (10) is history independent, which implies that the additional tax (or subsidy) on equipment capital can be applied at the firm level. We use this fact in Section 5, in which we describe how the constrained efficient allocation can be implemented in a competitive equilibrium environment with taxes.

4.3 The Relevant Case

We now impose two natural assumptions that make it possible to proceed further with our characterization results. In particular, we are able to sign the capital return wedge and determine whether the intertemporal wedge is higher for structure capital or equipment capital. The first assumption concerns the wages of the two types of labor, the second the pattern of binding incentive constraints.

Since wages are endogenous in our model, it is possible that skilled agents earn a lower wage rate than unskilled agents in a certain period (e.g., if the level of equipment capital is low or the supply of skilled labor is too high). However, this is not empirically relevant and, furthermore, it does not occur in any of our numerical simulations in Section 6. In the rest of the paper, we assume that in any period, skilled workers earn higher wages than unskilled workers. Formally:

Assumption 4. *In any period $t \geq 1$, $w_{s,t}^* > w_{u,t}^*$.*

There is no theoretical result that establishes the pattern of binding incentive constraints for general skill processes in dynamic Mirrleesian environments. However, with Assumption 4 and with the government providing insurance and redistribution, it is natural to assume that the only incentive constraints that bind are those that prevent the skilled from pretending to be unskilled. We call such incentive constraints *downward incentive constraints*.¹³ We find it useful to formalize this pattern of binding incentive constraints with the following assumption.

Assumption 5. *Suppose only downward incentive constraints bind.*

¹³In our environment, one can establish that this is the pattern of binding incentive constraints when skills are i.i.d. or permanent. We assume that skills are permanent in our quantitative analysis in Section 6.

Lemma 1. *Suppose Assumption 5 holds. Then, at the constrained efficient allocation, in any period $t \geq 1$,*

$$X_{t+1}^* < 0.$$

Proof. Under Assumption 5,

$$X_{t+1}^* = \sum_{\{h^t \in H^t\}} \mu_{t+1}^*(s|h^t) v' \left(\frac{l_{t+1}^*(h^t, u) w_{u,t+1}^*}{w_{s,t+1}^*} \right) l_{t+1}^*(h^t, u) \frac{\partial \frac{w_{u,t+1}^*}{w_{s,t+1}^*}}{\partial K_{e,t+1}^*}.$$

Using the definition of wages in equation (2), Assumption 2 implies

$$\frac{\partial \frac{w_{u,t+1}^*}{w_{s,t+1}^*}}{\partial K_{e,t+1}^*} < 0,$$

implying the result. □

Remember that X_{t+1}^* represents the effect of increasing equipment capital on the incentives. By Assumption 2, we know that increasing equipment capital in period $t + 1$ increases the wage premium in that period. This tightens the incentive constraints of the people who are skilled in period $t + 1$ and relaxes the incentive constraints of the people who are unskilled in period $t + 1$. By Assumption 5, the incentive constraints of the unskilled are slack, implying that the only effect of increasing equipment capital is tightening the incentive constraints of the period $t + 1$ skilled agents. Hence, the total effect of increasing $K_{e,t+1}$ on incentives is negative (i.e., $X_{t+1}^* < 0$).

Proposition 3. *Suppose Assumption 5 holds. Then, in the constrained efficient allocation, in any period $t \geq 1$ and history h^t ,*

$$\tilde{F}_1(K_{s,t+1}^*, K_{e,t+1}^*, L_{s,t+1}^*, L_{u,t+1}^*) < \tilde{F}_2(K_{s,t+1}^*, K_{e,t+1}^*, L_{s,t+1}^*, L_{u,t+1}^*),$$

$$\tau_{e,t}^*(h^t) > \tau_{s,t}^*(h^t).$$

Proof. The first part of the proposition follows from Proposition 2 and Lemma 1, whereas the second part follows from Corollary 1 and Lemma 1. □

The first part of this proposition says that the physical return on equipment capital should be higher than the physical return on structure capital. This result is intuitive: by decreasing the level of equipment capital, the planner decreases the skill premium and thus indirectly redistributes from the skilled to the unskilled agents. Decreasing the level of equipment capital increases its return due to diminishing marginal returns. This intuition

shows that there is an extra reason to depress equipment capital accumulation relative to structure capital, which implies that the wedge on equipment capital should be higher than that on structure capital. This is the second part of the proposition.

As we show formally in Section 5, to create a higher wedge on equipment capital in a competitive market environment in which agents hold and trade capital, the government should tax equipment capital at a higher rate than structure capital. Note that this result is in contrast with the actual U.S. tax code in which structure capital is taxed at a higher rate than equipment capital.

4.4 Intratemporal Wedges

The intratemporal (labor) wedge measures the distortions in the consumption labor decision by comparing the marginal rate of substitution between consumption and labor with the marginal product of labor (i.e., the wage rate):

$$\tau_{y,t}(h^t) = 1 - \frac{v'(l_t(h^t))}{w_t(h_t)u'(c_t(h^t))}.$$

The famous no distortion at the top result proven originally by Sadka (1976) and Seade (1977) states that in a static Mirrleesian economy, if the distribution of skills has a finite support, then the consumption-labor decision of the agent with the highest skill level should not be distorted. A counterpart of this proposition holds in a standard dynamic Mirrleesian model with discrete skill types: in any period, an agent who is currently of the highest skill level should face a zero labor wedge. This result does not depend on the stochastic process governing skills but requires that only the downward incentive constraints bind.¹⁴

In this section, we show that the no distortion at the top result does not hold in the presence of capital-skill complementarity even if only the downward constraints bind. This result is in line with Stiglitz (1982), who shows that when two types of labor are imperfect substitutes, then the top agents' labor supply decision should be distorted.¹⁵ In particular, Stiglitz shows that the top agents' labor income should be subsidized. The intuition is that, under imperfect substitutability between the two types of labor, increasing the labor supply of the skilled agents decreases the wage ratio of the skilled relative to the unskilled agents. Thus, increasing skilled labor supply creates indirect redistribution. In order to induce the skilled agents to work more, the government has to subsidize their labor supply at the margin.

¹⁴Farhi and Werning (2013) and Golosov, Troshkin, and Tsyvinski (2011) prove this result for dynamic economies with continuous support under certain assumptions on the primitives. Huggett and Parra (2010) prove it for a dynamic economy in which skill types are permanent.

¹⁵Rothschild and Scheuer (2012) generalize the subsidy result in Stiglitz (1982) to an economy with occupational choice and a continuum of types.

Mechanically, in the planning problem, increasing the skilled labor supply in a given period decreases the skill premium, which relaxes the skilled agents' incentive constraints associated with that period. With capital-skill complementarity in the production function, skilled and unskilled labor become imperfect substitutes, and hence Stiglitz's result is generalized to dynamic economies with capital accumulation.¹⁶

Proposition 4. *Suppose Assumption 5 holds. In any period $t \geq 1$, the optimal labor wedge of the agent who is currently skilled is negative.*

Proof. Relegated to Appendix A. □

5 Implementation

In this section, we show how the constrained efficient allocation can be implemented in an incomplete markets environment with taxes. We say that a tax system implements the constrained efficient allocation in a market if the constrained efficient allocation arises as an equilibrium of this market arrangement under the given tax system. The constrained efficient allocation can be implemented in many different ways. We provide an implementation in which the tax system mimics the actual U.S. tax code in the sense that capital tax differentials are created at the firm level through depreciation allowances that differ from actual economic depreciation. Therefore, creating the optimal capital tax differentials would not require further complications to the U.S. tax code.

We begin by describing the market arrangement. We assume that markets are incomplete in that agents can only trade uncontingent claims to future consumptions (i.e. they can save and borrow at a net risk-free rate, r_t). We denote an agent's savings as a_t . There is a static representative firm that rents capital at a net interest rate r_t and labor to produce the output good.¹⁷ The wage rates for skilled and unskilled are $w_{s,t}$ and $w_{u,t}$ and are taken as given by the firm.

Government and Taxes. There is a government that needs to finance $\{G_t\}_{t=1}^{\infty}$, an exogenously given sequence of government consumption. The government taxes consumers' savings and labor income (in a nonlinear and history-dependent way). The government also

¹⁶Grochulski and Piskorski (2010) prove that the no distortion at the top result does not hold in all but the last period in a finite horizon Mirrleesian economy with full persistence in innate skills and human capital accumulation.

¹⁷The assumption that the firm does not accumulate capital is innocuous and is made for convenience only. This setup is equivalent to one in which the firm, instead of the consumers, accumulates capital and makes the capital accumulation decisions.

taxes the firms' capital income net of depreciation. The statutory depreciation allowance can differ from economic depreciation, as in the U.S. tax code. This is how the government implements differential taxation of capital income, both in the U.S. economy and in our implementation.

Taxes on consumers are specified following Kocherlakota (2005). Let $\tau_{y,t} : \mathbb{R}_+^t \rightarrow \mathbb{R}$ denote the labor tax schedule, where $\tau_{y,t}(y^t)$ is the labor income tax an agent with labor income history $y^t = (y_1, \dots, y_t)$ pays in period t . Labor income in history h^t is defined as $y_t(h^t) = w_t(h_t) \cdot l_t(h^t)$. There are also linear taxes on people's asset holdings, which may depend on their income history. Letting $\tau_{a,t} : \mathbb{R}_+^t \rightarrow \mathbb{R}$ denote linear tax rates on asset holdings, an agent with income history y^t pays $\tau_{a,t}(y^t)(1+r_t)a_t(h^{t-1})$, where $a_t(h^{t-1})$ denotes the agent's asset holdings.¹⁸

Unlike in Kocherlakota (2005), the government also has access to a linear, history-independent tax on firm profits, denoted by $\tau_{f,t} : \mathbb{R}_+ \rightarrow \mathbb{R}$. The corporate tax code also includes statutory depreciation allowances, $(\bar{\delta}_{s,t}, \bar{\delta}_{e,t})_{t=1}^\infty$, where the firm is allowed to deduct $\bar{\delta}_{i,t}K_{i,t}$ from its tax base in period t .

Consumer's Problem. Taking prices $(r_t, w_{s,t}, w_{u,t})_{t=1}^\infty$ and taxes $(\tau_{y,t}, \tau_{a,t})_{t=1}^\infty$ as given, a consumer solves

$$\max_{c,y,a} \sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} \left[u(c_t(h^t)) - v\left(\frac{y_t(h^t)}{w_t(h^t)}\right) \right]$$

subject to

$$c_t(h^t) + a_{t+1}(h^t) \leq y_t(h^t) - \tau_{y,t}(y^t(h^t)) + [1 - \tau_{a,t}(y^t(h^t))] (1 + r_t)a_t(h^{t-1}),$$

$$a_1 = K_{s,1}^* + K_{e,1}^*,$$

c, y are nonnegative.

Firm's Problem. Taking $(r_t, w_{s,t}, w_{u,t})_{t=1}^\infty$ and taxes $(\tau_{f,t}, \bar{\delta}_{s,t}, \bar{\delta}_{e,t})_{t=1}^\infty$ as given, the firm solves

$$\max_{K_{e,t}, K_{s,t}, L_{s,t}, L_{u,t}} \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r_t)(K_{s,t} + K_{e,t}) - \frac{w_{s,t}}{z_s} L_{s,t} - \frac{w_{u,t}}{z_u} L_{u,t} - \tau_{f,t} \Pi_{f,t},$$

where $\Pi_{f,t}$ is the firm's capital income net of depreciation allowances in period t and is given

¹⁸One can show that the government can implement the constrained efficient allocation by taxing only the return from capital $r_t a_t(h^{t-1})$, which would be more in line with what we observe in the U.S. tax code.

by

$$\Pi_{f,t} = \left[F(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \bar{\delta}_{s,t}K_{s,t} - \bar{\delta}_{e,t}K_{e,t} - \frac{w_{s,t}}{z_s}L_{s,t} - \frac{w_{u,t}}{z_u}L_{u,t} \right].$$

Notice that, as in the U.S. corporate tax code, we assume a flat tax on the firm's capital income net of depreciation allowances.

Equilibrium. Given a tax system $(\tau_{y,t}, \tau_{a,t}, \tau_{f,t}, \bar{\delta}_{s,t}, \bar{\delta}_{e,t})_{t=1}^{\infty}$, an equilibrium is an allocation for consumers, $(c_t(h^t), l_t(h^t), a_{t+1}(h^t))_{t=1}^{\infty}$, an allocation for the firm, $(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^{\infty}$, and prices $(r_t, w_{s,t}, w_{u,t})_{t=1}^{\infty}$ such that $(c_t(h^t), y_t(h^t), a_{t+1}(h^t))_{t=1}^{\infty}$ solves the consumer's problem, $(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^{\infty}$ solves the firm's problem, and markets clear:

$$\begin{aligned} \sum_{h^t \in H^t} \pi_t(h^t) c_t(h^t) + K_{s,t+1} + K_{e,t+1} + G_t &= \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \\ K_{s,t} + K_{e,t} &= \sum_{h^t \in H^t} \pi_t(h^t) a_t(h^{t-1}), \\ L_{s,t} = \sum_{\{h^t \in H^t | h_t = s\}} \pi_t(h^t) \frac{y_t(h^t)}{w_{s,t}} z_s, & \quad L_{u,t} = \sum_{\{h^t \in H^t | h_t = u\}} \pi_t(h^t) \frac{y_t(h^t)}{w_{u,t}} z_u. \end{aligned}$$

The government's period-by-period budget balance is implied by Walras' law:

$$\sum_{h^t \in H^t} \pi_t(h^t) [\tau_{a,t}(y^t(h^t))(1 + r_t)a_t(h^{t-1}) + \tau_{y,t}(y^t(h^t))] + \tau_{f,t}\Pi_{f,t} = G_t.$$

In what follows, we describe an optimal tax system that implements the constrained efficient allocation in the market setup described above. Before doing so, we provide a formal definition of our notion of implementation.

Implementation. A tax system $(\tau_{y,t}, \tau_{a,t}, \tau_{f,t}, \bar{\delta}_{s,t}, \bar{\delta}_{e,t})_{t=1}^{\infty}$ implements the constrained efficient allocation $(c_t^*(h^t), l_t^*(h^t), K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)_{t=1}^{\infty}$ if an allocation for consumers $(c_t^*(h^t), l_t^*(h^t), a_{t+1}(h^t))_{t=1}^{\infty}$ and an allocation for the firm $(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)_{t=1}^{\infty}$ jointly with the tax system and prices $(r_t, w_{s,t}, w_{u,t})_{t=1}^{\infty}$ constitute an equilibrium.

5.1 Optimal Tax System

First, we construct the optimal tax system. Then, we prove that it implements the constrained efficient allocation. Finally, we characterize its properties.

Optimal Tax System. We begin by describing optimal savings taxes. Set the taxes on

people's savings as

$$\begin{aligned}\tau_{a,t+1}^*(y^{t+1}) &= 1 - \frac{u'(c_t^*(h^t))}{\beta u'(c_{t+1}^*(h^{t+1})) \tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)}, \text{ if } y^{t+1} \in Y^{t+1*} \\ \tau_{a,t+1}^*(y^{t+1}) &= 1, \text{ if else,}\end{aligned}$$

where we define $y^t(h^t) = (y_m(h^m))_{m=1}^t$, $Y^{t*} = \{y^t : y^t = y^{t*}(h^t), h^t \in H^t\}$, and $\forall j$

$$y_t^*(h^{t-1}, j) = \frac{\partial F(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)}{\partial L_{j,t}} z_j l_t^*(h^{t-1}, j).$$

In words, Y^{t*} is the set of labor income histories observed at the constrained efficient allocation. Set labor income taxes such that if $y^t \in Y^{t*}$, then $\tau_{y,t}^*(y^t)$ and $a_{t+1}^*(h^t)$ are defined to satisfy the flow budget constraints every period:

$$c_t^*(h^t) + a_{t+1}^*(h^t) = y_t^*(h^t) - \tau_{y,t}^*(y^t) + [1 - \tau_{a,t}^*(y^t)] \tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) a_t^*(h^{t-1}).$$

If $y^t \notin Y^{t*}$, then

$$\tau_{y,t}^*(y^t) = 2y_t.$$

Finally, set

$$\begin{aligned}\tau_{f,t}^* &= 1 - \frac{F_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) - \delta_s}{F_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) - \delta_e} \\ \bar{\delta}_{s,t}^* &= r_t^* + \delta_s \\ \bar{\delta}_{e,t}^* &= \delta_e.\end{aligned}$$

Observe that at the corporate level, the optimal tax system works in the same way as the U.S. tax code. In the U.S. corporate tax code, there is a single statutory corporate tax rate, and differences between effective tax rates on equipment and structure capital income are created through differences in tax depreciation rules, as explained in detail in Section 2.

Proposition 5. *The optimal tax system described above implements the constrained efficient allocation.*

Proof. Relegated to Appendix A. □

Properties of Optimal Capital Taxes. Next we summarize and discuss the properties of optimal capital taxes.

1. Capital income is taxed twice, once at the consumer level via savings taxes and once at the corporate level (double taxation of capital).
2. The statutory corporate tax is strictly positive in the relevant case:

$$\tau_{f,t}^* = 1 - \frac{F_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) - \delta_s}{F_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) - \delta_e} > 0,$$

where the inequality comes from Proposition 3.

3. Even though there is a single statutory tax rate on all of corporate income, τ_t^f , the effective taxes on capital income at the firm level differ across different capital assets because of differences in the statutory depreciation allowances. The firm is allowed to expense all of its user cost of structure capital, whereas the depreciation allowance on equipment capital is equal to its economic depreciation. Using equation (2), these properties imply that the optimal effective *corporate* tax rate on structure capital is zero and the optimal effective *corporate* tax rate on equipment capital is equal to the statutory corporate tax rate $\tau_{f,t}^*$.¹⁹
4. Properties 2 and 3 imply that the government uses capital income taxes to collect revenue at the corporate level, as in the U.S. tax code.
5. Expected taxes on consumers' asset holdings are zero as in Kocherlakota (2005):

$$E_t \{1 - \tau_{t+1}^{a*}(y^{t+1})\} = E_t \left\{ \frac{u'(c_t^*(h^t))}{\beta u'(c_{t+1}^*(h^{t+1})) \tilde{F}_{1,t+1}^*} \right\} = 1.$$

This property follows from equation (6) in Proposition 2.

¹⁹Note that there are other implementations in which both capital types can be taxed positively as long as the tax rates create the efficient capital return wedge. In that case, the savings' tax that the consumers face would have to be adjusted.

6 Quantitative Analysis

The main goal of this section is to analyze the quantitative importance of differential taxation of capital in a version of our model calibrated to the U.S. economy. We assume that agents' skill types are permanent: if an agent enters the labor market skilled (unskilled), she remains so forever. Since there is no labor income risk in this environment, the only role of taxation is redistribution (along with financing government consumption).²⁰ Permanent skills is a natural assumption given that in the data we associate skill levels with educational attainment. In addition, there is empirical evidence that initial conditions account for a large part of the cross-sectional variation in lifetime earnings.²¹

We first calibrate the model parameters to the U.S. economy using a competitive equilibrium framework with the actual U.S. tax code and government consumption level. We then solve a social planning problem in which the planner “inherits” the capital levels from the competitive equilibrium and needs to finance the same level of government consumption. We solve for the whole time series of the constrained efficient allocation, thus taking into account the transition to a new steady state. We find that the optimal capital taxes on equipment capital should be higher than those on structure capital. Specifically, in our benchmark calibration, the optimal tax differential increases from 27.6% in the first period to 39.5% in the steady state. We also assess the welfare consequences of differential capital taxation. We find that the welfare gains of optimal differential capital taxation can be as high as 0.4% in terms of lifetime consumption.

The assumption of permanent skill types also allows for a sharper qualitative characterization of the constrained efficient allocation and the associated optimal taxes. Therefore, before going to the quantitative analysis, in Section 6.1 we provide a sharper characterization of the qualitative properties of the optimal tax system under permanent skill types. The most important result of this section is the optimality of positive taxes on equipment capital in the long run.

6.1 Social Planning Problem and Long-Run Capital Taxes

We first define the social planning problem for the environment with permanent skill types. We assume a utilitarian social planner who puts an equal weight on all agents. We fur-

²⁰A related exercise has been performed in Werning (2007), who characterizes the constrained efficient allocation qualitatively in a standard neoclassical growth model with permanent types. Huggett and Parra (2010) compute the constrained efficient allocation in a partial equilibrium overlapping generations environment. Unlike the focus in these papers, we focus on the role of differential capital asset taxation.

²¹Keane and Wolpin (1997) estimate this number as high as 90%, Huggett, Ventura, and Yaron (2011) over 60%, and Storesletten, Telmer, and Yaron (2004) almost 50%.

ther assume that Assumption 4 is satisfied (i.e., the skilled agents' wage is higher than the unskilled agents' wage). The social planning problem (SPP) reads:²²

$$\begin{aligned} & \max_{\{(c_{j,t}, l_{j,t})_j, K_{s,t}, K_{e,t}, L_{u,t}, L_{s,t}\}_{t=0}^{\infty}} \sum_{j=u,s} \pi_j \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{j,t}) - v(l_{j,t})] \quad \text{s.t.} \\ \forall t : G_t + \sum_j \pi_j c_{j,t} + K_{s,t+1} + K_{e,t+1} & \leq \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \\ \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{s,t}) - v(l_{s,t})] & \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[u(c_{u,t}) - v\left(\frac{l_{u,t} w_{u,t}}{w_{s,t}}\right) \right]. \end{aligned}$$

To simplify notation, we denote the consumption and labor allocation of an agent of (permanent) type j in period t by $c_{j,t}$ and $l_{j,t}$.

In this problem, the only private information that people have is their permanent skill type. People report their type only once in the first period. As a result, the planner faces only two incentive constraints. In the problem defined above, we omit the incentive constraint that prevents unskilled types from pretending to be skilled types because that incentive constraint never binds. This result follows from Assumption 4 and the fact that the utilitarian social planner puts equal weight on each agent.

Proposition 6 provides a sharper characterization of the optimal intertemporal wedges for the permanent type case.

Proposition 6. *Suppose skill types are permanent and Assumption 4 holds. Then the optimal intertemporal wedges on the two types of capital satisfy the following properties.*

1. *In all periods, the optimal wedge on equipment capital is strictly positive, whereas the optimal wedge on structure capital is zero.*
2. *If the optimal allocation converges to a steady state, then the optimal wedge on equipment capital is strictly positive at the steady state.*

Proof. Relegated to Appendix A. □

²²Note that factor prices are endogenous in this SPP (i.e., we take general equilibrium effects into account). We would not be able to assess the role of differential capital taxation in a partial equilibrium environment, which most quantitative papers in the NDPF literature consider. As Farhi and Werning (2012) show, considering general equilibrium effects is important even with a standard Cobb-Douglas production function. They find that welfare gains of a particular partial tax reform are greatly overestimated if general equilibrium effects are not taken into account. Another exception in this regard is Fukushima (2010), who computes the constrained efficient allocation in a general equilibrium overlapping generations environment with a specific assumption on the persistence of skill shocks.

Part 1 of Proposition 6 calls for positive taxation of equipment capital and zero taxation of structure capital in every period.²³ To understand this result, remember that by Assumption 1, a change in the level of structure capital does not affect the skill premium. Therefore, there is no indirect redistribution motive to distort structure capital accumulation. The only reason why it might be optimal to tax structure capital in this environment is to discourage the double deviation of saving too much and reporting to be of a different skill type in the next period. With permanent skill types, agents do not report after the initial period, and hence, this reason disappears. In contrast, taxing equipment capital has the extra benefit of decreasing the skill premium, thus providing indirect redistribution. Therefore, the planner finds it optimal to tax equipment capital. Observe that since there is no need to tax consumers' savings, all capital taxation can be done at the firm level.

Part 2 of Proposition 6 says that the optimal wedge on equipment capital is positive in steady state. This result is interesting because it shows that the indirect redistribution channel calls for taxing equipment capital not only in the short run but also in the long run. This result is in contrast with the famous Chamley-Judd result of the long-run optimality of zero capital taxation in the Ramsey literature.²⁴ Observe that it is meaningful to discuss steady-state taxation because skill types are permanent and, hence, steady states can exist.

6.2 Calibration

To calibrate the parameters in the social planning problem, we assume that the steady state of the competitive equilibrium defined in Section 5 represents the current U.S. economy. In this stationary competitive equilibrium (SCE), we use the current U.S. tax system rather than the optimal taxes. We first fix a number of parameters to values from the data or from the literature. We then calibrate the remaining parameters so that the SCE matches the U.S. data along selected dimensions.

We assume that the period utility function takes the form

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

One period in our model corresponds to one year. In the benchmark case, we use $\sigma = 2$ and $\gamma = 1$. These are within the range of values that have been considered in the literature.

²³The optimality of not taxing structure capital is related to Werning (2007), who shows that zero capital taxation is optimal in every period in a dynamic Mirrleesian model with standard Cobb-Douglas production function, a single capital good, and permanent types.

²⁴See Chamley (1986) and Judd (1985). Note that in the standard Ramsey problem with linear capital and labor income taxes and permanent types, the optimal tax on both types of capital is zero in the long run even with our production function that features equipment capital-skill complementarity.

Table 1: Benchmark Parameters

Parameter	Symbol	Value	Source
Relative risk aversion	σ	2	
Inverse Frisch elasticity	γ	1	
Structure capital depreciation rate	δ_s	0.056	GHK
Equipment capital depreciation rate	δ_e	0.124	GHK
Share of structure capital in output	α	0.117	KORV
Measure of elasticity of substitution between equipment capital K_e and unskilled labor L_u	η	0.401	KORV
Measure of elasticity of substitution between equipment capital K_e and skilled labor L_s	ρ	-0.495	KORV
Tax on labor income	τ_y	0.27	HSV
Tax on consumers' capital income	τ_a	0.15	U.S. tax code
Corporate tax on structure capital income	τ_s^f	0.32	Gravelle (2011)
Corporate tax on equipment capital income	τ_e^f	0.26	Gravelle (2011)
Government consumption	G/Y	0.16	NIPA
Relative supply of skilled workers	p_s/p_u	0.778	U.S. Census
Share of skilled workers' wealth	ζ	0.686	U.S. Census

The acronyms GHK, KORV, and HSV stand for Greenwood, Hercowitz, and Krusell (1997), Krusell, Ohanian, Ríos-Rull, and Violante (2000), and Heathcote, Storesletten, and Violante (2010). NIPA stands for the National Income and Product Accounts.

We further assume that the production function takes the same form as in Krusell, Ohanian, Ríos-Rull, and Violante (2000):

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

Krusell, Ohanian, Ríos-Rull, and Violante (2000) estimate α , ρ , η , and we use their estimates, but they do not estimate ω and ν . We calibrate these parameters to U.S. data, as we explain in detail below. One can show that this production function satisfies Assumptions 1 – 3 if $\eta > \rho$, which is what Krusell, Ohanian, Ríos-Rull, and Violante (2000) find.

As for government policies, we assume that the government consumption-to-output ratio equals 16%, which is close to the average ratio in the United States during the period 1980 – 2012, as reported in the National Income and Product Accounts (NIPA) data.²⁵ We follow Heathcote, Storesletten, and Violante (2010) and assume a flat labor income tax rate $\tau_y = 27\%$ (for a discussion of the construction of this number, see Domeij and Heathcote (2004)). Gravelle (2011) documents that because of differences in tax depreciation

²⁵We find that including government consumption in the model is important for its quantitative properties. However, government consumption is mostly omitted in the quantitative exercises in the NDPF literature. An exception is Fukushima (2010), who uses a similar number for the government consumption-to-output ratio.

Table 2: Benchmark Calibration Procedure

Parameter	Value	Target	Data and SCE	Source
ω	0.477	Labor share	2/3	NIPA
β	0.985	Capital-to-output ratio	2.9	NIPA and FAT
ν	0.657	Skill premium $\frac{w_s}{w_u}$	1.8	HPV
ϕ	67.8	Labor supply	1/3	

The acronym HPV stands for Heathcote, Perri, and Violante (2010). NIPA stands for the National Income and Product Accounts, and FAT stands for the Fixed Asset Tables.

rates, the effective tax rates on structure capital and equipment capital differ at the firm level. Specifically, the effective corporate tax rate on structure capital $\tau_s^f = 32\%$, and the effective corporate tax rate on equipment capital $\tau_e^f = 26\%$. The capital income tax rate at the consumer level $\tau_a = 15\%$ as in the U.S. tax code. This implies an overall tax on structure capital $\tau_s = 1 - .85 \cdot (1 - .32) = 42.2\%$ and an overall tax on equipment capital $\tau_e = 1 - .85 \cdot (1 - .26) = 37.1\%$. These numbers are in line with a 40% tax, which Domeij and Heathcote (2004) report for the aggregate capital stock. We set the ratio of skilled to unskilled agents to be consistent with the 2011 US Census data.

Notice that the stationary competitive equilibrium is not unique in our environment with permanent types. Let us denote the steady-state asset holdings of a skilled agent by a_s and of an unskilled agent by a_u . Given aggregate capital levels K_s, K_e consistent with the SCE, any asset distribution of the form $p_s a_s = \zeta(K_s + K_e)$ and $p_u a_u = (1 - \zeta)(K_s + K_e)$ with $\zeta \in (0, 1)$ can arise in the SCE. We pick a ζ so that the SCE asset distribution matches the observed asset distribution between skilled and unskilled agents in the 2010 U.S. Census data. Table 1 summarizes our benchmark parameters.

This leaves us with several parameters to be determined. We cannot identify z_u and z_s separately from the remaining parameters of the production function and therefore set them to $z_u = z_s = 1$. We then calibrate the parameter that controls the income share of equipment capital ω , the parameter that controls the income share of unskilled labor ν , the labor disutility parameter ϕ , and the discount factor β . We calibrate these parameters so that (i) the labor share equals 2/3 (approximately the average labor share in 1980 – 2010 as reported in the NIPA data), (ii) the capital-to-output ratio equals 2.9 (approximately the average of 1980 – 2011 as reported in the NIPA and FAT data), (iii) the skill premium (i.e., the ratio of skilled to unskilled wages $\frac{w_s}{w_u}$) equals 1.8 (as reported by Heathcote, Perri, and Violante (2010) for the 2000s), and (iv) the aggregate labor supply in steady state equals 1/3 (as is common in the macro literature). Table 2 summarizes our calibration procedure.

6.3 Quantitative Results

In this section, we discuss the quantitative properties of the optimal tax system for the U.S. economy. We do so by solving the SPP defined above with parameters calibrated in Section 6.2 to match the U.S. economy.²⁶ We solve the SPP assuming that the economy converges to a steady state in 200 periods. We verify that changing the number of periods does not affect the results. In other words, the economy gets very close to steady state long before period 200.

We divide the discussion of our results into four parts. First, we compare the steady-state properties of the optimal tax system to the current U.S. tax system. Then we discuss the evolution of the optimal taxes along the transition path to the new steady state. Next, we discuss the welfare consequences of differential capital asset taxation. Finally, we provide a brief summary of our sensitivity analysis results.

Steady-State Comparison. In this section, we discuss the properties of the optimal tax system at steady state and compare it with the current U.S. tax system. The first column of Table 3 summarizes the U.S. tax system, which we use when solving for the SCE. The second column reports its counterpart in the optimal tax system. The first two rows of Table 3 report capital income taxes net of depreciation.²⁷ We find that the equipment capital tax τ_e is substantial at the steady state of the solution to the SPP. It is 39.54% – that is, 39.54 percentage points higher than the tax on structure capital τ_s , which is zero. This is in contrast with the effective tax rates in the United States where structure capital is taxed by 5.1 percentage points more than equipment capital overall, combining capital income taxes at the consumer and corporate levels. As for the labor wedges, they are 27% for both types of labor in the SCE because we approximate the U.S. labor income tax code by a 27% linear tax. At the steady state of the solution to the SPP, the labor wedge for unskilled labor $\tau_y(u)$ is 26.6%, which is almost the same as in the SCE. The skilled labor wedge $\tau_y(s)$, on the other hand, is -11.14%. Both higher taxes on equipment capital and marginal subsidies on skilled labor are used to create indirect redistribution from the skilled

²⁶We further assume that the planner starts with the same initial capital levels as those in the SCE. Finally, following Conesa, Kitao, and Krueger (2009), we assume that the planner needs to finance the same level of government consumption as in the SCE.

²⁷We report capital income taxes net of depreciation rather than the capital wedges defined in Section 4 because they correspond to the taxes used in the U.S. tax code. With a slight abuse of notation, we denote these capital income taxes as we denoted the capital wedges. In the column labeled “SPP,” the capital income taxes are recovered from the constrained efficient allocation by using the following definition for each (permanent) type $j \in \{u, s\}$, capital type i , and period t : $\tau_{i,t}(j) \equiv 1 - \left(\frac{u'(c_{t,j})}{\beta u'(c_{t+1,j})} - 1 \right) / (F_{K_{i,t+1}} - \delta_i)$. Part 1 of Proposition 6 implies that these taxes only depend on time and not on agent type; therefore, we only report one number (time series).

to the unskilled.

Table 3: Steady-State Comparison of Wedges

	SCE	SPP
τ_e	37.10%	39.54%
τ_s	42.20%	0.00%
$\tau_y(u)$	27.00%	26.60%
$\tau_y(s)$	27.00%	-11.14%

Table 4 shows how the optimal tax system achieves indirect redistribution by comparing the allocations at the SCE and the SPP. The higher tax on equipment capital discourages the accumulation of equipment capital relative to structure capital at the SPP in comparison with the SCE. At the same time, the marginal subsidy on skilled labor income gives rise to a higher ratio of skilled to unskilled labor. Both labor and capital taxes decrease the skill premium at the SPP. This way, the planner provides indirect redistribution from the skilled to the unskilled.²⁸

Table 4: Steady-State Comparison of Allocations

	SCE	SPP
K_e/K_s	1.02	0.93
L_s/L_u	0.82	1.11
Skill premium	1.80	1.47

The marginal subsidy on skilled labor income seems to imply that there is redistribution from the unskilled to the skilled at the SPP. However, recall that optimal taxes are nonlinear in labor income. In this case, at a given income level, the average income tax can be very different from the marginal income tax.²⁹ As a consequence, a tax system can be progressive overall even though the marginal taxes are regressive. This is precisely what happens at the optimal tax system. To assess the overall progressivity of the optimal tax system, we compute a measure of average labor taxes that an agent has to pay at the steady state of the SPP. This measure is defined as $1 - \frac{c_j}{w_j l_j}$ for agents of type j , as in Farhi and Werning

²⁸Observe that the planner makes skilled agents work more not only to provide indirect redistribution but also for efficiency reasons.

²⁹Suppose, for example, that the tax formula for an agent with income \$200,000 is $T(y) = \$100,000 - 0.1 \cdot y$. This agent pays \$80,000 in taxes, implying an average tax of 40%, even though he receives a marginal subsidy of 10%.

(2013). We find that the optimal average labor taxes are progressive: 6% for the unskilled and 18% for the skilled. Therefore, the optimal labor taxes do provide redistribution from the skilled to the unskilled.

Transition. This section summarizes the evolution of the optimal taxes (wedges) along the transition to the new steady state. The left panel of Figure 1 shows that the optimal structure capital income tax (net of depreciation) is 0 and the equipment capital tax is positive in all periods. These properties are in line with Proposition 6. The equipment capital tax is growing over time. To understand this fact, we look at the evolution of the stocks of the two types of capital, shown in the left panel of Figure 2. It shows that both capital stocks are growing along the transition path.³⁰ As the quantity of equipment capital grows, so does the skill premium (see Figure 3). To discourage the growth of the skill premium, the planner finds it optimal to increase the tax on equipment capital.³¹

Optimal labor wedges are almost constant along the transition, as shown in the right panel of Figure 1. It is easy to prove that labor wedges would in fact be exactly constant in a version of our model without equipment capital-skill complementarity. Figure 1 suggests that the extra distortions in labor wedges arising from equipment capital-skill complementarity are approximately constant over time. Since skilled labor is subsidized, skilled agents work more than unskilled agents in each period, as shown in the right panel of Figure 2. As the economy grows, agents become richer, and because of the income effect, they decrease their labor supply even though labor wedges do not change much.

Figure 3 depicts the evolution of the optimal skill premium over time. First, the optimal skill premium is much lower in each period than in the U.S. data. This result suggests that the current U.S. tax system does not generate enough indirect redistribution. Second, the skill premium is increasing over time because the equipment capital level increases. This result implies that an increasing skill premium is optimal in a growing economy, even if the government cares about equality.

³⁰This is because the planner inherits an inefficiently low level of capital stock from the SCE. The capital stock at the SCE is low due to the inefficiently high overall level of capital taxes.

³¹Note that since equipment capital accumulation is discouraged by higher taxes, it grows at a lower rate than structure capital. However, as Figure 2 shows, in the initial period, the planner increases the amount of equipment capital much more than the amount of structure capital. The reason is that at the steady state of the SPP, the return to equipment capital is about 1% higher than the return to structure capital. However, this difference is about twice as large in the initial period due to the shift toward more skilled labor. The planner therefore finds it optimal to increase equipment capital more in order to correct the suboptimal equipment-structure capital mix that he inherited from the SCE.

Figure 1: Optimal Taxes/Wedges

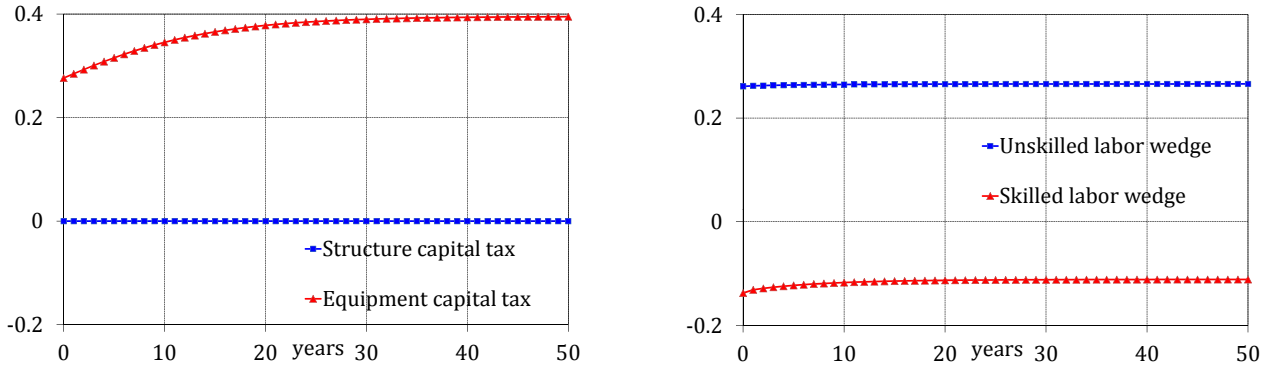


Figure 2: Factors of Production at the Constrained Efficient Allocation

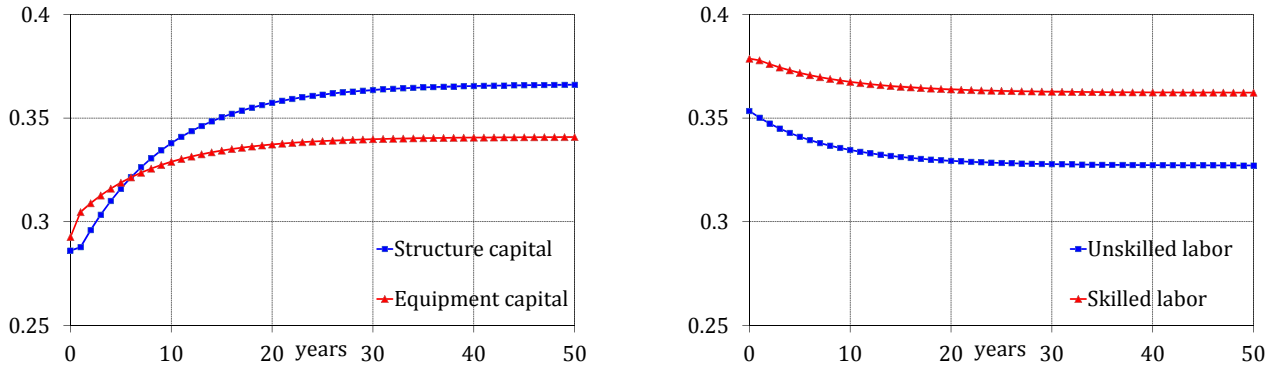
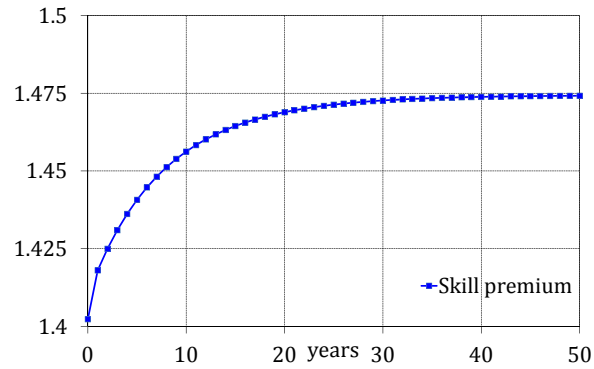


Figure 3: Skill Premium at the Constrained Efficient Allocation



Welfare Gains of Optimal Differential Taxation of Capital. To assess the quantitative importance of differential taxation of capital (DTC), we solve an additional version of the social planning problem. In this problem, the planner is unrestricted in his choice of labor taxes, but he is not allowed to tax the two types of capital differentially. We call this tax system the optimal *nondifferential taxation of capital* (optimal NDTC). We state the planning problem that gives rise to the optimal NDTC in Appendix B. We find that the welfare gain of optimal DTC relative to optimal NDTC is 0.05% for the benchmark parameters.³² This is a nonnegligible welfare gain of differential taxation. We find that the welfare gain can be as high as 0.35% for reasonable parameter values, as we discuss in more detail in the next section.

The welfare gains of reforming capital taxes are even larger if the starting point is a tax system in which structure capital is taxed more heavily than equipment capital. This is exactly what we see in the current U.S. tax system in which equipment capital is taxed 5% less than structure capital. To assess the welfare gains of reforming the U.S. capital taxes, we solve another version of the planning problem. In this problem, the planner is unrestricted in his choice of labor taxes, but he must use the capital income taxes as in the U.S. tax code. We call this tax system current DTC. We state the planning problem that gives rise to the current DTC in Appendix B. We find that the welfare gains of moving from current DTC to optimal DTC are 0.19% in the benchmark and can be as high as 0.40%.

The difference between current DTC and NDTC corresponds to the welfare gains associated with eliminating the current tax differentials in the spirit of the recent proposal of the Obama administration. Such reform implies a positive welfare gain because it goes in the right direction of taxing equipment capital more heavily than structure capital. In the benchmark, the welfare gain is sizable at 0.14% in terms of lifetime consumption.

Our results suggest considerable welfare gains of reforming the U.S. capital tax system. Observe that in all the exercises we perform, the planner is allowed to use any nonlinear labor taxes. This flexibility in the labor tax code allows the planner to use labor income taxes to substitute for optimal DTC for redistribution in scenarios in which he is not allowed to use DTC optimally. If labor taxes were not flexible (e.g., if they were fixed to the U.S. tax code), the welfare implications of having access to the optimal DTC would likely be larger. In this sense, our results provide a lower bound on the welfare gains of DTC.

We also compute who gains and who loses by the capital tax reforms. First, we find that moving from the optimal NDTC to the optimal DTC helps the unskilled and hurts

³²We measure the welfare gains of allocation x relative to allocation y as a fraction by which the consumption in allocation y would have to be increased in each date and state in order to make its welfare equal to the welfare of allocation x .

Table 5: Sensitivity Analysis

	Benchmark			Benchmark		
	1	2	4	2	2	2
σ	1	1	1	0.5	1	2
γ	1	1	1	0.5	1	2
Optimal taxes						
τ_e	24.39%	39.54%	54.84%	49.23%	39.54%	22.88%
τ_s	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
$\tau_y(u)$	18.77%	26.60%	33.75%	26.43%	26.60%	24.90%
$\tau_y(s)$	-5.29%	-11.14%	-21.23%	-16.46%	-11.14%	-5.08%
Welfare gains						
Optimal NDTC to opt. DTC	0.01%	0.05%	0.14%	0.10%	0.05%	0.01%
Current DTC to opt. DTC	0.24%	0.19%	0.21%	0.20%	0.19%	0.22%

the skilled. This is not surprising given that the optimal DTC features a higher equipment capital tax and a lower skill premium. This way, the planner is able to redistribute more efficiently in optimal DTC. Interestingly, we find that moving from current DTC to optimal DTC is Pareto improving (i.e., both types benefit). The reason is that the overall level of capital taxes at the current DTC is inefficiently high. Under the optimal DTC, capital taxes are lower and the economy accumulates more capital along the transition to the new (efficient) steady state (see Figure 2). This increases the productivity of both types of agents, and they both benefit from this reform.

Sensitivity Analysis. Table 5 summarizes the results of our sensitivity analysis. It reports the capital taxes as well as the labor wedges at the steady state of the SPP. In addition, it reports the welfare gains of the optimal DTC relative to the optimal NDTC and the welfare gains of the optimal DTC relative to the current DTC.

The first (vertical) panel of Table 5 reports the sensitivity of our results to changes in σ , the curvature of u (utility from consumption). With higher curvature, the planner wants to provide more redistribution. Therefore, the indirect redistribution channel becomes more important. Hence, as σ increases, the tax on equipment capital as well as the marginal subsidy to skilled labor increase. Similarly, the welfare gain of optimal DTC relative to optimal NDTC increases. With $\sigma = 4$, the welfare gains of optimal differential capital taxation are 0.14% in terms of lifetime consumption – three times larger than in the benchmark.

The second (vertical) panel of Table 5 reports the sensitivity of our results to changes in γ , the curvature of v (disutility from labor). We find that as γ becomes smaller, the tax on equipment capital and the skilled labor subsidy, as well as the welfare gain of optimal DTC relative to optimal NDTC increase. To understand this result, consider a world with

linear disutility of labor and no incentive constraints. In such a world, the planner would equalize the marginal products to the two types of labor, implying an optimal wage ratio of one. This implies that as γ decreases, the planner would like to bring the marginal products of labor closer to each other. This can be achieved at the constrained efficient allocation in two ways: higher subsidies on skilled labor and higher equipment capital taxes. As we document in Table 5, the planner uses both channels: as γ decreases, the optimal subsidy on skilled labor $\tau_y(s)$ and the optimal tax on equipment capital τ_e increase. Therefore, as γ decreases, the importance of differential capital taxation increases. As a consequence, the welfare gains of optimal DTC relative to optimal NDTC increase.

We have also solved the model for the other values of $(\sigma, \gamma) \in \{1, 2, 4\} \times \{0.5, 1, 2\}$. We find that the welfare gains of optimal DTC relative to optimal NDTC are as high as 0.23%, and the welfare gains of optimal DTC relative to current DTC are as high as 0.28% for $\sigma = 4$ and $\gamma = 0.5$. We also consider a higher degree of complementarity between equipment capital and skilled labor, namely, $\eta = 0.79$.³³ For this degree of complementarity and $\sigma = 4$ and $\gamma = 0.5$, the welfare gains of optimal DTC relative to optimal NDTC are 0.35%, and the welfare gains of optimal DTC relative to current DTC are 0.40%.

7 Conclusion

The effective marginal tax rates on returns to capital assets differ substantially depending on the capital asset type. In particular, the marginal tax rate on capital structures is about 5% higher than the marginal tax rate on capital equipments. In this paper, we assess the optimality of differential capital asset taxation both theoretically and quantitatively from the perspective of a government whose aim is to provide insurance and redistribution.

Contrary to the actual practice in the U.S. tax code, we show that it is optimal to tax equipment capital at a higher rate than structure capital. Intuitively, in an environment with equipment-skill complementarity, taxing equipment capital and hence depressing its accumulation decreases the skill premium, providing indirect redistribution from the skilled to the unskilled agents. In a similar vein, it is optimal to subsidize skilled labor at the margin. The marginal subsidy increases skilled labor supply, and since labor types are imperfect substitutes, it decreases the skill premium.

In a quantitative version of our model, we find that the tax rate on equipment capital should be at least 27 percentage points higher than the tax rate on structure capital during transition and at the steady state. Furthermore, we find that the welfare gains of optimal

³³This value has been used, for example, in He and Liu (2008), who use the same production function, and comes from an empirical study by Duffy, Papageorgiou, and Perez-Sebastian (2004).

differential capital taxation can be as high as 0.4% of lifetime consumption.

We compute the welfare gains in an environment in which (i) the government is unrestricted in its choice of labor income taxes, and (ii) there are no labor income shocks. It would be interesting to assess the welfare consequences of optimal differential capital asset taxation in a model with labor income shocks in which labor income taxes are the same as in the U.S. tax code. Such an analysis would allow us to evaluate the welfare gains of partially reforming the U.S. tax code with respect to capital taxes. We conjecture that these welfare gains could be even larger than those that we find in this paper. We leave this question as a task for future research.

References

- ATKINSON, A. B., AND J. E. STIGLITZ (1976): “The Design of Tax Structure: Direct versus Indirect Taxation,” *Journal of Public Economics*, 6(1–2), 55–75.
- AUERBACH, A. J. (1979): “The Optimal Taxation of Heterogeneous Capital,” *Quarterly Journal of Economics*, 93(4), 589–612.
- AUERBACH, A. J. (1983): “Corporate Taxation in the United States,” *Brookings Papers on Economic Activity*, 14(2), 451–514.
- AUERBACH, A. J., K. A. HASSETT, AND S. D. OLINER (1994): “Reassessing the Social Returns to Equipment Investment,” *Quarterly Journal of Economics*, 109(3), 789–802.
- CHAMLEY, C. (1986): “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, 54(3), 607–622.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): “Taxing Capital? Not a Bad Idea after All,” *American Economic Review*, 99(1), 25–48.
- DIAMOND, P. A., AND J. A. MIRRLEES (1971): “Optimal Taxation and Public Production I: Production Efficiency,” *American Economic Review*, 61(1), 8–27.
- DOMEIJ, D., AND J. HEATHCOTE (2004): “On the Distributional Effects of Reducing Capital Taxes,” *International Economic Review*, 45(2), 523–554.
- DUFFY, J., C. PAPAGEORGIOU, AND F. PEREZ-SEBASTIAN (2004): “Capital-Skill Complementarity? Evidence from a Panel of Countries,” *Review of Economics and Statistics*, 86(1), 327–344.
- FARHI, E., AND I. WERNING (2012): “Capital Taxation: Quantitative Explorations of the Inverse Euler Equation,” *Journal of Political Economy*, 120(3), 398–445.
- FARHI, E., AND I. WERNING (2013): “Insurance and Taxation over the Life Cycle,” *Review of Economic Studies*, 80(2), 596–635.
- FELDSTEIN, M. (1990): “The Second Best Theory of Differential Capital Taxation,” *Oxford Economic Papers*, 42(1), 256–267.
- FERNANDES, A., AND C. PHELAN (2000): “A Recursive Formulation for Repeated Agency with History Dependence,” *Journal of Economic Theory*, 91(2), 223–247.
- FLUG, K., AND Z. HERCOWITZ (2000): “Equipment Investment and the Relative Demand for Skilled Labor: International Evidence,” *Review of Economic Dynamics*, 3, 461–485.
- FUKUSHIMA, K. (2010): “Quantifying the Welfare Gains From Flexible Dynamic Income Tax Systems,” Working paper.
- GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): “Optimal Indirect and Capital Taxation,” *Review of Economic Studies*, 70(3), 569–587.

- GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2011): “Optimal Dynamic Taxes,” NBER Working Papers 17642.
- GRAVELLE, J. G. (1994): *The Economic Effects of Taxing Capital Income*. Cambridge MA: MIT Press.
- (2003): “Capital Income Tax Revisions and Effective Tax Rates,” CRS Report for Congress 17642, Congressional Budget Office.
- (2011): “Reducing Depreciation Allowances to Finance a Lower Corporate Tax Rate,” *National Tax Journal*, 64(4), 1039–1054.
- GREEN, E. (1987): “Lending and the Smoothing of Uninsurable Income,” in *Contractual Arrangements for Intertemporal Trade*, ed. by E. C. Prescott, and N. Wallace. Minneapolis: University of Minnesota Press.
- GREENWOOD, J., Z. HERCOWITZ, AND P. KRUSELL (1997): “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review*, 87(3), 342–362.
- GROCHULSKI, B., AND T. PISKORSKI (2010): “Risky Human Capital and Deferred Capital Income Taxation,” *Journal of Economic Theory*, 145(3), 908–943.
- HE, H., AND Z. LIU (2008): “Investment-Specific Technological Change, Skill Accumulation, and Wage Inequality,” *Review of Economic Dynamics*, 11(2), 314–334.
- HEATHCOTE, J., F. PERRI, AND G. L. VIOLANTE (2010): “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967–2006,” *Review of Economic Dynamics*, 13(1), 15–51.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2010): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118(4), 681–722.
- HUGGETT, M., AND J. C. PARRA (2010): “How Well Does the U.S. Social Insurance System Provide Social Insurance?,” *Journal of Political Economy*, 118(1), 76–112.
- HUGGETT, M., G. VENTURA, AND A. YARON (2011): “Sources of Lifetime Inequality,” *American Economic Review*, 101(7), 2923–2954.
- JUDD, K. L. (1985): “Redistributive Taxation in a Simple Perfect Foresight Model,” *Journal of Public Economics*, 28(1), 59–83.
- KEANE, M. P., AND K. I. WOLPIN (1997): “The Career Decisions of Young Men,” *Journal of Political Economy*, 105(3), 473–522.
- KOCHERLAKOTA, N. R. (2005): “Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation,” *Econometrica*, 73(5), 1587–1621.
- (2010): *The New Dynamic Public Finance*. Princeton, NJ: Princeton University Press.

- KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis,” *Econometrica*, 68(5), 1029–1054.
- MIRPLEES, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38(114), 175–208.
- NAITO, H. (1999): “Re-examination of Uniform Commodity Taxes under a Non-linear Income Tax System and its Implication for Production Efficiency,” *Journal of Public Economics*, 71(2), 165–188.
- ROGERSON, W. P. (1985): “Repeated Moral Hazard,” *Econometrica*, 53(1), 69–76.
- ROTHSCHILD, C., AND F. SCHEUER (2012): “Redistributive Taxation in a Roy Model,” Working paper.
- SADKA, E. (1976): “On Income Distribution, Incentive Effects and Optimal Income Taxation,” *Review of Economic Studies*, 43(2), 261–267.
- SAMUELSON, P. A. (1964): “Tax Deductibility of Economic Depreciation to Insure Invariant Valuations,” *Journal of Political Economy*, 72(6), 604–606.
- SEADE, J. K. (1977): “On the Shape of Optimal Tax Schedules,” *Journal of Public Economics*, 7(2), 203–235.
- STIGLITZ, J. E. (1982): “Self-Selection and Pareto Efficient Taxation,” *Journal of Public Economics*, 17(2), 213–240.
- STORESLETTEN, K., C. I. TELMER, AND A. YARON (2004): “Consumption and Risk Sharing over the Life Cycle,” *Journal of Monetary Economics*, 51(3), 609–633.
- WERNING, I. (2007): “Optimal Fiscal Policy with Redistribution,” *Quarterly Journal of Economics*, 122(3), 925–967.

Appendix

A Proofs

A.1 Proof of Proposition 4

For any t , let $\bar{h}^t = (h^{t-1}, s)$ and for $m \leq t$, let \bar{h}^m be the predecessor in period m . Consider the first-order optimality conditions with respect to $c_t(\bar{h}^t)$ and $l_t(\bar{h}^t)$:

$$\begin{aligned} u'(c_t^*(\bar{h}^t))A_t^*(\bar{h}^t) &= \lambda_t^* \pi_t(\bar{h}^t) \\ v'(l_t^*(\bar{h}^t)) \left(A_t^*(\bar{h}^t) - \frac{B_t^* \pi_t(\bar{h}^t)}{v'(l_t^*(\bar{h}^t))} \right) &= \lambda_t^* w_t^*(\bar{h}_t) \pi_t(\bar{h}^t), \end{aligned}$$

where, letting χ be the indicator function,

$$A_t^*(\bar{h}^t) = \left(\beta^{t-1} \pi_t(\bar{h}^t) + \sum_{m=1}^t \beta^{t-m} \pi_t(\bar{h}^t | \bar{h}^m) \mu_m^*(\bar{h}^m) [\chi_{\{s\}}(\bar{h}_m) - \chi_{\{u\}}(\bar{h}_m)] \right)$$

and

$$B_t^* = \sum_{\{h^t \in H^t | h_t = s\}} \mu_t^*(h^t) v' \left(\frac{l_t^*(h^{t-1}, u) w_t^*(u)}{w_t^*(s)} \right) l_t^*(h^{t-1}, u) \frac{\partial \frac{w_t^*(u)}{w_t^*(s)}}{\partial L_{s,t}^*} z_s.$$

By Assumption 3, we have $B_t^* > 0$, which implies

$$\tau_{y,t}(\bar{h}^t) = 1 - \frac{A_t^*(\bar{h}^t)}{\left(A_t^*(\bar{h}^t) - \frac{B_t^* \pi_t(\bar{h}^t)}{v'(l_t^*(\bar{h}^t))} \right)} < 0.$$

□

A.2 Proof of Proposition 5

First, define prices as $r_t^* = F_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) - \delta_s$ and $w_{j,t}^* = \frac{\partial F(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)}{\partial L_{j,t}^*}$. Second, observe that the only budget feasible income strategy for the household is the one that corresponds to some agent's income process at the constrained efficient allocation (i.e., for any history h^t in any period t , y^t should be in Y^{t*}).

Third, we claim that if an agent chooses an income strategy y' , where this means $y'_t(h^t) = y_t^*(\hat{h}^t)$ for some \hat{h}^t , then facing the prices defined above, the agent also chooses (c', a') , meaning that $c'_t(h^t) = c_t^*(\hat{h}^t)$ and $a'_{t+1}(h^t) = a_{t+1}^*(\hat{h}^t)$ for all h^t, t . If we can show this claim,

then the result that agent will actually choose the constrained efficient allocation follows, since the constrained efficient allocation is incentive compatible.

To see this claim, take an agent that follows income strategy y' . His problem is

$$\max_{c,a} \sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} \left[u(c_t(h^t)) - v \left(\frac{y_t^*(\hat{h}^t)}{w_t^*(h^t)} \right) \right]$$

subject to

$$c_t(h^t) + a_{t+1}(h^t) \leq y_t^*(\hat{h}^t) - \tau_{y,t}^*(y^{t*}(\hat{h}^t)) + \left[1 - \tau_{a,t}^*(y^{t*}(\hat{h}^t)) \right] (1 + r_t^*) a_t(h^{t-1}),$$

$$a_1 \leq K_{s,1}^* + K_{e,1}^*,$$

c is nonnegative.

The first-order conditions to this problem are the budget constraint holding with equality and

$$u'(c_t(h^t)) = (1 + r_{t+1}^*) \beta \sum_{h^{t+1}|h^t} \pi_{t+1}(h^{t+1}|h^t) u'(c_{t+1}(h^{t+1})) [1 - \tau_{t+1}^{a*}(y^{t+1*}(\hat{h}^t, h_{t+1}))].$$

Clearly, the agent's problem is concave, and hence these first-order conditions are necessary and sufficient for optimality provided that a relevant transversality condition holds.

By construction of the labor tax code and the prices, $c_t^*(\hat{h}^t)$ and $a_{t+1}^*(\hat{h}^t)$ satisfy the flow budget constraints. To see that they also satisfy the Euler equation above, observe that wealth taxes are constructed such that

$$1 - \tau_{a,t+1}^*(y^{t+1*}(\hat{h}^t, h_{t+1})) = \frac{u'(c_t^*(\hat{h}^t))}{\beta u'(c_{t+1}^*(\hat{h}^t, h_{t+1})) \tilde{F}_1^*(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})}.$$

We also need to make sure that the firm chooses the right allocation. The firm's optimality conditions for labor are satisfied at the constrained efficient allocation by construction of wages. The optimality conditions for capital are

$$\begin{aligned} (K_{s,t}) &: \tilde{F}_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r_t^*) - \tau_{f,t}^* [F_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \bar{\delta}_{s,t}] = 0 \\ (K_{e,t}) &: \tilde{F}_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r_t^*) - \tau_{f,t}^* [F_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \bar{\delta}_{e,t}] = 0. \end{aligned}$$

The first condition above holds again by construction of r_t^* and $\bar{\delta}_{s,t}^*$ whereas the second one holds by construction of $\tau_{f,t}^*$, r_t^* and $\bar{\delta}_{e,t}^*$. \square

A.3 Proof of Proposition 6

Let $\lambda_t \beta^{t-1}$ and μ be the multipliers on the feasibility and incentive constraints in the social planner's problem. When types are permanent, the counterparts of conditions (6) and (7) in Proposition 2 become

$$u'(c_{j,t}^*) = \beta \tilde{F}_{1,t+1}^* u'(c_{j,t+1}^*) \quad (11)$$

and

$$u'(c_{j,t}^*) = \beta \tilde{F}_{2,t+1}^* u'(c_{j,t+1}^*) \left(1 + \frac{X_{t+1}^*}{\tilde{F}_{2,t+1}^*} \right), \quad (12)$$

where

$$X_{t+1}^* = \frac{1}{\lambda_{t+1}^*} \mu^* v' \left(\frac{l_{u,t+1}^* w_{u,t+1}^*}{w_{s,t+1}^*} \right) l_{u,t+1}^* \frac{\partial \left(\frac{w_{u,t+1}^*}{w_{s,t+1}^*} \right)}{\partial K_{e,t+1}^*}.$$

Part 1. Equation (11) together with the definition of intertemporal wedges proves that the structure capital wedge is zero in all periods. By the assumption of equipment capital-skill complementarity, we have

$$\frac{\partial \left(\frac{w_{u,t+1}^*}{w_{s,t+1}^*} \right)}{\partial K_{e,t+1}^*} < 0.$$

Since $\mu^* > 0$, we have $X_{t+1}^* < 0$. Using the definition of the equipment capital wedge and equation (12), we have for both types of agents

$$\tau_{e,t}^* = -\frac{X_{t+1}^*}{\tilde{F}_{2,t+1}^*} > 0. \quad (13)$$

This finishes the proof of part 2 of Proposition 6.

Part 2. Suppose the optimal allocation converges to a steady state. Letting the allocation without any time subscripts denote this steady-state allocation, we have

$$X^* = \frac{1}{\lambda^*} \mu^* v' \left(\frac{l_u^* w_u^*}{w_s^*} \right) l_u^* \frac{\partial \left(\frac{w_u^*}{w_s^*} \right)}{\partial K_e^*} < 0.$$

Thus, we have

$$\tau_e^* = -\frac{X^*}{\tilde{F}_2^*} > 0.$$

□

B Alternative Tax Systems

B.1 Optimal Nondifferential Taxation of Capital

In optimal NDTC, the planner is not allowed to tax the two types of capital differentially. This means that for any history h^t we impose (using the definitions of the taxes on capital income net of depreciation),

$$\tau_{s,t}(h^t) := 1 - \frac{u'(c_t(h^t))}{\beta \tilde{F}_{1,t+1} E_t \{u'(c_{t+1}(h^{t+1})) | h^t\}} = 1 - \frac{u'(c_t(h^t))}{\beta \tilde{F}_{2,t+1} E_t \{u'(c_{t+1}(h^{t+1})) | h^t\}} =: \tau_{e,t}(h^t).$$

This condition can be simplified to an equality constraint on net returns: $F_{1,t+1} - \delta_s = F_{2,t+1} - \delta_e$, which we impose for each period. The optimal NDTC planning problem reads:

$$\begin{aligned} & \max_{\{(c_{j,t}, l_{j,t})_j, K_{s,t}, K_{e,t}, L_{u,t}, L_{s,t}\}_{t=0}^{\infty}} \sum_{j=u,s} \pi_j \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{j,t}) - v(l_{j,t})] \quad \text{s.t.} \\ \forall t : & G_t + \sum_j \pi_j c_{j,t} + K_{s,t+1} + K_{e,t+1} \leq \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \\ & \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{s,t}) - v(l_{s,t})] \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[u(c_{u,t}) - v\left(\frac{l_{u,t} w_{u,t}}{w_{s,t}}\right) \right], \\ \forall t : & F_1(K_{s,t+1}, K_{e,t+1}, L_{s,t+1}, L_{u,t+1}) - \delta_s = F_2(K_{s,t+1}, K_{e,t+1}, L_{s,t+1}, L_{u,t+1}) - \delta_e. \end{aligned}$$

As in the optimal DTC, we compute the approximate solution to this problem by assuming that the economy converges to steady state in 200 periods.

B.2 Current Differential Taxation of Capital

In the current DTC, the planner must use the capital income taxes as in the U.S. tax code. This means that we impose the following set of constraints for both types of agents in each period (here, τ_s and τ_e are fixed over time and taken from the data, namely, $\tau_s = 0.422$ and $\tau_e = 0.371$):

$$\begin{aligned} 1 - \tau_s &= \frac{\frac{u'(c_{t,s})}{\beta u'(c_{t+1,s})} - 1}{F_{1,t+1} - \delta_s}, \\ 1 - \tau_e &= \frac{\frac{u'(c_{t,s})}{\beta u'(c_{t+1,s})} - 1}{F_{2,t+1} - \delta_e}, \end{aligned}$$

$$1 - \tau_s = \frac{\frac{u'(c_{t,u})}{\beta u'(c_{t+1,u})} - 1}{F_{1,t+1} - \delta_s},$$

$$1 - \tau_e = \frac{\frac{u'(c_{t,u})}{\beta u'(c_{t+1,u})} - 1}{F_{2,t+1} - \delta_e}.$$

Note that any three of these constraints imply the fourth. Therefore, we write the current DTC planning problem, ignoring the fourth:

$$\begin{aligned} & \max_{\{(c_{j,t}, l_{j,t})_j, K_{s,t}, K_{e,t}, L_{u,t}, L_{s,t}\}_{t=0}^{\infty}} \sum_{j=u,s} \pi_j \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{j,t}) - v(l_{j,t})] \quad \text{s.t.} \\ \forall t : G_t + \sum_j \pi_j c_{j,t} + K_{s,t+1} + K_{e,t+1} & \leq \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \\ \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{s,t}) - v(l_{s,t})] & \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[u(c_{u,t}) - v\left(\frac{l_{u,t} w_{u,t}}{w_{s,t}}\right) \right], \\ \forall t : \beta u'(c_{t+1,s}) [(1 - \tau_s)(F_{1,t+1} - \delta_s) + 1] & = u'(c_{t,s}) \\ \forall t : \beta u'(c_{t+1,s}) [(1 - \tau_e)(F_{2,t+1} - \delta_e) + 1] & = u'(c_{t,s}) \\ \forall t : \beta u'(c_{t+1,u}) [(1 - \tau_s)(F_{1,t+1} - \delta_s) + 1] & = u'(c_{t,u}) \end{aligned}$$

As in the optimal DTC, we compute the approximate solution to this problem by assuming that the economy converges to steady state in 200 periods.