Machines, Buildings, and Optimal Dynamic Taxes

Ctirad Slavík\textsuperscript{a} and Hakki Yazici\textsuperscript{b}

\textsuperscript{a}Goethe University Frankfurt, Frankfurt, Germany. Email: slavik@econ.uni-frankfurt.de

\textsuperscript{b}Sabanci University, Istanbul, Turkey. E-mail: hakkiyazici@sabanciuniv.edu

February 22, 2014

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. Under a plausible assumption, 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 the new steady state. The welfare gains of optimal differential capital taxation can be as high as 0.4\% of lifetime consumption.


\textit{Keywords:} Differential capital asset taxation, equipment capital, structure capital, equipment-skill complementarity.

\textsuperscript{*}Corresponding author: Ctirad Slavík, Grüneburgplatz 1 (Campus Westend) House of Finance, Room 3.49 60323 Frankfurt am Main, Germany. Telephone: +49 (0)69 798 33808.
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 non-residential 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 non-residential 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 are as high as 0.4% of lifetime consumption for reasonable parameter values.

We study dynamic optimal taxes in an economy in which people are heterogeneous in

\(^1\) The capital tax differentials are created through tax depreciation allowances that differ from actual depreciation rates. We explain this in detail and provide further information on the historical evolution of capital tax differentials in the U.S. tax code in Appendix A.

terms of their skills, and the government uses capital and labor income taxes to provide redistribution (insurance). In our benchmark model, we consider an environment with permanent skills. We then generalize our main theoretical results to an environment with stochastic skills. Our approach to optimal dynamic taxation follows the recent New Dynamic Public Finance literature in the sense that taxes are allowed to be arbitrary functions of people’s past and current incomes.

The key feature of our environment is equipment-skill complementarity in the production technology. 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 for the U.S. economy provided by 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. Structure capital is neutral in terms of its complementarity with skilled and unskilled labor. More generally, Flug and Hercowitz (2000) provide evidence for equipment-skill complementarity for a large panel of countries.

Equipment-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 and/or 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 capital taxation.

We assess the quantitative importance of differential capital taxation using the model with permanent skills calibrated to the U.S. economy. 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.

The skill premium is about 40% in the first period after the optimal tax reform, and rises over the transition to 48% in the new steady state. Thus, the ‘optimal’ skill premium in any period is significantly lower than 80%, the empirical estimate for the current U.S. economy. This suggests that the optimal tax system relies much more on indirect redistribution than the current U.S. tax system. We also find that the optimal skill premium is rising over the transition because the economy is growing, and hence, the level of equipment capital increases. This result is interesting as it suggests that, even if the government cares about equality, an increasing skill premium is optimal in a growing economy.

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 the tax rates on both types of capital are restricted to be equal to the values in the U.S. tax code. The additional welfare gains of allowing for differential capital taxation are 0.19% in terms of lifetime consumption in the benchmark and can be as high as 0.40% within the set of reasonable parameter values.

We focus on the redistribution and insurance provision role of differential capital taxation. 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 its redistributive and insurance provision role.

**Related Literature.** Our paper relates to three distinct strands of literature. First, in their seminal paper Diamond and Mirrlees (1971) show that tax systems should maintain productive efficiency. In an environment with multiple capital types, this result implies that all capital should be taxed at the same rate. However, Auerbach (1979) and Feldstein (1990)
show that it might be optimal to tax capital differentially if the government is exogenously restricted to a narrower set of fiscal instruments than in Diamond and Mirrlees (1971). Our paper is different in the sense that the optimality of differential capital taxation stems from redistribution and/or insurance motives.

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 are allowed to change stochastically over time and 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 the current paper. In addition, our paper contributes 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, see e.g., Fukushima (2010), Huggett and Parra (2010), Farhi and Werning (2013), and Golosov, Troshkin, and Tsyvinski (2013).

Our paper is also 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. Naito (1999) shows that the uniform commodity taxation result of Atkinson and Stiglitz (1976) and productive efficiency result of Diamond and Mirrlees (1971) are no longer valid under imperfect labor substitutability. Ales, Kurnaz, and Sleet (2014) analyze a static optimal tax problem in which agents with different skills are assigned to tasks (occupations). They calculate optimal taxes for the U.S. economy for the 1970s and the 2000s and compare them to their empirical counterparts. In addition, they analyze the impact of technical change on optimal taxes. 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. In Section 2, we lay out the model for the
case of permanent skills. In Section 3, we show that differential capital taxation is optimal in this environment. Section 4 generalizes our main results to an environment with stochastic skills. Section 5 discusses our quantitative results, and Section 6 concludes. Appendix A discusses differential taxation of capital assets in the U.S. tax code, Appendix B contains the proofs, Appendix C provides a formal implementation of the constrained efficient allocation in an incomplete markets environment, and Appendix D defines alternative social planning problems that we analyze in Section 5.

2 Model

There is a continuum of measure one of agents who live for infinitely many periods. They differ in their skill levels: they are born either skilled or unskilled, \( h \in H = \{u, s\} \). A fraction \( \pi_h \) of agents belong to skill group \( h \). In the main body of the paper, we assume that agents’ skills are permanent. Permanent skills is a natural assumption given that in our quantitative analysis we associate skill levels with educational attainment. In Section 4, we show that our main theoretical results remain valid for a general stochastic skill process.

Production Technology. An agent of skill level \( h \) produces \( l \cdot z_h \) units of effective \( h \) type labor when he works \( l \) units of labor. There are two different occupational sectors in this economy: a skilled occupation in which only skilled agents are allowed to work and an unskilled occupation 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 occupation. The second assumption can be rationalized as follows. We assume that agents get the same disutility from working in the two occupations. Therefore, a skilled agent will choose to work in the skilled occupation as long as he gets a higher wage in the skilled occupation. 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.

Output is produced according to a production function \( Y = F(K_s, K_e, L_s, L_u) \), where
$L_s, L_u, K_s$ and $K_e$ denote the aggregate amounts of effective skilled labor, effective unskilled labor, structure capital, and equipment capital. Output can be used for consumption or can be converted to structure or equipment capital one-for-one. The economy is initially endowed with $K_{s,1}^*$ and $K_{e,1}^*$ units of the capital goods. We define a function $\tilde{F}$ that gives the total wealth of the economy: $\tilde{F} = F + (1 - \delta_s)K_s + (1 - \delta_e)K_e$, where $\delta_s$ and $\delta_e$ denote the depreciation rates of structure and equipment capital. We define $F_i(\cdot)$ and $\tilde{F}_i(\cdot)$ as partial derivatives of $F$ and $\tilde{F}$ with respect to the $i^{th}$ argument.

**Wages.** Agents of type $h \in H$ receive wage $w_{h,t}$ in period $t$ for one unit of their labor:

\[
  w_{s,t} = F_3(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) \cdot z_s, \quad w_{u,t} = F_4(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) \cdot z_u.
\] (1)

**Equipment-Skill Complementarity.** Following Krusell, Ohanian, Ríos-Rull, and Violante (2000), we assume that the production technology features equipment-skill complementarity, which means that the degree of complementarity between equipment capital and skilled labor is higher than that between equipment capital and unskilled labor. This 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 (skill premium) 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.** $F_3(\cdot)/F_4(\cdot)$ is independent of $K_s$.

**Assumption 2.** $F_3(\cdot)/F_4(\cdot)$ is strictly increasing in $K_e$.

**Assumption 3.** $F_3(\cdot)/F_4(\cdot)$ is strictly decreasing in $L_s$ and strictly increasing in $L_u$. 

6
Assumptions (1) - (3) are maintained throughout the paper without further reference.

Preferences. An agent of type \( h \) evaluates a consumption-labor sequence, \( (c_{h,t}, l_{h,t})_{t=1}^{\infty} \), with a utility function that is time-separable and separable between consumption and labor,

\[
\sum_{t=1}^{\infty} \beta^{t-1} [u(c_{h,t}) - v(l_{h,t})],
\]

where \( \beta \in (0, 1) \) is the discount factor, \( u, v : \mathbb{R}_+ \rightarrow \mathbb{R} \). We assume that \( u', -u'', v', v'' > 0 \).

Allocation. An allocation is \( x = ((c_{h,t}, l_{h,t})_{h \in H}, K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^{\infty} \).

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

\[
\sum_{h=u,s} \pi_h c_{h,t} + K_{s,t+1} + K_{e,t+1} + G_t \leq \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}), \quad (2) \\
L_{h,t} = \pi_h l_{h,t} z_h, \quad \text{for } h \in H \quad \text{and} \quad K_{s,1} \leq K_{s,1}^*, K_{e,1} \leq K_{e,1}^*.
\]

Here, \( \{G_t\}_{t=0}^{\infty} \) is a sequence of exogenously given wasteful government consumption.

Optimal Tax Problem. As in the U.S. tax code, we assume that taxes only depend on people’s incomes, and not directly on their skills, occupations, wages, or labor supplies. We do not model why the government does not use this information in the tax code (there could be constitutional, administrative or other reasons), but rather focus on the best policy given the existing fiscal framework. Following Mirrlees (1971) and the recent New Dynamic Public Finance literature, we impose no further restrictions on the tax code; specifically, taxes can be arbitrarily nonlinear functions of income histories.

Following Kocherlakota (2010), we make no explicit mention of private information in motivating why we restrict attention to tax systems that depend only on incomes. However, the fact that the government can condition taxes only on income implies that the optimal tax problem is isomorphic to a social planning problem, in which agents are privately informed about their skills, occupations, wages, and labor supplies. Income and consumption are public information. In the planning problem, each agent reports his skill type to the planner.
and receives an allocation as a function of his report.\(^4\) The set of allocations available to the planner is constrained by incentive compatibility constraints, which ensure that agents do not misreport their types.\(^5\)

Our strategy is to first characterize the solution to the planning problem and then use this characterization to back out properties of an optimal tax system. We now formally define the notion of incentive compatibility.

**Incentive Compatibility.** With permanent types, people report their type only once in the first period. Moreover, since we assume that agents cannot switch occupations, an agent can only mimic the other type’s income level by adjusting his labor hours. As a result, the planner faces only two incentive constraints.

We say that an allocation is incentive compatible if and only if for all \( h \in H \)

\[
\sum_{t=1}^{\infty} \beta^{t-1} \left[ u(c_{h,t}) - v(l_{h,t}) \right] \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[ u(c_{j,t}) - v(l_{j,t} w_{j,t} w_{h,t}) \right],
\]

where \( j \) denotes the complement of \( h \) in the set \( H \).

**Social Planning Problem.** We analyze the problem of a planner who maximizes a Utilitarian objective with equal weights on all agents. The social planning problem is

\[
\max_x \sum_{h \in H} \pi_h \sum_{t=1}^{\infty} \beta^{t-1} \left[ u(c_{h,t}) - v(l_{h,t}) \right] \quad \text{s.t.} \quad (1), (2), (3), \text{ and (4)}. 
\]

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

---

\(^4\)Agents only report their skill types, because given that income is observable and skilled (unskilled) agents can only work in the skilled (unskilled) occupation, knowing an agent’s skill type reveals all his private information.

\(^5\)The restriction to direct truth-telling mechanisms is without loss of generality because of the following argument. 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 Appendix C, we prove that there is a tax system that implements 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.
3 Optimality of Differential Taxation of Capital

In this section, we uncover the economic mechanism that calls for differential capital taxation. We show that, with equipment-skill complementarity, as long as only the incentive constraint that prevents skilled agents from pretending to be unskilled binds, the optimal tax on equipment capital is strictly higher than the optimal tax on structure capital. We begin by formalizing the assumption on the pattern of binding incentive constraints.

Assumption 4. The incentive constraint (4) binds for \( h = s \) and is slack for \( h = u \) at the solution to the social planning problem.

In all numerical simulations in Section 5, in which we parameterize the model to match the U.S. data, the skilled wage is higher than the unskilled wage in every period. However, in our environment with endogeneous wages, it is not possible to guarantee that skilled wages are always higher than unskilled wages without making very restrictive assumptions on \( F \). Without monotonic wages, it is not possible to determine the pattern of binding incentive constraints. Therefore, we proceed directly with Assumption 4, see Stiglitz (1982) for the same approach. We verify that Assumption 4 is satisfied in all our quantitative exercises.

3.1 Capital Return Wedge

In the standard growth model with two types of capital, aggregate savings are allocated between the two types of capital in a way that equates their marginal returns. Proposition 1 below shows that this is not true in the constrained efficient allocation, meaning it is optimal to create a wedge between the marginal returns to structure and equipment capital. This result forms the basis for the optimality of differential taxation of capital: to create the optimal wedge in the market equilibrium, the two types of capital should be taxed differently.

Proposition 1. Suppose Assumption 4 holds. Then, at the constrained efficient allocation, in any period \( t \geq 2 \),

\[
\tilde{F}_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) < \tilde{F}_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}).
\]
Proof. Let $\lambda t\beta^{t-1}$ be the multiplier on period $t$ feasibility constraint and $\mu$ be the multiplier on skilled agents’ incentive constraint. The first order optimality conditions with respect to the two types of capital are:

\[
\begin{align*}
(K_{e,t}) : \lambda_{t-1}^* &= \beta \left[ \lambda_t^* \tilde{F}_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) + X_t^* \right], \\
(K_{s,t}) : \lambda_{t-1}^* &= \beta \lambda_t^* \tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*),
\end{align*}
\]

where

\[
X_t^* = \mu^* v^t \left( \frac{w_{u,t}^*}{w_{s,t}^*} \right) K_{e,t}^* \frac{\partial}{\partial K_{e,t}^*} \frac{w_{u,t}^*}{w_{s,t}^*}.
\]

By equipment-skill complementarity, we have $\partial \left( \frac{w_{u,t}^*}{w_{s,t}^*} \right) / \partial K_{e,t}^* < 0$. Since $\mu^* > 0$, we have $X_t^* < 0$. Using $X_t^* < 0$ together with the first-order conditions gives the result. □

Because of equipment-skill complementarity, increasing the level of equipment capital in period $t$ decreases the wage ratio $w_{u,t}^*/w_{s,t}^*$. This makes it more profitable for the skilled agents to pretend to be unskilled and, hence, tightens the skilled incentive constraint. From a planning perspective, this means that increasing equipment capital has an extra negative return, $X_t^* < 0$, in addition to the physical return, $\tilde{F}_2^* t$, where $F_i^* t$ denotes $\tilde{F}_i(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)$. Since structure capital is neutral, changing its level does not affect the incentive constraint, and hence its only return is its physical return, $\tilde{F}_1^* t$. In order for the overall return on the two types of capital to be equal, the physical return on equipment capital must higher than the physical return on structure capital at the constrained efficient allocation.

This result is intuitive: decreasing the level of equipment capital has an additional marginal benefit for the planner, because it decreases the skill premium and thus indirectly redistributes from the skilled to the unskilled. Decreasing the level of equipment capital increases its return above the return on structure capital due to diminishing marginal returns. This intuition shows that there is an extra reason to depress equipment capital accumulation relative to structure capital. This implies that equipment capital should be taxed at a higher rate than structure capital as we show in Section 3.2.

Two features of the model are key for the optimality of the capital return wedge. First,
if equipment capital was also neutral in terms of its complementarity with the two types of labor, then, \( X^*_t = 0 \), and hence, it would be efficient to equate the physical marginal returns to the two types of capital. Second, if the government could condition taxes on skill types, it could redistribute via type-specific lump-sum taxes at zero efficiency cost and would not need the indirect (and distortionary) channel of redistribution, which works through the capital return wedge. In terms of the planning problem, this would mean that skills were not private information but publicly known. As a result, there would be no incentive constraints, and hence, \( X^*_t = 0 \), and the optimal capital return wedge would again be zero.

### 3.2 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. In Proposition 2, we characterize the properties of optimal wedges (distortions) that a planner has to create in the intertemporal allocation of resources in order to implement the constrained efficient allocation in a competitive market environment, in which people are allowed to save through both types of capital. Formally, the optimal intertemporal wedge that the planner has to create for an agent of type \( h \) for capital of type \( i \in \{ s, e \} \) from period \( t \) to \( t+1 \) is defined as \( \tau^*_{i,t+1}(h) = 1 - \frac{u'(c^*_{h,t})}{\beta \tilde{F}^*_i,t+1 u'(c^*_{h,t+1})} \).

**Proposition 2.** Suppose Assumption 4 holds. Then,

1. In all periods \( t \geq 2 \), the optimal wedge on equipment capital is strictly positive and independent of agent type, whereas the optimal wedge on structure capital is zero, i.e.,
   
   for all \( h \in H \), \( \tau^*_{e,t}(h) > \tau^*_{s,t}(h) = 0 \).

2. If a steady state of the constrained efficient allocation exists, then the optimal wedge on equipment capital is strictly positive at the steady state.

**Proof.** Relegated to Appendix B. \( \square \)

Part 1 of Proposition 2 calls for zero taxation of structure capital and positive taxation of equipment capital in every period. Recall 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. In addition, it follows from the uniform commodity taxation result of Atkinson and Stiglitz (1976) that in the absence of skill risk, it is optimal not to tax structure capital. 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. Finally, the proposition also shows that the capital tax rates are type independent.

Part 2 of Proposition 2 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 long run optimality of zero capital taxation in the Ramsey literature shown by Chamley (1986) and Judd (1985).

### 3.3 Intratemporal Wedges

Our model has interesting implications for intratemporal wedges (i.e., marginal labor income taxes) as well. The optimal intratemporal wedge in period $t$ for an agent of skill type $h$, defined as $\tau_{y,t}(h) = 1 - v'(l_{h,t})/\left[w_{h,t}^* u'(c_{h,t})\right]$, measures the efficient distortion that the planner needs to create in this agent’s intratemporal allocation of consumption and labor in period $t$. 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. Huggett and Parra (2010) proves this result for a dynamic Mirrleesian economy in which skill types are permanent and a version of our Assumption 4 holds. In Proposition 3, we show that the no distortion at the top result does not hold in the

---

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

7 If Assumption 4 is not satisfied, it will still be generically optimal to tax the two types of capital differentially, as we show explicitly in a more general environment in Section 4. However, in that case, it is not possible to determine which capital good will be taxed at a higher rate.
presence of equipment-skill complementarity. In particular, we show that the skilled agents’ labor income should be subsidized.

**Proposition 3.** Suppose Assumption 4 holds. Then, in any period \( t \geq 1 \), the optimal intratemporal wedge of the skilled agent is negative, i.e., \( \tau_{y,t}^*(s) < 0 \).

**Proof.** Relegated to Appendix A. \qed

The intuition for this result is as follows. Under the equipment-skill complementarity assumption, skilled and unskilled labor are imperfect substitutes. This implies that increasing the labor supply of the skilled agents decreases the skill premium which means that increasing skilled labor supply creates indirect redistribution. In order to encourage the supply of skilled labor, the government finds it optimal to subsidize skilled labor at the margin. This result is in line with Stiglitz (1982), who shows that when two types of labor are imperfect substitutes, the more productive agents’ labor supply should be subsidized.

### 4 Generalization to Stochastic Skills

In the model laid out in Section 2, agents’ skill types are permanent. In this section, we allow for agents’ skills to evolve stochastically over time. This level of generality is useful because it allows us to establish that our main theoretical results remain valid if people’s skills change after they enter the labor market, or if we take a dynastic interpretation of our model in which skills change from one generation to another. Notice that in this environment with stochastic skills the government uses taxes to provide insurance in addition to providing redistribution and financing public spending.

We first show that differential taxation of capital is optimal for any stochastic skill process. Moreover, under an assumption regarding the pattern of binding incentive compatibility conditions, it is optimal to tax equipment capital at a higher rate than structure capital.

The environment is the same as in Section 2 except that people are born identical, but their skills evolve stochastically over time. A skill realization in period \( t \) is denoted by
A partial skill history in period \( t \) is denoted by \( h^t = (h_1, h_2, \ldots, h_t) \in H^t \), where \( H^t \) denotes the set of all period \( t \) histories. Let \( \pi_t(h^t) \) be the unconditional probability of \( h^t \).

**Wages.** An agent of type \( h \) in period \( t \) receives a wage \( w_{h,t} \), defined in equation (1), independent of his skill history before period \( t \). For expositional convenience, in this section, we use \( w_t(h_t) \) instead of \( w_{h,t} \) to denote wages.

**Preferences.** Preferences are now defined over stochastic processes of consumption and labor, \((c_t, l_t)_{t=0}^{\infty}\), where \( c_t, l_t : H^t \to \mathbb{R}_+ \), using an ex ante expected discounted utility function,

\[
\sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} \left[ u(c_t(h^t)) - v(l_t(h^t)) \right]. \tag{5}
\]

**Allocation.** An allocation is \( x = (c_t, l_t, K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^{\infty} \).

**Feasibility.** An allocation is feasible if in any period \( t \geq 1 \),

\[
L_{h,t} = \sum_{\{h^t \in H^t | h_t = h\}} \pi_t(h^t)l_t(h^t)z_h \quad \text{for } h \in H, \quad \text{and} \quad K_{s,1} \leq K_{s,1}^*, K_{e,1} \leq K_{e,1}^*. \tag{7}
\]

**Incentive Compatibility.** Define \( \sigma_t : H^t \to H \). A reporting strategy is \( \sigma = (\sigma_t)_{t=1}^{\infty} \). Let \( \Sigma \) denote the set of all reporting strategies. The truth-telling strategy, which we denote by \( \sigma^* \), prescribes reporting the true type at each and every node: for all \( h^t \), \( \sigma^*_t(h^t) = h_t \). Let \( \sigma^t(h^t) = (\sigma_1(h^1), \ldots, \sigma_t(h^t)) \) denote the history of reports along history \( h^t \). 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 \( \sigma \) given an allocation \( x \). We say that an allocation \( x \) is incentive compatible if and only if for all \( \sigma \in \Sigma \), \( W(\sigma^*|x) \geq W(\sigma|x) \).

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. This allows us to replace the incentive compatibility constraints defined above with a sequence of temporary
incentive constraints, one for each node. We say that an allocation \( x \) is temporary incentive compatible if and only if, in any period \( t \) and at any node \( h^{t-1} \) and for all \( h_t \in H \),

\[
\begin{align*}
  u(c_t(h^{t-1}, h_t)) - v(l_t(h^{t-1}, h_t)) &+ \sum_{m=t+1}^{\infty} \sum_{h^m \in H^m} \pi_m(h^m) \beta^{m-t} \left[ u(c_m(h^m)) - v(l_m(h^m)) \right] \\
  \geq u(c_t(h^{t-1}, h^*_t)) - v \left( \frac{l_t(h^{t-1}, h^*_t) w_t(h^*_t)}{w_t(h_t)} \right) + \sum_{m=t+1}^{\infty} \sum_{h^m \in H^m} \pi_m(h^m) \beta^{m-t} \left[ u(c_m(\tilde{h}^m)) - v(l_m(\tilde{h}^m)) \right],
\end{align*}
\]

where \( h^*_t \) is the complement of \( h_t \) in the set \( H \), \( H^m \) denotes the set of period \( m \) histories that follow from \( h^t \), i.e., \( H^m \equiv \{ h^m \in H^m : h^m \succ h^t \} \), and \( \tilde{h}^m = (h^{t-1}, h^*_t, h_{t+1}, ..., h_m) \) is identical to \( h^m \) except in period \( t \). From now on, we use (8) to represent incentive compatibility.\(^8\)

**Social Planning Problem.** The social planning problem that defines the constrained efficient allocation is: \( \max_x (5) \) s.t. (1), (6), (7), and (8).

**Optimality of Differential Capital Taxation.** Now we prove the optimality of differential taxation of capital for the general environment with skill shocks. First, we define the intertemporal wedge for an agent with skill history \( h^t \) and for capital of type \( i \in \{s,e\} \) from period \( t \) to period \( t + 1 \), as

\[
\tau_{i,t+1}(h^t) = 1 - \frac{u'(c_t(h^t))}{\beta F_{i,t+1} E_t \{ u'(c_{t+1}(h^{t+1})) | h^t \}}.
\]

The first part of Proposition 4 generalizes Proposition 1 by showing that it is optimal to create a wedge between the marginal returns to structure and equipment capital when skills evolve stochastically over time. The second part of Proposition 4 shows that the optimal intertemporal wedges for structure and equipment capital are different. Thus, optimality of differential taxation of capital does not depend on the permanent skill types assumption maintained in the main model.

---

\(^8\)Temporary incentive constraints were first shown to be necessary and sufficient for incentive compatibility by Green (1987) for an environment with i.i.d. shocks. Fernandes and Phelan (2000) generalized this result to environments with persistent shocks. To be precise, we need to use two more assumptions to be able to use temporary incentive constraints. First, we need each 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.
Proposition 4. 1. At the constrained efficient allocation, in any period $t \geq 2$,

$$\tilde{F}_1(K^*_s,t, K^*_e,t, L^*_s,t, L^*_u,t) = \tilde{F}_2(K^*_s,t, K^*_e,t, L^*_s,t, L^*_u,t) + X^*_t / \lambda^*_t,$$

where

$$X^*_t = \sum_{h^t \in H^t} \mu^*_t(h^t) v' \left( \frac{l^*_t(h^{t-1}, h^*_t) w^*_t(h^*_t)}{w^*_t(h_t)} \right) l^*_t(h^{t-1}, h^*_t) \frac{\partial w^*_t(h^*_t)}{\partial K^*_e,t}$$

and $\lambda^*_t \beta^{t-1}$ and $\mu_t(h^t)$ are Lagrange multipliers on period $t$ feasibility constraint and the incentive constraint at history $h^t$.

2. (a) The optimal wedge on structure capital in any period $t \geq 2$ and history $h^{t-1}$ satisfies $\tau^*_s(t, h^{t-1}) \geq 0$. The inequality is strict if and only if there is no $h \in H$ such that $\pi_t(h^{t-1}, h | h^{t-1}) = 1$.

(b) The optimal wedge on equipment capital in any period $t \geq 2$ and history $h^{t-1}$ is

$$[1 - \tau^*_e(h^{t-1})] = [1 - \tau^*_s(h^{t-1})] \cdot \left[ 1 + X^*_t / \left( \lambda^*_t \tilde{F}^*_2 \right) \right].$$  \(10\)

Proof. Relegated to Appendix B.

The idea behind the first part of Proposition 4 is very similar to the one for Proposition 1: under equipment-skill complementarity, increasing the amount of equipment capital has an effect on incentives, summarized by the term $X^*_t$. In contrast, changing the amount of structure capital does not affect incentives. As a result, it is optimal to create a wedge between the physical returns to the two types of capital. The main distinction from the permanent type model is that, in the case with stochastic skills, a change in period $t$ equipment capital affects all the binding incentive constraints in that period. Thus, $X^*_t$ measures the cumulative effect of a change in equipment capital on all the binding incentive constraints. Since at this level of generality it is not possible to determine the pattern of binding incentive constraints, the sign of $X^*_t$ is ambiguous.

Part 2(a) of Proposition 4 states that the intertemporal wedge on structure capital is positive if there is skill risk. Intuitively, if we allow an agent to save at the marginal rate of
return to structure capital, he will save more than the efficient level. 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. The government achieves that with a positive wedge on structure capital.\footnote{The positive wedge on structure capital follows from the inverse Euler equation, see equation (16) in Appendix B. This condition was first derived by Rogerson (1985) and then generalized by Golosov, Kocherlakota, and Tsyvinski (2003). The inverse Euler equation does not hold for equipment capital because of the effect that equipment capital has on incentives. We derive a modified version of the inverse Euler equation for equipment capital in Appendix B, see equation (17).} Naturally, with permanent types there is no skill risk and, hence, no reason to tax structure capital, as we have already shown in Proposition 2.

Equation (10) in part 2(b) of the proposition is a version of the no-arbitrage condition for this economy. The equation shows that the intertemporal wedge on equipment capital can be decomposed into two parts. First, the government wants to discourage savings in equipment capital for the same reason that it wants to discourage savings in structure capital, which is captured by 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 the optimal wedge between the returns to the two types of capital. 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.

A Special Case. In Assumption 5 below, we assume that the only incentive constraints that bind are those that prevent the skilled from pretending to be unskilled. We call these incentive constraints downward incentive constraints. There is no theoretical result that establishes the pattern of binding incentive constraints for general skill processes in dynamic Mirrleesian environments, even when wages are exogeneous.\footnote{Downward incentive constraints are the only binding incentive constraints when skills are i.i.d. and wages are exogeneous.} Indeed, there are examples in which some upward incentive constraints bind. In this regard, Assumption 5 is stronger than Assumption 4, which we use in Section 3.

**Assumption 5.** In any period $t \geq 1$, history $h^t$, only downward incentive constraints bind.
capital return wedge, and show that the optimal equipment capital wedge is higher than the optimal structure capital wedge. These results are summarized by the following proposition.

**Proposition 5.** Suppose Assumption 5 holds. Then, in any period \( t \geq 2 \) and history \( h^{t-1} \),

\[
\bar{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) < \bar{F}_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) \quad \text{and} \quad \tau_{e,t}^*(h^{t-1}) > \tau_{s,t}^*(h^{t-1}).
\]

**Proof.** Relegated to Appendix B. \( \square \)

**Intratemporal Wedges.** Under Assumption 5, Proposition 6 generalizes the optimality of subsidizing skilled labor supply, shown for the permanent type case in Section 3.3, for the environment in which skills evolve stochastically over time. First, define the optimal intratemporal wedge at history \( h^t \) as \( \tau_{y,t}^*(h^t) = 1 - v'(l_t^*(h^t))/\left(w_t^*(h_t)u'(c_t^*(h^t))\right) \).

**Proposition 6.** Suppose Assumption 5 holds. In any period \( t \geq 1 \) and history \( h^{t-1} \),

\( \tau_{y,t}^*(h^{t-1}, s) < 0 \).

**Proof.** Relegated to Appendix B. \( \square \)

**Implementation.** In Appendix C, we provide an implementation of the constrained efficient allocation through a tax system in a competitive market environment in which agents trade a risk free bond and capital. The implementation result holds for any stochastic process, including the permanent type model. An interesting feature of this tax system is that the optimal tax differentials across equipment and structure capital can be implemented at the firm level, as is the case in the current U.S. tax system. This is possible because, as the second term on the right-hand side of equation (10) shows, the differential between optimal intertemporal wedges of structure and equipment capital is history independent in any period. Another notable feature of our implementation is that, the optimal tax system mimics the actual U.S. tax code in the sense that capital tax differentials are created through depreciation allowances that differ from actual economic depreciation. Therefore, creating the optimal capital tax differentials would not require complicating the U.S. tax code further.
5 Quantitative Analysis

The main goal of this section is to analyze the quantitative importance of differential taxation of capital in a calibrated version of our model. As in the main part of the paper, we assume that agents’ skill types are permanent. Since there is no labor income risk in this environment, the only role of taxation is redistribution (along with financing government consumption). 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.\(^{11}\)

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 with endogenous factor prices in which the planner “inherits” the initial capital stocks from the steady state of the competitive equilibrium.\(^{12}\) We solve for the whole time series of the constrained efficient allocation, thus taking into account the transition to a new steady state. Then, we recover the optimal wedges (taxes) from the constrained efficient allocation. In line with Proposition 2, we find that the optimal taxes on equipment capital are 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.

\(^{11}\)Keane and Wolpin (1997) estimate that initial conditions account for 90% of the cross-sectional variation in life-time earnings. Huggett, Ventura, and Yaron (2011) estimate this number to be over 60%, and Storesletten, Telmer, and Yaron (2004) estimate it to be almost 50%.

\(^{12}\)We would not be able to assess the role of differential capital taxation in a partial equilibrium environment, because the skill premium would not be affected by changes in the level of equipment capital. To the contrary, most quantitative papers in the NDPF literature consider partial equilibrium environments. As Farhi and Werning (2012) show, considering general equilibrium effects might be important even with a standard production function without complementarities.
5.1 Calibration

To calibrate the parameters in the social planning problem, we assume that the steady state of the competitive equilibrium defined in Appendix C represents the current U.S. economy. 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.

One period in our model corresponds to one year. We assume that the period utility function takes the form $u(c) - v(l) = c^{1-\sigma}/(1-\sigma) - \phi l^{1+\gamma}/(1+\gamma)$. In the benchmark case, we use $\sigma = 2$ and $\gamma = 1$. These are within the range of values that have been considered in the literature. We further assume that the production function takes the same form as in Krusell, Ohanian, Ríos-Rull, and Violante (2000):

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

Using U.S. data Krusell, Ohanian, Ríos-Rull, and Violante (2000) estimate $\alpha, \rho, \eta$, and we use their estimates, but they do not estimate $\omega$ and $\rho$. We calibrate these parameters to U.S. data, as we explain in detail below. This production function satisfies Assumptions 1 – 3 if $\eta > \rho$, which is what Krusell, Ohanian, Ríos-Rull, and Violante (2000) find.

We assume that the government consumption-to-output ratio equals 16%, which is close to the average ratio in the United States during the period 1980 – 2012, as reported in the National Income and Product Accounts (NIPA) data. We follow Heathcote, Storesletten, and Violante (2010) and assume a flat labor income tax rate of $\tau_y = 27\%$ (for a discussion of the construction of this number, see Domeij and Heathcote (2004)). Gravelle (2011) documents that because of differences in tax depreciation rates, the effective tax rates on structure capital and equipment capital differ at the firm level. Specifically, she estimates the effective corporate tax rate on structure capital to be 32%, and that on equipment capital to be 26%. The capital income tax rate at the consumer level is 15% in the U.S. tax code. This implies
Table 1: Benchmark Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative risk aversion</td>
<td>$\sigma$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Inverse Frisch elasticity</td>
<td>$\gamma$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Structure capital depreciation rate</td>
<td>$\delta_s$</td>
<td>0.056</td>
<td>GHK</td>
</tr>
<tr>
<td>Equipment capital depreciation rate</td>
<td>$\delta_e$</td>
<td>0.124</td>
<td>GHK</td>
</tr>
<tr>
<td>Share of structure capital in output</td>
<td>$\alpha$</td>
<td>0.117</td>
<td>KORV</td>
</tr>
<tr>
<td>Measure of elasticity of substitution between structure capital</td>
<td>$\eta$</td>
<td>0.401</td>
<td>KORV</td>
</tr>
<tr>
<td>and unskilled labor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure of elasticity of substitution between equipment capital</td>
<td>$\rho$</td>
<td>-0.495</td>
<td>KORV</td>
</tr>
<tr>
<td>and skilled labor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tax on labor income</td>
<td>$\tau_y$</td>
<td>0.27</td>
<td>HSV</td>
</tr>
<tr>
<td>Overall tax on structure capital income</td>
<td>$\tau_s$</td>
<td>0.422</td>
<td>Gravelle (2011)</td>
</tr>
<tr>
<td>Overall tax on equipment capital income</td>
<td>$\tau_e$</td>
<td>0.371</td>
<td>Gravelle (2011)</td>
</tr>
<tr>
<td>Government consumption</td>
<td>$G/Y$</td>
<td>0.16</td>
<td>NIPA</td>
</tr>
<tr>
<td>Relative supply of skilled workers</td>
<td>$p_s/p_u$</td>
<td>0.778</td>
<td>U.S. Census</td>
</tr>
<tr>
<td>Share of skilled workers’ wealth</td>
<td>$\zeta$</td>
<td>0.686</td>
<td>U.S. Census</td>
</tr>
</tbody>
</table>

This table reports the benchmark parameters that we take directly from the data or the literature. 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), respectively. NIPA stands for the National Income and Product Accounts.

an overall tax on structure capital $\tau_s = 1 - 0.85 \cdot (1 - 0.32) = 42.2\%$ and an overall tax on equipment capital $\tau_e = 1 - 0.85 \cdot (1 - 0.26) = 37.1\%$. These numbers are in line with a 40% tax on aggregate capital that is reported by Domeij and Heathcote (2004). We assume that unspent government tax revenue is distributed back to the agents in a lump-sum manner, which implies that in the SCE average taxes are in general not equal to marginal taxes. We set the ratio of skilled to unskilled agents to be consistent with the 2011 US Census data.

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 $\zeta$ 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 the benchmark parameters that we take directly from the data or the literature.

This leaves us with several parameters to be determined. We cannot identify $z_u$ and
Table 2: Benchmark Calibration Procedure

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Target</th>
<th>Data and SCE</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount factor</td>
<td>$\beta$</td>
<td>0.985</td>
<td>Capital-to-output ratio</td>
<td>2.9</td>
<td>NIPA and FAT</td>
</tr>
<tr>
<td>Disutility of labor</td>
<td>$\phi$</td>
<td>67.8</td>
<td>Labor supply</td>
<td>1/3</td>
<td></td>
</tr>
<tr>
<td>Production function parameter</td>
<td>$\omega$</td>
<td>0.477</td>
<td>Labor share</td>
<td>2/3</td>
<td>NIPA</td>
</tr>
<tr>
<td>Production function parameter</td>
<td>$\nu$</td>
<td>0.657</td>
<td>Skill premium $\frac{w_s}{w_u}$</td>
<td>1.8</td>
<td>HPV</td>
</tr>
</tbody>
</table>

This table reports our benchmark calibration procedure. The production function parameters $\nu$ and $\omega$ control the income share of equipment capital, skilled and unskilled labor in output. 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.

$z_s$ 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 $\omega$, the parameter that controls the income share of unskilled labor $\nu$, the labor disutility parameter $\phi$, and the discount factor $\beta$. 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 Fixed Asset Tables), (iii) the skill premium 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 commonly used in the macro literature). Table 2 summarizes our calibration procedure.

5.2 Quantitative Results

In this section, we analyze the quantitative properties of the optimal tax system. We do so by solving the social planning problem (SPP) defined in Section 3.2 with parameters calibrated in Section 5.1 to the U.S. economy. We assume that the planner inherits initial

---

13We 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.
Table 3: Steady-State Comparison of Wedges

<table>
<thead>
<tr>
<th>Tax (wedge) on equipment capital $\tau_e$</th>
<th>SCE</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax (wedge) on structure capital $\tau_s$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tax (wedge) on unskilled labor $\tau_y(u)$</td>
<td>27.00%</td>
<td>26.60%</td>
</tr>
<tr>
<td>Tax (wedge) on skilled labor $\tau_y(s)$</td>
<td>27.00%</td>
<td>-11.14%</td>
</tr>
</tbody>
</table>

This table compares the tax rates in the stationary competitive equilibrium (column SCE) and wedges at the steady state of the solution to the social planning problem (column SPP).

We first discuss the properties of the optimal tax system in steady state and compare it to the current U.S. tax system. The first column of Table 3 summarizes the current U.S. tax system. The second column reports its counterpart in the optimal tax system at the steady state. The first two rows of Table 3 report capital income taxes net of depreciation.\(^{14}\) We find that the equipment capital tax $\tau_e$ 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 $\tau_s$, which is zero. This is in contrast with the current effective tax rates in the United States where structure capital is taxed by 5.1 percentage points more than equipment capital overall. 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 in line with our theoretical results from Section 3.

\(^{14}\) We report capital income taxes net of depreciation rather than the capital wedges defined in Section 3.2 because the former 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 denoted “SPP,” the capital income taxes are recovered from the constrained efficient allocation by using the following definition for each type $h \in H$, capital type $i$, and period $t$: $\tau_{i,t+1}(h) \equiv 1 - \left( \frac{u'(c_{h,t+1})}{\beta u'(c_{h,t+1})} - 1 \right) / \left( F_{i,t+1} - \delta_i \right)$. Part 1 of Proposition 2 implies that these taxes only depend on time and not on agent type; therefore, we only report one number (time series).
Table 4: Steady-State Comparison of Allocations

<table>
<thead>
<tr>
<th></th>
<th>SCE</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_e/K_s$</td>
<td>1.02</td>
<td>0.93</td>
</tr>
<tr>
<td>$L_s/L_u$</td>
<td>0.82</td>
<td>1.11</td>
</tr>
<tr>
<td>$w_s/w_u$</td>
<td>1.80</td>
<td>1.47</td>
</tr>
</tbody>
</table>

This table compares allocations in the stationary competitive equilibrium (column SCE) and at the steady state of the solution to the social planning problem (column SPP). $K_e/K_s$ denotes the equipment-to-structure capital ratio, $L_s/L_u$ denotes the skilled-to-unskilled labor ratio and $w_s/w_u$ denotes the skill premium.

The higher taxes on equipment capital relative to structure capital, together with marginal subsidies on skilled labor, are used to indirectly redistribute from the skilled to the unskilled. 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 to the SCE. At the same time, the marginal subsidy on skilled labor income increases the ratio of skilled to unskilled labor. Both capital and labor taxes decrease the skill premium at the SPP. This way, the planner provides indirect redistribution from the skilled to the unskilled.

The marginal subsidy on skilled labor income seems to imply that there is direct 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 quite different from the marginal income tax.\footnote{Suppose, e.g., 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 gets a marginal subsidy of 10%.} 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 $1 - c_h/(w_h l_h)$ for agents of type $h$, following Farhi and Werning (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...
direct redistribution from the skilled to the unskilled.\textsuperscript{16}

**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 2. The equipment capital tax is growing over time. To understand this finding, 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. The overall capital stock is growing in the constrained efficient allocation because the planner inherits an inefficiently low level of capital from the SCE, which is due to the inefficiently high overall level of capital taxes at the SCE. As the quantity of equipment capital grows, so does the skill premium (see Figure 3). The government wants to prevent an unfettered growth of the skill premium because of its adverse redistributive effects. To keep the growth of the skill premium under control, the planner finds it optimal to increase the tax on equipment capital.\textsuperscript{17}

Optimal labor wedges are almost constant along the transition, as shown in the right panel of Figure 1. In fact, Werning (2007) shows that with our utility function labor wedges are exactly constant over time in a permanent-type model without equipment-skill complementarity. Figure 1 suggests that the extra distortions in labor wedges arising from equipment-skill complementarity are also 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, both types of agents become richer, and because of the income effect, they decrease their labor supply even though labor wedges

\textsuperscript{16}The non-linear nature of the optimal labor income tax code also explains how government budget is balanced under the optimal tax system. Table 3 seems to suggest that - except for a small increase in equipment capital taxes - government revenue from all other sources declines significantly when the economy moves from the current system to the optimal one. However, with a non-linear tax system the total amount of labor income taxes collected can increase even if the marginal taxes decline.

\textsuperscript{17}We check the validity of this intuition by conducting exercises, in which the planner inherits inefficiently high amounts of capital from the SCE. In those cases, as our intuition suggests, the planner decreases both types of capital over the transition to the new steady state, and optimal equipment taxes decline over the transition period.
This figure shows the paths of optimal taxes (wedges) along the transition to the new steady state at the solution to the social planning problem.

Figure 3 depicts the evolution of the optimal skill premium over time. First, the optimal skill premium is much lower in each period than it is 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.

**Welfare Gains of Optimal Differential Taxation of Capital.** We assess the importance of optimal differential taxation of capital by answering the following question: how much of the welfare gains of the full reform (which we call optimal DTC in this section) do we give up if we restrict the government to use the current capital taxes and allow it only to choose the labor income tax code optimally? To answer this question, we solve an additional 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 differential taxation of capital* (current DTC). We state the planning problem that gives rise to the current DTC in Appendix D. We find that, for the benchmark parameters, reforming the current tax system to the optimal DTC implies 0.19%
Figure 2: Factors of Production at the Constrained Efficient Allocation

This figure shows the paths of factors of production along the transition to the new steady state at the solution to the social planning problem.

Figure 3: Skill Premium at the Constrained Efficient Allocation

This figure shows the path of the skill premium $w_s/w_u$ along the transition to the new steady state at the solution to the social planning problem.
more welfare gains than reforming labor taxes alone (i.e., moving to the current DTC). The additional gains of optimal DTC can be as high as 0.40% for reasonable parameter values, as we discuss in more detail in the sensitivity analysis below.

In addition, we solve a version of the social planning problem, in which the planner is unrestricted in his choice of labor taxes, but 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 D. We find that the welfare gains of the current DTC falls 0.14% short of the welfare gains of the optimal NDTC for the benchmark parameters. This difference in welfare gains can be as high as 0.27% for reasonable parameter values.

We also assess how people rank the different capital tax reforms. We find that relative to the current DTC the optimal DTC helps both types. The reason is that the overall level of capital taxes at the current DTC is inefficiently high. Under the optimal DTC, structure capital taxes are zero while equipment capital taxes are virtually unchanged. As a result, there is more capital of both types at the optimal DTC. This increases the productivity of both types of agents, and they both benefit from this reform. We also find that relative to the current DTC, the optimal NDTC helps the skilled and hurts the unskilled.

**Sensitivity Analysis.** In each sensitivity exercise, we change the parameter of interest and redo the calibration procedure. Table 5 summarizes the sensitivity results. We report optimal taxes only for the optimal DTC reform. With a higher $\sigma$, the curvature of utility from consumption, the planner wants to provide more redistribution. Therefore, the indirect redistribution channel becomes more important. Hence, as $\sigma$ increases, the tax on equipment capital as well as the marginal subsidy to skilled labor increase. Table 5 also reports the

---

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

19 These results suggest that setting capital tax rates to a uniform rate, as proposed recently by President Obama’s administration, might imply substantial welfare gains. However, our results here are only suggestive, since that proposal only involves reforming capital taxes, but would leave labor taxes intact. We evaluate the consequences of such a proposal in a world with multiple layers of heterogeneity across agents in Slavik and Yazici (2014).
Table 5: Sensitivity Analysis

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th></th>
<th>Benchmark</th>
<th>Benchmark</th>
<th></th>
<th>Benchmark</th>
<th>Benchmark</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 4</td>
<td>1 1 1</td>
<td>1 0.5 1 2</td>
<td>2 2 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Optimal taxes

- $\tau_e$: 24.39% 39.54% 54.84%
- $\tau_s$: 0.00% 0.00% 0.00%
- $\tau_y(u)$: 18.77% 26.60% 33.75%
- $\tau_y(s)$: -5.29% -11.14% -21.23%

Optimal DTC vs. current DTC

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>1 2 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>1 1 1</td>
</tr>
</tbody>
</table>

Optimal NDTC vs. current DTC

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>1 2 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>0.5 1 2</td>
</tr>
</tbody>
</table>

This table reports the results of the sensitivity analysis. In the first column, $\sigma$ is intertemporal elasticity of substitution, $\gamma$ is the inverse Frisch elasticity of labor supply, $\tau_e$ is the equipment capital tax (wedge), $\tau_s$ is the structure capital tax (wedge), $\tau_y(u)$ is the labor wedge of the unskilled agents, $\tau_y(s)$ is the labor wedge of the skilled agents. For all taxes (wedges) we report their steady state values. Opt. DTC refers to optimal differential taxation of capital, i.e., a reform that reforms both labor and capital taxes. Current DTC refers to current differential taxation of capital, i.e., a tax reform that reforms labor taxes, but leaves capital taxes at their current values. Opt. NDTC refers to optimal nondifferential taxation of capital, i.e., a reform in which the planner is free to adjust labor taxes, but must set the tax rate on equipment capital equal to the tax rate on structure capital.

The sensitivity of our results to changes in $\gamma$, the curvature of disutility from labor. We find that, as $\gamma$ decreases, the tax on equipment capital and the skilled labor subsidy increase.

As we report in the penultimate row of Table 5, the welfare gains of the optimal DTC reform are around 0.20% higher than the gains of the current DTC reform for all values of $\sigma$ and $\gamma$ we consider.\(^\text{20}\) The welfare gains of optimal NDTC relative to current DTC are decreasing in $\sigma$ and increasing in $\gamma$, as shown in the last row of Table 5. The reason is that with a larger $\sigma$ or lower $\gamma$, the optimal capital tax differential is larger, as one can see in the rows denoted by $\tau_e$ and $\tau_s$ in Table 5. Therefore, optimal NDTC, which forces capital taxes to be uniform, is more restrictive and implies smaller welfare gains for higher $\sigma$ or lower $\gamma$.

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 current DTC are as high as 0.28% for

\(^{20}\) We also compute the welfare gains of optimal DTC under alternative social welfare weights. We find that if the planner cares more about the unskilled, the welfare gains of optimal DTC are larger. This is intuitive: not being able to use one of the channels of indirect redistribution optimally has more severe welfare consequences when society care more about redistribution.
\( \sigma = 4 \) and \( \gamma = 0.5 \). We also consider a higher elasticity of substitution between equipment capital and unskilled labor, namely, \( \eta = 0.79 \). This value has been used in He and Liu (2008), who use the same production function, and is based on an empirical estimate by Duffy, Papageorgiou, and Perez-Sebastian (2004). For this value of \( \eta \) and \( \sigma = 4 \) and \( \gamma = 0.5 \), the welfare gains of optimal DTC relative to current DTC are 0.40%.

6 Conclusion

The effective marginal tax rates on returns to capital assets differ substantially depending on the capital asset type in the U.S. tax code. 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 redistribution and insurance. Contrary to the actual practice in the U.S. tax code, we show that, under a plausible assumption, 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 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 differential capital taxation can be as high as 0.4% of lifetime consumption.
Acknowledgements

We thank the associate editor, Christopher Sleet, an anonymous referee, as well as Laurence Ales, Alex Bick, Felix Bierbrauer, V. V. Chari, Nicola Fuchs-Schündeln, Kenichi Fukushima, Mike Golosov, Christian Hellwig, Larry Jones, Juan Pablo Nicolini, Fabrizio Perri, Dominik Sachs, Florian Scheuer, Yuichiro Waki, and participants at various conferences, workshops and seminars for their comments. Yazici greatly appreciates IIES (Stockholm University) for their hospitality during his visit.

References


Appendix

A 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 point, let $\rho_i$ be the return to capital type $i$, $\delta_i$ 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 $\tau$ be the statutory tax rate, the effective tax rate, call it $\tau_i$, 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.\(^{21}\)

\(^{21}\)Of 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
Gravelle (2003) concisely summarizes the historical evolution of tax differentials. 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 reduce the tax differentials between equipment capital and structure capital. These policy changes resulted in effective 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.

B Proofs

B.1 Proof of Proposition 2

We call the multipliers on period $t$ feasibility and skilled agents’ incentive constraint to be $\lambda_t \beta^{t-1}$ and $\mu$, respectively. The first-order optimality conditions with respect to $c_{s,t}$ and $c_{u,t}$ are given by, for $t \geq 1$,

\begin{align}
(c_{s,t}) : (\pi_s + \mu^*) u'(c_{s,t}^*) &= \lambda^*_t \pi_s \\
(c_{u,t}) : (\pi_u - \mu^*) u'(c_{u,t}^*) &= \lambda^*_t \pi_u,
\end{align}

rate on capital. Inflation also affects the real value of depreciation deductions because these deductions are based on the historical acquisition costs. For a detailed description of how effective tax rates are calculated, see Gravelle (1994).

\footnote{For a more detailed description of how tax differentials between capital equipments and structures have changed between 1950 and 1983, see Auerbach (1983).}
which imply for all \( h \in H \) and \( t \geq 2 \),
\[
\frac{\lambda_{t-1}^*}{\lambda_t^*} = \frac{u'(c_{h,t-1}^*)}{u'(c_{h,t}^*)}.
\] (13)

The first order optimality conditions with respect to the two types of capital for \( t \geq 2 \) are:

\[(K_{e,t}) : \lambda_{t-1}^* = \beta \left[ \lambda_t^* \tilde{F}_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) + X_t^* \right],\]
\[(K_{s,t}) : \lambda_{t-1}^* = \beta \lambda_t^* \tilde{F}_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*),\] where
\[X_t^* = \mu^* u' \left( \frac{l_{u,t}^* w_{u,t}^*}{w_{s,t}^*} \right) l_{u,t}^* \frac{\partial}{\partial K_{e,t}^*} \left( \frac{w_{u,t}^*}{w_{s,t}^*} \right).\]

Combining the first-order conditions with respect to capital with (13), we get for both \( h \in H \) and \( \forall t \geq 2 \),
\[u'(c_{h,t-1}^*) = \beta \tilde{F}_{1,t}^* u'(c_{h,t}^*),\] (14)
\[u'(c_{h,t}^*) = \beta \tilde{F}_{2,t}^* u'(c_{h,t}^*) \left( 1 + \frac{X_t^*}{\lambda_t^* \tilde{F}_{2,t}^*} \right).\] (15)

**Part 1.** Equation (14) together with the definition of intertemporal wedges proves that the structure capital wedge is zero in all periods for all \( h \in H \). Since \( \mu^* > 0 \), we have \( X_t^* < 0 \). Using the definition of the equipment capital wedge and equation (15), we have for all \( h \in H \) and \( t \geq 2 \),
\[\tau_{e,t}^*(h) = - \frac{X_t^*}{\lambda_t^* \tilde{F}_{2,t}^*} > 0.\]
This finishes the proof of part 1 of Proposition 2.

**Part 2.** Suppose a steady state of the constrained efficient allocation exists. Letting the allocation without any time subscripts denote this steady-state allocation, we have
\[X^* = \mu^* u' \left( \frac{l_s^* w_u^*}{w_s^*} \right) l_u^* \frac{\partial}{\partial K_e^*} \left( \frac{w_u^*}{w_s^*} \right) < 0.\]
Thus, we have
\[ \tau^*_e = -\frac{X^*}{\lambda^*_t F^*_2} > 0. \]

**B.2 Proof of Proposition 3**

As before, we call the multipliers on period \( t \) feasibility and skilled agents’ incentive constraint to be \( \lambda_t \beta^{-1} \) and \( \mu_t \), respectively. Consider the first-order optimality conditions with respect to \( c_{s,t} \) and \( l_{s,t} \):

\[
\begin{align*}
    u'(c^*_{s,t}) A^*_t &= \lambda^*_t \pi_s \\
    v'(l^*_{s,t}) \left( A^*_t - \frac{B^*_t}{v'(l^*_{s,t})} \right) &= \lambda^*_t w^*_{s,t} \pi_s, \quad \text{where} \\
    A^*_t &= (\pi_s + \mu^*) \\
    B^*_t &= \mu^* v' \left( \frac{l^*_{u,t} w^*_{u,t}}{w^*_{s,t}} \right) l^*_{u,t} \frac{\partial w^*_{s,t}}{\partial L^*_{s,t}} z_s \pi_s.
\end{align*}
\]

By the first order conditions above, we have \( A^*_t > 0 \) and \( \left( A^*_t - \frac{B^*_t}{v'(l^*_{s,t})} \right) > 0 \). By Assumption 3, we have \( B^*_t > 0 \). These imply

\[ \tau^*_{y,t}(s) = 1 - \frac{A^*_t}{\left( A^*_t - \frac{B^*_t}{v'(l^*_{s,t})} \right)} < 0. \]

**B.3 Proof of Proposition 4**

**Part 1.** Let \( \lambda_t \beta^{-1} \) and \( \mu_t(h^t) \) be multipliers on period \( t \) feasibility constraint and the incentive constraint at history \( h^t \), respectively. Under Assumptions 1 and 2, the result follows from
the first-order conditions with respect to the two capital types:

\[
(K_{s,t}) : -\lambda_{t-1}^* + \beta \lambda_t^* F_1(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*) = 0,
\]

\[
(K_{e,t}) : -\lambda_{t-1}^* + \beta \lambda_t^* F_2(K_{s,t}^*, K_{e,t}^*, L_{s,t}^*, L_{u,t}^*)
+ \beta \sum_{\{h^t \in H^t\}} \mu_t^*(h^t) v' \left( \frac{l_t^*(h_t^t, h_t^o) w_t^*(h_t^o)}{w_t^*(h_t)} \right) l_t^*(h_t^t, h_t^o) \frac{\partial w_t^*(h_t^o)}{\partial K_{e,t}^*} = 0.
\]

**Part 2.** Also, under Assumption 1, it follows directly from Golosov, Kocherlakota, and Tsyvinski (2003) that in any period \( t \geq 1 \) and following any history \( h^t \), the constrained efficient allocation satisfies the following inverse Euler equation:

\[
\frac{1}{u'(c_t^*(h^t))} = \frac{1}{\beta \bar{F}_{1,t+1}} E_t \left\{ \frac{1}{u'(c_{t+1}^*(h^{t+1}))} \left| h^t \right. \right\}.
\] (16)

Part 2(a) of the proposition follows from the definition of intertemporal wedges in equation (9), equation (16) and the conditional version of Jensen’s inequality. Part 2(b) of the proposition follows from Part 1 and equation (9), which defines intertemporal wedges for both types of capital. \( \square \)

There is a modified inverse Euler equation for equipment capital. We derive it here for the sake of completeness. It follows from the first part of Proposition 4 and equation (16):

\[
\frac{1}{u'(c_t^*(h^t))} = \frac{1}{\beta (\bar{F}_{2,t+1}^* + X_{t+1}^*/\lambda_t^*)} E_t \left\{ \frac{1}{u'(c_{t+1}^*(h^{t+1}))} \left| h^t \right. \right\}. \] (17)

**B.4 Proof of Proposition 5**

First, under Assumption 5, for \( t \geq 2 \)

\[
X_t^* = \sum_{\{h_t^t \in H_t \}} \mu_t^*(h_t^t, s) v' \left( \frac{l_t^*(h_t^t, u) w_{u,t}^*}{w_{s,t}^*} \right) l_t^*(h_t^t, u) \frac{\partial w_{u,t}^*}{\partial K_{e,t}^*}.
\]
Assumption 2 implies $\partial w^*_t / \partial K_{e,t}^* < 0$, implying that, in any period $t \geq 2$, $X_t^* < 0$. The first part of the proposition then follows from the first part of Proposition 4 and $X_t^* < 0$, whereas the second part follows from the second part of Proposition 4 and $X_t^* < 0$.

B.5 Proof of Proposition 6

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 $l_t(\bar{h}_t)$:

$$u'(c_t^*(\bar{h}_t))A_t^*(\bar{h}_t) = \lambda_t^* \pi_t(\bar{h}_t),$$

$$v'(l_t^*(\bar{h}_t)) \left( A_t^*(\bar{h}_t) - \frac{B_t^* \pi_t(\bar{h}_t)}{v'(l_t^*(\bar{h}_t))} \right) = \lambda_t^* w_t^*(\bar{h}_t) \pi_t(\bar{h}_t),$$

where, letting $\chi$ be the indicator function,

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

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

By the first order conditions above, we have $A_t^*(\bar{h}_t) > 0$, $\left( A_t^*(\bar{h}_t) - \frac{B_t^* \pi_t(\bar{h}_t)}{v'(l_t^*(\bar{h}_t))} \right)$ and by Assumption 3, $B_t^* > 0$. These imply

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

C 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 non-contingent 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 representative firm that rents capital at a net interest rate $r_t$ and labor to produce the output good. 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 taxes the firms’ capital income net of depreciation. The statutory depreciation allowance can differ from economic depreciation, as in the U.S. tax code.

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 $y^t = (y_1, ..., y_t)$ pays in period $t$. Labor income in history $h^t$ is defined as $y_t(h^t) = w_t(h_t) \cdot l_t(h^t)$. Labor income $y_t$ and labor $l_t$ are one to one for a given level of wages, which means it is sufficient to use either one of them when we define an allocation. In the rest of this section, we use $y_t$. There are also linear taxes on people’s asset holdings, which 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.$^{24}$

---

$^{23}$The 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.

$^{24}$One can show that the government can implement the constrained efficient allocation by taxing only the
Unlike in Kocherlakota (2005), the government also has access to a sequence of linear taxes on firm’s capital income, denoted by \((\tau_{f,t})_{t=1}^{\infty}\). The corporate tax code also includes statutory depreciation allowances, \((\bar{\delta}_{s,t}, \bar{\delta}_{e,t})_{t=1}^{\infty}\). The firm is allowed to deduct \(\bar{\delta}_{t,t}K_{i,t}\) 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

\[
\max_{c,y,a} \sum_{t=1}^{\infty} \sum_{h^{t} \in H^{t}} \pi_{t}(h^{t}) \beta^{t-1} \left[ u(c_{t}(h^{t})) - v \left( \frac{y_{t}(h^{t})}{w_{t}(h^{t})} \right) \right] \\
\text{s.t.} \\
c_{t}(h^{t}) + a_{t+1}(h^{t}) \leq y_{t}(h^{t}) - \tau_{y,t}(y_{t}(h^{t})) + \left[ 1 - \tau_{a,t}(y_{t}(h^{t})) \right] (1 + r_{t})a_{t}(h^{t-1}), \\
a_{1} = K_{s,1}^{*} + K_{e,1}^{*}, \quad c, y \text{ are nonnegative.}
\]

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

\[
\max_{K_{s,t},K_{e,t},L_{s,t},L_{u,t}} \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r_{t})(K_{s,t} + K_{e,t}) - \frac{w_{s,t}}{z_{s}}L_{s,t} - \frac{w_{u,t}}{z_{u}}L_{u,t} - \tau_{f,t}\Pi_{f,t},
\]

where \(\Pi_{f,t}\) is the firm’s capital income net of depreciation allowances in period \(t\) given 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, \((c_{t}(h^{t}), y_{t}(h^{t}), a_{t+1}(h^{t}))_{t=1}^{\infty}\), 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 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.
lem, \((K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})_{t=1}^\infty\) solves the firm’s problem, and markets clear:

\[
\sum_{h^t \in H^t} \pi_t(h^t)c_t(h^t) + K_{s,t+1} + K_{e,t+1} + G_t = \tilde{F}(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}),
\]

\[
K_{s,t} + K_{e,t} = \sum_{h^t \in H^t} \pi_t(h^t)a_t(h^{t-1})
\]

\[
L_{s,t} = \sum_{\{h^t \in H^t | h_t = s\}} \pi_t(h^t) \frac{y_t(h^t)}{w_{s,t}} z_s,
\]

\[
L_{u,t} = \sum_{\{h^t \in H^t | h_t = u\}} \pi_t(h^t) \frac{y_t(h^t)}{w_{u,t}} z_u.
\]

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

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

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 \((\tau_{y,t}, \tau_{a,t}, \tau_{f,t}, \tilde{\delta}_{s,t}, \tilde{\delta}_{e,t})_{t=1}^\infty\) implements the constrained efficient allocation \((c^*_t(h^t), y^*_t(h^t), K^*_s,t, K^*_e,t, L^*_s,t, L^*_u,t)_{t=1}^\infty\) if an allocation for consumers \((c^*_t(h^t), y^*_t(h^t), a_{t+1}(h^{t+1}))_{t=1}^\infty\) 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.

**C.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

\[
\tau^*_{a,t+1}(y^{t+1}) = 1 - \frac{u'(c^*_t(h^t))}{\beta u'(c^*_{t+1}(h^{t+1}))F_1(K^*_s,t, K^*_e,t, L^*_s,t, L^*_u,t)}, \text{ if } y^{t+1} \in Y^{t+1*}
\]

\[
\tau^*_{a,t+1}(y^{t+1}) = 1, \text{ if else,}
\]
where we define \( y_t^i(h^t) = (y_m(h^m))_{m=1}^t, Y^{ts} = \{y_t^i : y_t^i = y_t^{ts}(h^t), h^t \in H^t \}. \)

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

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

If \( y_t^i \notin Y^{ts} \), then \( \tau^*_y(y_t^i) = 2y_t^i. \)

Finally, set

\[
\tau^*_f = 1 - \frac{F_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \delta_s}{F_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \delta_e} \quad \delta^*_s = r_t^* + \delta_s, \quad \delta^*_e = \delta_e.
\]

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 A.

**Proposition 7.** The optimal tax system described above implements the constrained efficient allocation.

**Proof.** First, define prices as \( r_t^* = F_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \delta_s \) and \( \forall h \in H, w_{h,t}^* = \frac{\partial F(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})}{\partial L_{u,t}}z_h. \) 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 \( t, y_t^i \) should be in \( Y^{ts} \)).

Third, we claim that if an agent chooses an income strategy \( y_t' \), where this means \( y_t'(h^t) = y_t^*(h^t) \) for some \( h^t \), then facing the prices defined above, the agent also chooses \((c', a')\), meaning that \( c_t'(h^t) = c_t^*(h^t) \) and \( a_{t+1}'(h^t) = a_{t+1}^*(h^t) \) 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
\[
\max_{c,a} \sum_{t=1}^{\infty} \sum_{h^t \in H^t} \pi_t(h^t) \beta^{t-1} \left[ u(c_t(h^t)) - v \left( \frac{y^*_t(h^t)}{u^*_t(h^t)} \right) \right] \quad \text{s.t.}
\]
\[
c_t(h^t) + a_{t+1}(h^t) \leq y^*_t(h^t) - \tau^*_{a,t}(y^*(\hat{h}^t)) + \left[ 1 - \tau^*_{a,t}(y^*(\hat{h}^t)) \right] (1 + r^*_t) a_t(h^{t-1}),
\]
\[
a_1 \leq K^*_{s,1} + K^*_{e,1}, \quad c \text{ is nonnegative.}
\]

The first-order conditions to this problem are the budget constraint with equality and
\[
u'(c_t(h^t)) = (1 + r^*_{t+1}) \beta \sum_{h^{t+1}|h^t} \pi_{t+1}(h^{t+1}|h^t) u'(c_{t+1}(h^{t+1})) [1 - \tau^*_{a,t}(y^{t+1*}(\hat{h}^t, h_{t+1})].
\]

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
\[
1 - \tau^*_{a,t+1}(y^{t+1*}(\hat{h}^t, h_{t+1})) = \frac{u'(c^*_t(\hat{h}^t))}{\beta u'(c^*_{t+1}(\hat{h}^t, h_{t+1})) \bar{F}^*_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t})}.
\]

Finally, 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 firm’s optimality conditions for capital are

\[
(K_{s,t}) : \bar{F}_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r^*_t) - \tau^*_f \left[ F_1(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \bar{\delta}_{s,t} \right] = 0
\]
\[
(K_{e,t}) : \bar{F}_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - (1 + r^*_t) - \tau^*_f \left[ F_2(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}) - \bar{\delta}_{e,t} \right] = 0.
\]

At the constrained efficient allocation, the first condition above holds by construction of $r^*_t$ and $\bar{\delta}^*_s$ whereas the second one holds by construction of $\tau^*_f$, $r^*_t$ and $\bar{\delta}^*_e$. \hfill \Box
**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 if only downward incentive constraints bind: \( \forall t \geq 2 : \tau_{f,t} > 0 \), where the inequality comes from Proposition 5.

3. There is a single statutory tax rate on corporate income, \( \tau_{f} \), but 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. These optimal tax depreciation allowances 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 \( \tau_{f,t} \). \(^{25}\)

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):

\[
E_t \left\{ 1 - \tau_{t+1}^{a} (y_{t+1}) \right\} = E_t \left\{ \frac{u'(c_{t+1}^*(h^t))}{\beta u'(c_{t+1}^*(h^{t+1})) F_{1,t+1}^*} \right\} = 1.
\]

This property follows from the inverse Euler equation (16). \(^{25}\)

\(^{25}\)Note 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.
D Alternative Tax Systems

D.1 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 a set of constraints for both types of agents in each period (here, \( \tau_s \) and \( \tau_e \) are fixed over time and taken from the data, namely, \( \tau_s = 0.422 \) and \( \tau_e = 0.371 \)): for all \( h \in H \)

\[
1 - \tau_s = \frac{u'(c_{h,t})}{F_{1,t+1} - \delta_s} - 1 \quad \text{and} \quad 1 - \tau_e = \frac{u'(c_{h,t})}{F_{2,t+1} - \delta_e}, \tag{18}
\]

Observe that the restrictions that the current capital taxes impose on the set of allocations that the planner can choose, (18), is written down taking into account that, in the U.S. tax code, it is capital income net of depreciation that is being taxed. Also, notice that any three of these constraints in (18) imply the fourth. Therefore, we write the current DTC planning problem, ignoring the fourth:

\[
\max \sum_{h \in H} \pi_h \sum_{t=0}^{\infty} \beta^t \left[ u(c_{h,t}) - v(l_{h,t}) \right] \quad \text{s.t.} \quad \forall t \geq 1, \quad G_t + \sum_h \pi_h c_{h,t} + K_{s,t+1} + K_{e,t+1} \leq \tilde{F}(K_s,t, K_e,t, L_{s,t}, L_{u,t}),
\]

\[
\forall t \geq 1, \quad \sum_{t=1}^{\infty} \beta^{t-1} [u(c_{s,t}) - v(l_{s,t})] \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[ u(c_{u,t}) - v(l_{u,t} w_{u,t} w_{s,t}) \right],
\]

\[
\forall t \geq 1, \quad \beta u'(c_{s,t+1}) [(1 - \tau_s)(F_{1,t+1} - \delta_s) + 1] = u'(c_{s,t}),
\]

\[
\forall t \geq 1, \quad \beta u'(c_{s,t+1}) [(1 - \tau_e)(F_{2,t+1} - \delta_e) + 1] = u'(c_{s,t}),
\]

\[
\forall t \geq 1, \quad \beta u'(c_{u,t+1}) [(1 - \tau_s)(F_{1,t+1} - \delta_s) + 1] = u'(c_{u,t}).
\]
D.2 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 agents of both types $h \in H$ and all $t \geq 1$, we impose

$$1 - \tau_{s,t+1}(h) = 1 - \tau_{e,t+1}(h).$$

This restriction on taxes is equivalent to the following restriction on allocations: for all $h \in H$ and $t \geq 1$,

$$1 - \tau_{s,t+1}(h) = \frac{u'(c_{h,t})}{F_{1,t+1} - \delta_s} - 1 - \tau_{e,t+1}(h) = \frac{u'(c_{h,t})}{F_{2,t+1} - \delta_e} - 1.$$

This equation 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:

$$\max_{\{(c_{h,t}, l_{h,t}), K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}\}} \sum_{h=u,s} \pi_h \sum_{t=0}^{\infty} \beta^{t-1} [u(c_{h,t}) - v(l_{h,t})] \quad \text{s.t.}$$

$$\forall t \geq 1, \quad G_t + \sum_h \pi_h c_{h,t} + K_{s,t+1} + K_{e,t+1} \leq F(K_{s,t}, K_{e,t}, L_{s,t}, L_{u,t}),$$

$$\sum_{t=1}^{\infty} \beta^{t-1} [u(c_{s,t}) - v(l_{s,t})] \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[ u(c_{u,t}) - v\left(\frac{l_{u,t} w_{u,t}}{w_{s,t}}\right) \right],$$

$$\forall t \geq 1, \quad F_1(K_{s,t+1}, K_{e,t+1}, L_{s,t+1}, L_{u,t+1}) - \delta_s = F_2(K_{s,t+1}, K_{e,t+1}, L_{s,t+1}, L_{u,t+1}) - \delta_e.$$