Optimal Uniform Capital Taxation in A Partially Automated Society *

Hideto Koizumi

October 22, 2019

Abstract

A recent rapid-automation movement has been displacing routine labor and has sparked a series of discussion about taxation on automation such as a robot tax. However, the government's dilemma is that the planner may want to tax such physical capital that displaces routine labor—for example, industrial robots—for redistributive motives but does not want to tax other physical capital that increases such workers' productivities—for instance, collaborative robots (cobots). This paper studies a novel setting of the optimal capital taxation on physical capital in which there is asymmetric information on both labor types and capital types between the planner and market. In particular, my model focuses on a two-by-two scenario where there are two types of labor (routine and non-routine labor) and two types of capital (*displacing* and *reinstating* capital). Despite asymmetric information, I find that the optimal uniform capital tax rate over different types of capital is strictly positive, as long as the solution is interior.

Keywords: Automation, Optimal Capital Taxation, Optimal Taxation, Robot Tax

JEL Codes: H21, H25

*The author thanks Benjamin Lockwood, Juuso Toikka, Alexander Rees-Jones, and Kent Smetters for extensive guidance and instructions in this work. The author also thanks Eduardo Azevedo for comments and guidance, and thanks seminar audiences for useful feedback. This project is funded by Mack Institute Research Grants for presentations abroad.

"“I'm sure I can come up with a robot that isn't a robot, according to the [robot] tax code.”—Shu-Yi Oei, a Boston College law professor, excerpted from Wall Street Journal article “The ‘Robot Tax’ Debate Heats Up” on January 8, 2020.

1 Introduction

The advent of robots in the workplace has led to a debate about whether and how to tax this form of capital. According to Acemoglu and Restrepo (2017), an additional robot per thousand workers decreases the employment to population ratio in the U.S. by approximately 0.18-0.34 percentage points and wages by 0.25-0.5 percent. Given that robots are assumed to displace routine-job workers, Guerreiro et al. (2019) theoretically found that the optimal robot tax rate is strictly positive for redistributive motives. The intuition behind their result is similar to the indirect taxation idea by Naito (1999). By taxing the complement of high-skilled non-routine workers and the substitute of low-skilled routine workers, the planner can indirectly transfer wealth from the poor to the rich, in addition to an income tax.

However, administrative tasks of identifying such displacing capital are not easy. While we want to tax physical capital that displaces routine labor for redistribution of wealth, we do not want to tax capital that increases the productivity of such labor[1]. For example, there are industrial robots that displace assembly line workers, while there are so-called collaborative robots, *cobots*, designed to augment the productivity of assembly line workers. Even in the cases where one capital uses much more sophisticated technology than the other, there are still high administrative costs for screening tasks. There are numerous examples of these

[1]Sachs and Kotlikoff (2012); Sachs et al. (2015) describe a similar problem with a focus on skill differences across ages. Unlike my study, they do not have asymmetric information on capital types—in their words, robots vs. traditional capital.

cases such as self-check-out cash registers and conventional cash registers, self-driving trucks and conventional trucks, and so forth. We want to tax only capital that displaces routine labor, but identifying such capital is administratively costly and may not even be feasible. Furthermore, even if the planer pays these administrate costs and impose a special tax (e.g., a robot tax from Guerreiro et al. (2019) and automation tax suggested by Acemoglu et al. (2020)), tax avoidance will be a concern as encapsulated by the excerpt introduced in the very beginning[2].

To address this issue, this paper analyzes a novel setting of the optimal capital taxation where there is asymmetric information between the planner and market on both labor and capital types. My model focuses on a two-by-two scenario: two types of labor and two types of capital. Workers are categorized into routine and non-routine labor[3]. Households are born into one of the labor types and face sufficiently high barriers to switching between the two sectors, while they can adjust labor supply amounts[4]. Capital is dichotomized into (i) *displacing capital*, which is a complement to non-routine labor but a perfect substitute to routine labor such as robots and self-driving trucks and (ii) *reinstating capital*, which is more complementary to routine labor than to non-routine labor such as cobots and conventional trucks. The two types of capital are perfect substitutes. My main interests lie in settings where the government cannot identify or discriminate between labor types and between capital types. The planner wants to maximize the utilitarian social welfare function by

[2]See, e.g., Slemrod and Yitzhaki (2002) and Slemrod and Kopczuk (2002)

[3]I do not use the conventional terms of low versus high skills since some routine work requires relatively higher skills than low skill.

[4]Naito (1999), Slavík and Yazici (2014), and Guerreiro et al. (2019) make the same assumption on the occupational choice barriers. Saez (2004) allows for occupational choice but not labor supply amounts, and find no distortion in production to be optimal. Gomes et al. (2017) allow for both dimensions to vary and find production inefficiency at optimum.

setting tax schedules.

Note that the planner faces the following asymmetric information problem in capital markets. On one hand, by taxing *just* displacing capital, the planner can reduce the wage gap between the two. Taxing a complement of non-routine workers will decrease non-routine-worker wage rates while taxing a substitute for routine workers will increase routine-worker wage rates. The reduction of the wage gap will relax the incentive compatibility constraint of non-routine workers to mimic and reduce their hours of work to that of routine workers—that is, a reduction in the information rent of non-routine types—in the planner's welfare maximization program. Note that the informational rent is two-fold: the planner can only observe income level but cannot observe labor supply amounts and cannot observe agent types. The informational gain from the tax achieves greater redistribution.

On the other hand, taxing *just* reinstating capital will decrease the wage rates of not only non-routine workers but also routine workers since reinstating capital is a complement to both types of labor. A decrease in the routine-workers' wage rates will decrease welfare and possibly dominate the positive redistribution effects from taxing displacing capital. Therefore, it is not immediate to see what the optimal uniform capital tax would be if there is asymmetric information on capital types as well.

Despite these competing forces, I find that positive taxes are useful for redistributive purposes even in the presence of asymmetric information. That is, even when the planner can distinguish between neither labor types nor capital types, the optimal uniform capital tax rate is strictly positive, regardless of the degree of gross-complementarity between routine labor and reinstating capital, as long as the solution is interior. By interior solution, I mean a solution at which all the four inputs are used with positive amounts in production.

Note that there are two corner solutions: one without any use of routine labor and thus reinstating capital, and the other without any use of displacing capital and with the composite of routine labor and reinstating capital and non-routine labor. At the former corner solution, only displacing capital and non-routine labor will be used in production. Then, the setting becomes equivalent to one of the corner solutions from Guerreiro et al. (2019), and the optimal capital tax rate is zero. This is because there is no redistributive gain from capital taxation in the absence of routine labor. In contrast, at the latter corner solution, only the composite of routine labor and reinstating capital together with non-routine labor will be used in production. In this case, the optimal capital tax rate is generally negative since this setting in the absence of displacing capital tends to be the opposite case of the interior solution of Guerreiro et al. (2019). These are trivial cases and thus will not be explored in the main text.

A novel key force driving the result of non-trivial interior solution cases comes from the cost difference between the two types of capital. A representative producer has to use two different inputs for the joint production by routine workers and reinstating capital, while the producer needs only one input for the production by displacing capital. Then, for an interior solution, the price of reinstating capital tends to be lower than that of displacing capital; otherwise, the solution can be at corner and can use only displacing capital. Thus, if we impose a uniform (linear) tax on both types of capital, the tax burden is placed disproportionally on displacing capital than reinstating capital since displacing capital is more expensive. This differential tax burden reduces the wage gap between two worker types. The reduced wage gap relaxes the incentive compatibility constraints of non-routine workers, resulting in the first-order informational gain.

As for the magnitude of the optimal tax rate, I use three reference points. One is the first-best outcome in which the planner can differentiate between two types of labor. This is a rather ideal, impractical situation, while this outcome sets the upper bound of welfare.

Another reference point is the zero-capital-tax outcome. This is the lower bound of welfare when the planner only uses optimal labor income taxes but not capital taxes. The last one is the second-best outcome in which the government can differentiate between the two types of capital and thus can impose differential rates of capital tax. This benchmark is useful since this is an analogous case to the existing models such as Guerreiro et al. (2019). I numerically show that the third-best outcome of the main setting is comparable to the second-best outcome in terms of total welfare, indicating that the main result is not practically negligible. With respect to the consumption amount of routine-work households, the uniform capital taxation can increase the amount by 0.03 to 14% when the costs of capital are still high, depending on the degree of complementarity between routine workers and reinstating capital. This magnitude dies out when technological progress reduces the costs of capital over time.

The main contribution of this paper is two-fold. First, this paper studies a novel setting of the optimal capital taxation where there is asymmetric information on both labor and capital types between the social planner and the market. This setting is especially relevant to the current economy and debate over taxing robots. Second, the findings of this paper provide a simple and practical solution that has a significant policy implication in the current, partially automated society.

1.1 Literature Review

Among many related papers[5], the closest paper to mine is Guerreiro et al. (2019)[6]. They question if robots should be ever taxed, using the task-based framework emphasized by Autor et al. (2003) and Acemoglu and Restrepo (2018). In their settings, there are two types of labor between whom the policymaker cannot identify, and one type of capital/intermediate goods—robots. Robots are complements to non-routine workers, but substitutes for routine workers. Guerreiro et al. (2019) find an optimal robot tax to be, in general, strictly positive. The intuition is similar to the one behind indirect taxation such as Naito (1999) and Slavík and Yazici (2014)[7]; the robot tax can be used for the planner's informational gain on labor types by decreasing the rent of non-routine workers to masquerade as routine workers. In other words, the robot tax can be used as another instrument to decrease the wage gap of the two worker types for a welfare gain at the expense of production loss that tends to be second-order[8].

I use this rich baseline framework from Guerreiro et al. (2019) and add two additional components[9]. One is reinstating capital introduced in the introduction[10]. The other one is asymmetric information on capital types. As emphasized in the introduction, these additions

[5]See (1) the direct taxation literature examplified by Diamond and Mirrlees (1971a,b), Atkinson and Stiglitz (1976), Deaton (1979), Saez (2001, 2004), Rothschild and Scheuer (2013), Scheuer (2014), and Gomes et al. (2017), (2) the indirect taxation literature examplified by Naito (1999), Saez (2002), Slavík and Yazici (2014), and (3) the new dynamic public finance literature examplified by Golosov et al. (2003), Kocherlakota (2005), Golosov et al. (2006), Golosov et al. (2011), Farhi and Werning (2013), Golosov et al. (2016), Stantcheva (2017).

[6]Thuemmel (2018) is similar to Guerreiro et al. (2019)

[7]Saez (2002) also has the same indirect taxation idea but focuses on household preferences

[8]The production loss due to the tax on non-labor inputs is cushioned by an increase in the labor supply of routine workers which are complements to non-routine workers, in response to the wage increase.

[9]Displacing capital in my model corresponds to robots in their model.

[10]Slavík and Yazici (2014) have two different types of capital in their dynamic model: structure and equipment. Yet, they have no asymmetric information on capital types in their settings and thus do not allow for the situation of our interests

are not just theoretically interesting exercises, but also important elements for a practical policy recommendation.

Note that while Diamond and Mirrlees (1971a) finds no production distortion to be optimal, this does not apply to my model. Diamond and Mirrlees (1971a) demonstrates that if the planner can levy differential taxes on all factors of production (input and output), then the economy should lie at the production efficient frontier. This assumption in Diamond and Mirrlees (1971a) implies that the planner can distinguish between different labor types, violating the premise of my model. In fact, if the government can differentiate between different labor types, production efficiency is achieved in my model as well. Furthermore, Atkinson and Stiglitz (1976) finds no production distortion to be optimal even with non-linear income taxes, but their result does not apply to this paper, either. They assume that workers with different productivities are perfect substitutes, violating the assumption of my model.

Another paper close to mine is Costinot and Werning (2018). They study the optimal robot tax rate using a sufficient statistics approach. The goal of their study "is not to sign the tax on robots, nor to explore a particular production structure, but instead to offer tax formulas highlighting key sufficient statistics needed to determine the level of taxes, with fewer structural assumptions" (p.4). Without imposing any structural assumption, they find an optimal non-zero robot tax should converge to zero over time as the process of automation deepens.

The important assumption in their setting is that the planner know which firms use "new technology" such as robots and which firms use "old technology" involving conventional non-robotic physical capital. One can translate displacing capital into the "new technology firm" and reinstating capital into “old technology firm.” In this sense, they study a scenario where there is no asymmetric information on capital types in my model without structural assumptions. In this paper, my main focus is on a case with asymmetric information on capital types.

2 Model Environment: Key Players in the Economy And Equilibrium Conditions

My main focus is on a static model. In my static model, a capital tax can be interpreted in a way that (a) the federal planner reduces the total amount (undiscounted sum) of tax depreciation deductions at the federal level, or (b) the state planner imposes a sales tax on physical capital at the state level[11].

This paper follows the notations from Guerreiro et al. (2019), so that the readers can closely compare two papers. There are two types of households, whose utility is based upon consumption of private and public goods and whose disutility comes from labor. One type of household provides routine labor, while the other type provides non-routine labor. Note that following Naito (1999), Slavík and Yazici (2014), and Guerreiro et al. (2019), analysis is based upon the assumption that households face high enough occupational barriers between the two sectors to switch job types, while they are allowed to change labor supply amounts[12].

[11] As for (a), the U.S. federal government allows firms to deduct tax for the purchase of depreciable assets by the amount equal to the the asset values over time following the depreciation schedule set by the Internal Revenue Service. In the static version of my model, the federal planner can impose a tax on physical capital by reducing the total deduction amounts of such purchase. For example, if a firm purchases \$70,000 depreciable assets that follows 7-year depreciation schedule, currently, the firm is allowed to deduct the entire \$70,000 in year one (written off as an expense) or deduct \$10,000 over seven years. The federal planner can impose 10% capital tax by reducing the total deduction amount to \$63,000.

[12] Saez (2004) allows workers of different types to switch their jobs/sector to work while prohibit them to change labor supply amounts. He models that in the long run, the relative supply to each

Routine labor, non-routine labor, and displacing and reinstating capital all contribute to the production of the single consumption good. There is a continuum of symmetric tasks in a unit range $[0, 1]$ where routine workers together with reinstating capital compete against displacing capital[13]. I incorporate tasks in my model not only to closely follow Guerreiro et al. (2019) for comparison, but also to precisely elicit what assumptions warrant the aggregation of production. Intermediate goods producers for both capital types come from the perfectly competitive external global market. The final good producer faces a production function featuring constant returns to scale and thus faces zero profits.

2.1 Household

The economy has a continuum of households with a unit measure. The unit measure of households are decomposed to π_n non-routine worker households and π_r routine worker households, where subscript n and r denote the non-routine and routine labor types, respectively. A household of type j derives utility from consumption of private goods, c_j, and public goods, G. Each household draws disutility from the hours of labor it supplies, l_j. Every household has a unit of time per period, leading to $l_j \leq 1$. A household of type j's

occupation becomes infinitely elastic, even though workers of different types can be imperfect substitutes in production. This leads to the resurrection of the Diamond-Mirrlees and Atkinson-Stiglitz no production distortion results. In contrast, by allowing for both occupational choice and labor supply adjustments, Gomes et al. (2017) found production inefficiency at optimum. To elicit a different channel of production distortion result, I shut down both channels by assuming no occupational choice but labor supply to be adjustable.

[13] See Autor et al. (2003) for the importance of tasks performed by routine workers.

optimization problem is

$$\underset{c_j, l_j}{\text{maximize}}\ U_j = u(c_j, l_j) + v(G)$$

$$\text{subject to } c_j \leq w_j l_j - T(w_j l_j),$$

where w_j denotes the wage rate of type j and $T(.)$ indicates the income tax schedule.

For convenience, write $u_x = \partial u(c,l)/\partial x$ where $x = c, l$ and $u_{xy} = \partial^2 u(c,l)/\partial x \partial y$. I make the standard concavity and convexity assumptions that, $u_c > 0, u_l < 0, u_{cc}, u_{ll} < 0$, and that consumption and leisure are both normal goods. Furthermore, I assume that the utility functions satisfy the single-crossing property and thus, we are given $u_{ll} l / u_l + 1 - u_{cl} l / u_c > 0$[14]. Additionally, I assume that $v'(G) > 0, v''(G) < 0$ and that $u(c,l)$ satisfies the standard Inada conditions for interior solutions.

2.2 Final good producers

The representative producer of the final good employs non-routine workers (N_n) for the single non-automatable task, routine labor $(n(i))$ for symmetric, automatable task i, and buys intermediate goods, a displacing capital $(x_d(i))$ and reinstating capital $(x_r(i))$, for the same automatable task i as well. For routine task i, the final producer faces the following task-production function:

$$y(i) = x_d(i) + \left[\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta) n(i)^{\frac{q-1}{q}}\right]^{\frac{q}{q-1}}, \tag{1}$$

[14]While the single-crossing condition is usually for reducing the dimension of continuously many labor types, it is still useful to ease the analysis.

where q indicates an elasticity of substitution between routine workers and reinstating capital. For the two inputs to be gross-complements, q is assumed to be between 0 and 1. With these routine tasks, the producer has the following final production function:

$$Y = A \left(\int_0^1 y(i)^{\frac{\rho-1}{\rho}} di \right)^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^{\alpha} \quad \alpha \in (0,1), \rho \in [0, \infty), \tag{2}$$

where ρ is an elasticity of substitution between tasks. Now, let m denote the average amount of displacing capital used in routine tasks —equivalently, m is the portion of tasks in which the producer only uses displacing capital $m = \int_0^1 \mathbb{1}\{x_d(i) > 0, x_r(i) = n(i) = 0\} di$. For convenience, since tasks are symmetric and have no difference in needs for productivity, I write that for the range $[0, m]$, the final good producers only use displacing capital, while for the range $(m, 1]$, the final good producer only uses the composite of routine workers and reinstating capital. Thus, m is a choice variable for the final good producer.

Then, we can divide the integral into the two different ranges and rewrite the production function as:

$$Y = A \left[\int_0^m x_d(i)^{\frac{\rho-1}{\rho}} di + \left(\int_m^1 \left[\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta) n(i)^{\frac{q-1}{q}} \right]^{\frac{\rho-1}{\rho}} di \right)^{\frac{q}{q-1}} \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^{\alpha},$$
$$\alpha \in (0,1), \rho \in [0, \infty). \tag{3}$$

Note that this way of defining automated and not-yet automated tasks is different from Acemoglu and Restrepo (2018). I abstract away from the *quality* of tasks a new technology destroys and creates but rather focus on the *quantity* of tasks being destroyed and created[15].

[15] In their model, tasks are not symmetric but vary by the needs for productivity. More precisely, the productivity demand increases with task index for labor but is constant for capital in their model.

With the production function, the final good producer's problem is to maximize profits,

$$Y - w_n N_n - w_r \int_m^1 n(i)di - (1+\tau_x) \int_0^1 \left(\phi_d x_d(i) + \phi_r x_r(i)\right) di, \tag{4}$$

where τ_x is an ad-valorem uniform tax rate on intermediate goods. Notice that my production function features constant returns to scale, leading to the zero profits property.

Proposition 1 below simplifies the production function (3). The intuition behind the derivation is that since all the marginal costs (ϕ_d, ϕ_r, w_r) for automatable tasks are constant across tasks, the parameter for the elasticity of substitution gets dropped. To derive it, we need the following technical assumption.

Assumption 2.1. *Every input for routine tasks, $x_d(i), x_r(i)$ and $n(i)$, has a positive measure for assigned tasks. That is, $x_d(i)$ has a positive measure on $i \in [0, m]$ and $x_r(j)$ and $n(j)$ have positive measures on $j \in (m, 1]$.*

With this assumption, the following proposition follows. For readability, define total amounts of every input as

$$X_d = \int_0^m x_d(i)di, \quad X_r = \int_m^1 x_r(i)di, \quad N_r = \int_m^1 n(i)di.$$

In my model, tasks are dichotomized into routine tasks and the single non-routine task. Acemoglu and Restrepo (2018) were able to find many interesting results based on their rich settings. However, their focus is to analyze the effects of automation on growth patterns and find conditions for balanced growth. In contrast, my main focus is to examine the sign of capital tax if not zero. Thus, for analytical tractability, I make an additional simplifying assumption on needs for labor productivity across different tasks.

Proposition 1. *Let Assumption 2.1 hold. Then, we can simplify the production function as*

$$Y = A\left(X_d + \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)^{1-\alpha} N_n^{\alpha} \tag{5}$$

Note that my production function contains that from Guerreiro et al. (2019) as a special case. To see this, let $q \longrightarrow \infty$ so that it becomes

$$Y = A\left(X_d + \beta X_r + (1-\beta)N_r\right)^{1-\alpha} N_n^{\alpha}.$$

If $\phi_d = \frac{1}{\beta}\phi_r$, then the production function becomes equivalent to that of Guerreiro et al. (2019).

2.3 Intermediate Goods Producers

Both types of capital are produced by perfectly-competitive intermediate goods producers in the global market and are used in automatable tasks $i \in [0,1]$. Following Guerreiro et al. (2019), I assume that tasks are symmetric—i.e., there is no difference among tasks in needs for productivity. Then, the price of capital type $k \in \{d, r\}$, $p_k(i)$, is same across tasks—that is, $p_k(i) = p_k \; \forall i$— and is equal to marginal costs of ϕ_k units of output. Since these producers are in the external market, a change in the income flows to these producers in response to capital taxation will not be considered. If the welfare function includes such income changes, then the optimal capital tax is expected to be even higher. This is because non-routine workers are expected to own more physical capital than routine workers, making greater room for redistribution. Therefore, my main result is a lower bound in this sense.

2.4 Government

The government determines tax rates and government spending amounts, subject to the budget constraint

$$G \leq \pi_r T\left(w_r l_r\right) + \pi_n T\left(w_n l_n\right) + \tau_x \int_0^1 \left[\phi_d x_d\left(i\right) + \phi_r x_r(i)\right] di \tag{6}$$

2.5 Equilibrium

An equilibrium is defined as the collection of a set of allocations $\{c_r, l_r, c_n, l_n, G, N_r, X_d, X_r, x_d(i), x_r(i), n(i), m\}$, prices $\{w_r, w_n, p_x\}$, and a tax system $\{T(.), \tau_x\}$ such that: (i) given prices and taxes, allocations solve the households' problem; (ii) given prices and taxes, allocations solve the firms' problem; (iii) the government budget constraint is satisfied; and (iv) markets clear.

The market-clearing conditions for routine and non-routine labor are given by

$$(1-m)n(i) = N_r = \pi_r l_r, \tag{7}$$

$$N_n = \pi_n l_n. \tag{8}$$

The market-clearing condition for displacing capital is

$$m x_d(i) = X_d, \tag{9}$$

while that for reinstating robots is

$$(1-m)x_r(i) = X_r, \tag{10}$$

Similarly, the market-clearing condition for the output market is

$$\pi_r c_r + \pi_n c_n + G \leq Y - \int_0^1 \left(\phi_d x_d(i) + \phi_r x_r(i)\right) di. \tag{11}$$

Also, note that we get an interior solution only if for a given level of total task-output, the total costs are the same between displacing capital and the composite of routine workers and reinstating capital. Equivalently,

$$(1+\tau_x)\phi_d = \frac{w_r n(i) + (1+\tau_x)\phi_r x_r(i)}{\left(\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta)n(i)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}} = \frac{w_r N_r + (1+\tau_x)\phi_r X_r}{\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}. \tag{12}$$

Then, the constant marginal cost of each input implies that at an interior solution, $x_d(i) = \left(\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta)n(i)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}$ for every i. This implies the following relation at an equilibrium with partial automation ($m \in (0,1)$):

$$x_d(i) = \frac{X_d}{m} = \frac{\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}{(1-m)}, \text{ for } i \in [0,m]. \tag{13}$$

Notice that using the fact that $mx_d(i) = X_d$ and $x_d(i) = \left(\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta)n(i)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}$

for every task i, one can also express m as

$$m = \frac{X_d}{X_d + \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta) N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}. \tag{14}$$

At an interior solution, by taking the first-order conditions of the profit function with respect to the inputs and plugging in the equilibrium conditions above, we get

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_x)\,\phi_d]^{\frac{1-\alpha}{\alpha}}}, \tag{15}$$

$$X_r = \left(\frac{1-\beta}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta}\right)^{\frac{q}{q-1}} \pi_r l_r, \tag{16}$$

$$w_r = (1-\beta)(1+\tau_x)\phi_d \left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta}\right)^{\frac{1}{q-1}}. \tag{17}$$

Note that when $0 < \left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} < \beta$, $w_r < 0$. In other words, as long as $\frac{\phi_r}{\phi_d} < \beta^{\frac{q}{q-1}}$, $w_r > 0$. Then, we can derive

$$m = 1 - \left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n} \left(\frac{\beta(1-\beta)}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta} + (1-\beta)\right)^{\frac{q}{q-1}}. \tag{18}$$

2.6 Main Result: Asymmetric Information at Both Household And Production Sides

In this section, the planner can observe neither household types nor capital types. What the planner observes are households' reported income levels. For the reasons highlighted in

Guesnerie (1998), I focus on a linear robot tax. Note that in contrast to Mirrlees (1971), the productivities of different agents are endogenous; they depend upon τ_x. This feature of my model induces trade-offs between redistribution and production efficiency by imposing a robot tax. To avoid degenerate cases, I focus on the case with partial automation, $0 < m < 1$. Furthermore, I assume that $w_n \geq w_r$ in an equilibrium.

In the Mirrlees' settings, the policymaker's problem is to choose allocations $\{c_j, l_j\}_{j=r,n}$, G and a uniform capital tax rate τ_x to maximize the following utilitarian social welfare:

$$\pi_r \omega_r \left[u\left(c_r, l_r\right) + v(G)\right] + \pi_n \omega_n \left[u\left(c_n, l_n\right) + v(G)\right], \tag{19}$$

where ω_j are social weights for household of type j. These weights are normalized in a way that $\pi_r\omega_r + \pi_n\omega_n = 1$. Following Guerreiro et al. (2019), I also focus on a case $\omega_r \geq 1$, so that the planner puts either equal or more weights on routine workers. The planner faces the following resource constraint[16]:

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \left(\frac{\tau_x + \alpha}{\alpha(1+\tau_x)}\right) + \frac{\pi_r w_r l_r}{(1+\tau_x)}. \tag{20}$$

The planner also faces households' incentive compatibility problems. For households of type j to choose bundles $\{c_j, l_j\}$, the bundle must yield at least as high as the utility level of any other arbitrary bundle choice $\{c, l\}$ that satisfies the budget constraint $c \leq w_j l - T(w_j l)$. This implies that $u(c_j, l_j) \geq u(c, l)$. In particular, routine workers must prefer their own-type bundle, $\{c_r, l_r\}$, to the bundle that they would obtain from masquerading the non-routine

[16] To get this, plug in equilibrium conditions of endogenous variables to the output market-clearing condition. The details of the derivation are in the proof for Lemma 1.

type by adjusting their hours of work, $\{c_n, w_n l_n / w_r\}$. Also, non-routine workers must prefer their own-type bundle, $\{c_n, l_n\}$, to the bundle that they would obtain from masquerading the routine type by adjusting their hours of work, $\{c_r, w_r l_r / w_n\}$. Note that an important assumption here is that mimicking the other type will not alter that worker's productivity[17]. We can write these as

$$u\left(c_n, l_n\right) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right), \tag{21}$$

$$u\left(c_r, l_r\right) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right). \tag{22}$$

Recall that the wages of two types of households are given by (15) and (17). These are necessary conditions for the household optimality and thus for an equilibrium.

In the following lemma, I show that the resource constraint (20) and the incentive compatibility constraints (21) and (22) are also sufficient conditions for an equilibrium. To establish that, the planner can set an income tax schedule such that for example, the government will appropriate the entire income if a household reports income levels other than $w_r l_r$ and $w_n l_n$.

Lemma 1. *The resource constraint (20) and the incentive compatibility constraints (21) and (22) are necessary and sufficient conditions for an equilibrium.*

Following Guerreiro et al. (2019), for analytical results, I first fix the level of τ_x and obtain optimal allocations. Second, I search for the optimal level of τ_x. Let $W(\tau_x)$ be the social welfare function (19) at the optimal allocation with a given level of τ_x. The social planner's problem at the second step is to find a maximizer $\tau_x^\star$, subject to the resource and incentive compatibility constraints. The optimal capital tax level has to satisfy the first-order

[17] See Scheuer and Werning (2016) for discussion on this point.

condition of the objective function. Following Guerreiro et al. (2019), I focus on cases where the IC constraint for non-routine workers binds, while that for routine workers slacks. This is a scenario that Stiglitz (1982) calls a *normal case*. This holds in my numerical exercises, as long as $w_n \geq w_r$.

Note that the expression for the net output in the right-hand side of (20) can be rewritten as:

$$\frac{\tau_x + \alpha}{\alpha\left(1+\tau_x\right)^{1/\alpha}} \frac{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\phi_d^{\frac{1-\alpha}{\alpha}}} \pi_n l_n + \pi_r l_r (1-\beta)\phi_d \left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta}\right)^{\frac{1}{q-1}} \tag{23}$$

Notice that the term $\dfrac{\tau_x + \alpha}{\alpha\left(1+\tau_x\right)^{1/\alpha}}$ is equal to one when $\tau_x = 0$ and strictly less than one when $\tau_x \neq 0$. This term measures the production inefficiency caused by the tax on capital.

As discussed in Guerreiro et al. (2019), in the absence of reinstating capital, the key force to make the capital tax positive is the reduction in the wage gap. This, in turn, relaxes the planner's information constraint from the non-routine workers' incentive compatibility, and therefore the planner can improve welfare by this informational gain, a similar intuition to Naito (1999) and Slavík and Yazici (2014). The problem with the uniform capital tax is that it may lower the wage of routine workers who have a higher marginal utility of income. This negative effect can dominate the informational gain, and thus it seems infeasible to determine the sign of the uniform tax with generality, at a glance.

Recall the main intuition behind the main result mentioned in the introduction. Due to the price differences between the two types of capital, a uniform capital tax indirectly imposes differential tax burdens on the two capital types. This additional key force makes the sign of the uniform tax to be always positive at an interior solution.

Proposition 2. *Let Assumption 2.1 hold. And suppose that the optimal allocation is such that the incentive compatibility constraint for non-routine workers binds, but that for routine workers does not bind. Then, at an interior solution (*$0 < m < 1$ *and* $l_r > 0$*), the optimal uniform linear capital tax is strictly positive, regardless of the value of* $q \in [0, 1]$*. The optimal capital tax rate generally satisfies*

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha}\frac{\pi_r l_r}{\pi_n}\left[\widetilde{\omega}_r\left(-u_l\left(c_r, l_r\right)\right) - (1-\beta)\phi_d\left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta}\right)^{\frac{1}{q-1}}\right] \tag{24}$$

where $\widetilde{\omega}_r = \omega_r/\mu$*, for* μ *being the Lagrange multiplier on the resource constraint.*

Now, we analyze what a practically implementable labor income tax would look like with this capital taxation. In fact, the forms of marginal rates of substitution between consumption and labor are the same as Guerreiro et al. (2019):

$$\frac{u_l\left(c_n, l_n\right)}{u_c\left(c_n, l_n\right)} = w_n\frac{\tau_x + \alpha}{\alpha\left(1+\tau_x\right)},$$

and

$$\frac{u_l\left(c_r, l_r\right) l_r}{u_c\left(c_r, l_r\right)} = \frac{\omega_r - \eta_n\frac{u_c\left(c_r, w_r l_r/w_n\right)}{u_c\left(c_r, l_r\right)}}{\omega_r - \eta_n\frac{u_l\left(c_r, w_r l_r/w_n\right)\frac{1}{w_n}}{u_l\left(c_r, l_r\right)\frac{1}{w_r}}}\frac{w_r l_r}{1+\tau_x},$$

where $\eta_n\pi_r$ is the multiplier for the IC constraint of non-routine workers. The properties of labor income tax at optimum are same with Guerreiro et al. (2019) as well: (i) non-routine workers are subsidized at the margin at partial automation, while (ii) routine workers are taxed at the margin.

2.7 First-Best: Perfect Discrimination of Labor Types

To see the ideal outcome and have a reference point for the main result, I define two different scenarios in this and the next section: the first-best and second-best outcomes. Following Guerreiro et al. (2019), the first-best allocation maximizes the weighted average of utilities, subject only to resource constraints without incentive compatibility constraints of households.

I denote such (positive) weights by ω_r for routine households and by ω_n for non-routine households. These weights are normalized in a way that $\pi_r\omega_r + \pi_n\omega_n = 1$. The planner's problem is to choose $\{c_r, l_r, c_n, l_n, G, m, \{x_d(i), x_r(i), n(i)\}\}$ to maximize utilitarian social welfare of

$$\pi_r\omega_r \left[u\left(c_r, l_r\right) + v(G)\right] + \pi_n\omega_n \left[u\left(c_n, l_n\right) + v(G)\right]. \tag{25}$$

Note that the first-best allocation implies production efficiency since robots' markets are perfectly competitive. In the first-best scenario, the government can perfectly distinguish and discriminate between the two labor types. Then, the government will whip more productive non-routine workers to work more while increase the transfer amount from non-routine households to routine households over time, so that the consumption levels are equalized between the two households. Non-routine workers are more productive in a sense that unlike routine workers, they are not replaceable. This is consistent with the first-order conditions of the planner's problem with respect to consumption amounts c_j. This results in a higher utility level for routine households. Since this creates incentives for non-routine households to slack and imitate routine households, this outcome is an ideal but unrealistic one.

I define the first-best allocation in the economy as the solution to the aforementioned utilitarian social welfare function, absent informational constraints. The absence implies that the planner can perfectly discriminate among two types of agents and enforce any allocation.

2.8 Second-best: Perfect Differentiation of Capital Types

I now analyze a situation where the government cannot perfectly discriminate between two types of labor but can perfectly distinguish between two types of capital. Denote by τ_d as the capital tax on displacing capital and by τ_r as that on reinstating capital. Both taxes are bounded from below but not above: $-1 \leq \tau_d < \infty$ and $-1 \leq \tau_r < \infty$.

At a partial automation level $0 < m < 1$, the differences from the main setting are

$$X_r = \left(\frac{1-\beta}{\left(\frac{(1+\tau_r)\phi_r}{\beta(1+\tau_d)\phi_d}\right)^{q-1} - \beta} \right)^{\frac{q}{q-1}} \pi_r l_r \tag{26}$$

$$w_r = (1-\beta)(1+\tau_d)\phi_d \left((1-\beta) \left(\frac{\beta}{\left(\frac{(1+\tau_r)\phi_r}{\beta(1+\tau_d)\phi_d}\right)^{q-1} - \beta} + 1 \right) \right)^{\frac{1}{q-1}}, \tag{27}$$

and

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[(1+\tau_d)\,\phi_d\right]^{\frac{1-\alpha}{\alpha}}}. \tag{28}$$

Also, the parameter of automation m becomes

$$m = 1 - \left(\frac{(1+\tau_d)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n}. \tag{29}$$

With this, in general, I obtain a result where the optimal capital tax rate for displacing capital is positive and that for reinstating capital is negative. The result is quite intuitive since the tax on displacing capital is similar to the one in Guerreiro et al. (2019) and can be used as an instrument to reduce the wage gap for the aforementioned informational gains, while that on reinstating capital can be used to boost the wage of routine labor at the expense of second-order production loss. The proof can be found in the Appendix.

Proposition 3. *Suppose that the optimal allocation is such that the incentive compatibility constraint for non-routine workers binds, but that for routine workers does not bind. Then, given partial automation ($m < 1$ and $l_r > 0$), and given Assumption 2.1, the optimal tax for displacing capital is always strictly positive ($\tau_d^\star > 0$) if that for reinstating capital is zero ($\tau_r = 0$), while that for reinstating capital is always strictly negative ($\tau_r^\star < 0$) if that for displacing capital is zero ($\tau_d = 0$), regardless of the value of $q \in [0, 1]$.*

Note that the planner tends to be able to just use one of the two instruments for optimality. Yet, the following corollary demonstrates some exceptions with the special case of $q \longrightarrow 0$.

Corollary 1. *Suppose that the optimal allocation is such that the incentive compatibility constraint for non-routine workers binds, but that for routine workers does not bind. Suppose also that $q \longrightarrow 0$. Then, given partial automation ($m < 1$ and $l_r > 0$), and given Assumption 2.1, the optimal capital tax rate on reinstating capital is full subsidy ($\tau_r = -1$) and the optimal capital tax rate on displacing capital is strictly positive, unless one of the two instruments can solely achieve the first-best allocations.*

The reason for 100% subsidy for reinstating capital comes from the Leontief functional form of the joint production of routine labor and reinstating capital. Note that the wage rate of routine workers in this case becomes

$$w_r = (1 + \tau_d)\phi_d - (1 + \tau_r)\phi_r.$$

We can see that the wage goes up by a linear rate of ϕ_r. The increase tends to dominate the second-order production inefficiency by the subsidy.

3 Numerical Analysis

For numerical exercises, I mostly follow the settings used in Guerreiro et al. (2019). I assume the following utility functional form from Ales, Kurnaz, and Sleet (2015) and Heathcote, Storesletten, and Violante (2017):

$$u(c_j, l_j) + v(G) = \log(c_j) - \zeta \frac{l_j^{1+\nu}}{1+\nu} + \chi \log(G).$$

This utility functional form features two desirable properties. One is that this form is consistent with balanced growth. The other is that the preferences are consistent with the empirical evidence from Chetty (2006).

ζ is set at 10.63. ν is set at 4/3, so that the Frisch elasticity becomes 0.75 that is consistent with the estimates discussed in Chetty et al. (2011). As discussed in Heathcote et al. (2017), χ is chosen to be 0.233 that makes the optimal ratio of government spending to output 18.9 percent, which is the weight observed in the U.S.

Furthermore, normalize A to one and choose $\alpha = 0.53$ and $\pi_r = 0.55$, which are estimates from Chen (2016).

To capture the technological progress over time, I consider a sequence of static equilibria in which the cost of capital geometrically falls over time, $\phi_k = \phi_{0,k} e^{-g_\phi t}$ for $k \in \{d, r\}$. To allow my model to match the decline in task content estimated by Acemoglu and Restrepo (2019) for the year 2000-2008, I set $g_\phi = 0.0083$. The period avoids the financial crisis but captures the time period when the automation movement took off in the U.S.

As for ϕ_k, Guerreiro et al. (2019) choose ϕ_d to be 0.4226, the lowest value that is consistent with no automation in their status-quo equilibria[18]. Note that there is no ϕ_r in their model. In my model, what matters the most is the difference between the two capital costs, rather than individual cost magnitudes. In my numerical exercises, unless noted, I choose $\phi_d = 0.45$ and ϕ_r to be 55% of ϕ_d. The reason to choose this difference is that it is consistent with ad-hoc evidence of cost differences between industrial robots and cobots[19]. Meanwhile, I will demonstrate the optimal capital tax rate of the main setting with respect to the whole space of the cost differences—that is, a matrix of different (ϕ_d, ϕ_r) composites.

3.1 First-best: Perfect Discrimination of Labor Types

Figure 1 shows the first-best outcome with the aforementioned parameter selections. Note that while the full automation takes place only asymptotically, a large portion of routine tasks will be displaced in the early periods. The decrease in the cost of displacing capital over time results in the increasing wage gap between the two household types. However, since in the first-best scenario, the government can perfectly distinguish and discriminate between the two labor types, the government will whip more productive non-routine workers to work

[18] The status-quo equilibria in their study are equilibria where there is no capital tax, and the planner uses the income tax schedule estimated in Heathcote et al. (2017)

[19] See, for example, `https://blog.robotiq.com/what-is-the-price-of-collaborative-robots`

more while increasing the transfer amount from non-routine households to routine households over time, so that the consumption levels are equalized between the two households. This is consistent with the first-order conditions of the planner's problem with respect to consumption amounts c_j. This results in a higher utility level for routine households. Since this creates incentives for non-routine households to slack and imitate routine households, this outcome is an ideal but unrealistic one.

3.2 Second-best: Perfect Identification of Capital Types

Figure 2 and 3a, 3b, and 3c demonstrate the numerical results of the second-best outcome with different functional choices of the joint production function involving routine workers and reinstating capital. Figure 2 shows the Leontief functional form version, which is the case the most deviant from Guerreiro et al. (2019). As the corollary above expected, the optimal tax rate of reinstating capital is -1 (full subsidy) while that of displacing capital is positive. The taxes and allocations are set at the levels so that the incentive constraint for non-routine households binds, resulting in the small utility difference between the two households.

Figure 3a, 3b, and 3c involve a Cobb-Douglas functional form assumption with different values of share between the two inputs: 0.5, 0.2, and 0.7. Notice that the magnitude of the optimal capital tax rate becomes larger when the share parameter is lower[20]. This is because the smaller share, the closer the economy becomes to the one in Guerreiro et al. (2019) where there is no reinstating capital. Furthermore, note that these cases are the least deviant from

[20] Recall that the planner can just use either one of the capital tax instruments in regular cases. While the optimal capital tax on displacing capital is higher for $\beta = 0.7$ than $\beta = 0.5$ case, the subsidy rate on reinstating capital is higher for $\beta = 0.5$ case. If one adds both instruments up, then one can see the magnitude of the capital tax is larger for $\beta = 0.5$ case in total.

Guerreiro et al. (2019). This is because Cobb-Douglas is the least complementary case. Thus, we see much lower subsidy rates on reinstating capital than the Leontief case.

3.3 Third-best: Main Settings

Figure 4 shows the third-best outcome that is the main setting outcome in this paper, with a Leontief functional form assumption for the joint production of routine labor and reinstating capital. I choose this functional form since this is the most deviant case from Guerreiro et al. (2019). Note that the optimal uniform capital tax rate has a decreasing trend over time. This is because the gains from the tax become smaller and the costs from distorting production become larger as the cost of capital decreases over time. More importantly, the cost difference between the two capital types becomes smaller since the technological progress rate is assumed to be the same between the two. The more expensive displacing capital benefits disproportionately more from technological progress. The reduced gap leads to a reduced rate of an increase in the wage rate of routine workers by the capital tax. We can see this from their equilibrium wage rate in the Leontief assumption:

$$w_r = (1 + \tau_x)(\phi_d - \phi_r).$$

Then, over time, the redistribution force with the capital tax becomes smaller, leading to near zero optimal capital tax rates asymptotically.

This motivates us to look into the optimal capital tax rate in different coordinates of (ϕ_d, ϕ_r). Figure 5 demonstrates this. As you can see in the figure, there is a hump in the optimal capital tax rate. When the gap between the two costs is large, the routine worker wage is so high that there does not need to be much redistribution to make. On the other hand, when there is only a small gap, then an increase in the capital tax rate will only marginally increase the routine labor wage. Therefore, there is a sweet spot in between for redistribution. Note that this result partially addresses the issue of magnitude. If the magnitude of the optimal capital tax rate is negligible compared to a case of Guerreiro et al. (2019) where the planner can tax only displacing capital, then the theoretical finding on the sign of the optimal uniform capital tax may have little importance. Figure 5 partially addresses this concern since the magnitude depends on the magnitude of the cost difference.

Furthermore, Figure 6a, b, c, and d provide a more direct answer. I look at the both Cobb-Douglas and Leontief cases that are the closest and most deviant from Guerreiro et al. (2019), respectively. I compare the welfare and routine households' consumption levels among all the scenarios of the first-best ("FB" in Figure 6a and 6b), the second-best ("SB"), the main third-best ("TB"), and the outcome with $\tau_x = 0$ ("Zero"). Recall that the second-best allocations allow the planner to differentially tax displacing and reinstating capital, an analogous case to Guerreiro et al. (2019). Moreover, since the optimal capital tax becomes close to zero after 2050 in the Leontief case, I compare the welfare levels between 2000 and 2050.

Figure 6a uses Cobb-Douglas with $\beta = 0.2$ and $\phi_{r,0} = 0.55\phi_{d,0}$ same as with one of the previous parameter value settings. As you can see, in this case, the difference between the second-best and third-best is so small that it is almost negligible, compared to the difference from the first-best and zero-capital-tax cases. This is because Cobb-Douglas is the least gross-complement case, and thus the differential tax burden effect from the price difference has relatively strong effects. Figure 6b is the same as Figure 6a except for the y-axis to be consumption amounts of routine households. In the early periods, the uniform capital tax can dramatically increase the amounts up to 17% compared to the zero capital tax case.

As for the Leontief case, Figure 6c shows that the welfare improvement of the third-best outcome is negligible, compared to that of the second-best especially in the beginning. Figure 6d demonstrate that the consumption levels of routine households also increase only up to 0.03% in the early periods.

These results imply that while the magnitude of the welfare improvement by the optimal uniform capital taxation depends upon the degree of complementarity between routine labor and reinstating capital, the magnitude can be comparable to the differential capital taxation case.

4 Extension: Dynamic Model

While I explore the extension to the dynamic counterpart of the static model in the Appendix, the addition of another capital type makes it extremely difficult to obtain both analytical and numerical results in the uniform taxation case. Nevertheless, given that Guerreiro et al. (2019) find even larger positive tax rates in the dynamic counterpart, that is likely to be the case in my setting as well.

5 Conclusion & Limitations

This paper finds that despite the asymmetric information problem, the optimal uniform capital tax rate over different types of capital is strictly positive, as long as the solution is interior. Here, I will acknowledge the limitations of this study. First, my model assumes away occupational transfer between routine and non-routine sectors for analytical tractability. While Gomes et al. (2017) find that allowing for both occupational choice and labor supply adjustments may result in production inefficiency at optimum, there might be some important insights in the extension to dynamic models with uncertainty and evolving abilities and skills.

Another limitation is that I assume perfectly competitive external global markets for intermediate goods producers. One possible extension is to relax this assumption and to incorporate R&D process for intermediate goods producers along the lines of Aghion et al. (2013). In this case, a uniform capital tax that results in the differential tax burden on different capital type would create a resource wedge in R&D between displacing and reinstating capital, leading to differential technological progress over time. This could further drive up the optimal uniform capital tax rate in the long run.

References

Acemoglu, D., Manera, A., and Restrepo, P. (2020). “Does the us tax code favor automation?” Technical report, National Bureau of Economic Research.

Acemoglu, D. and Restrepo, P. (2017). “Robots and jobs: Evidence from us labor markets”. *NBER working paper*, (w23285).

— (2018). “The race between man and machine: Implications of technology for growth, factor shares, and employment”. *American Economic Review*, 108(6):1488–1542.

— (2019). “Automation and new tasks: How technology displaces and reinstates labor”. Technical report, National Bureau of Economic Research.

Aghion, P., Akcigit, U., and Fernández-Villaverde, J. (2013). “Optimal capital versus labor taxation with innovation-led growth”. Technical report, National Bureau of Economic Research.

Atkinson, A. B. and Stiglitz, J. E. (1976). “The design of tax structure: direct versus indirect taxation”. *Journal of public Economics*, 6(1-2):55–75.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). “The skill content of recent technological change: An empirical exploration”. *The Quarterly journal of economics*, 118(4):1279–1333.

Chen, Y. G. (2016). “The rise of the machines and the u.s. labor market”. Technical report, mimeo, University of Maryland.

Chetty, R. (2006). “A new method of estimating risk aversion”. *American Economic Review*, 96(5):1821–1834.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). "Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins". *American Economic Review*, 101(3):471–75.

Costinot, A. and Werning, I. (2018). "Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation". Technical report, National Bureau of Economic Research.

Deaton, A. (1979). "Optimally uniform commodity taxes". *Economics Letters*, 2(4):357–361.

Diamond, P. A. and Mirrlees, J. A. (1971a). "Optimal taxation and public production i: Production efficiency". *The American Economic Review*, 61(1):8–27.

— (1971b). "Optimal taxation and public production ii: Tax rules". *The American Economic Review*, 61(3):261–278.

Farhi, E. and Werning, I. (2013). "Insurance and taxation over the life cycle". *Review of Economic Studies*, 80(2):596–635.

Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). "Optimal indirect and capital taxation". *The Review of Economic Studies*, 70(3):569–587.

Golosov, M., Troshkin, M., and Tsyvinski, A. (2011). "Optimal dynamic taxes". Technical report, National Bureau of Economic Research.

— (2016). "Redistribution and social insurance". *American Economic Review*, 106(2):359–86.

Golosov, M., Tsyvinski, A., Werning, I., Diamond, P., and Judd, K. L. (2006). "New dynamic public finance: A user's guide [with comments and discussion]". *NBER macroeconomics annual*, 21:317–387.

Gomes, R., Lozachmeur, J.-M., and Pavan, A. (2017). "Differential taxation and occupational choice". *The Review of Economic Studies*, 85(1):511–557.

Guerreiro, J., Rebelo, S., and Teles, P. (2019). "Should robots be taxed?" Technical report, National Bureau of Economic Research.

Guesnerie, R. (1998). *A contribution to the pure theory of taxation*. 25. Cambridge University Press.

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

Kocherlakota, N. R. (2005). "Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation". *Econometrica*, 73(5):1587–1621.

Mirrlees, J. A. (1971). "An exploration in the theory of optimum income taxation". *The review of economic studies*, 38(2):175–208.

Naito, H. (1999). "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency". *Journal of Public Economics*, 71(2):165–188.

Rothschild, C. and Scheuer, F. (2013). "Redistributive taxation in the roy model". *The Quarterly Journal of Economics*, 128(2):623–668.

Sachs, J. D., Benzell, S. G., and LaGarda, G. (2015). "Robots: Curse or blessing? a basic framework". Technical report, National Bureau of Economic Research.

Sachs, J. D. and Kotlikoff, L. J. (2012). "Smart machines and long-term misery". Technical report, National Bureau of Economic Research.

Saez, E. (2001). "Using elasticities to derive optimal income tax rates". *The review of economic studies*, 68(1):205–229.

— (2002). "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes". *Journal of Public Economics*, 83(2):217–230.

— (2004). "Direct or indirect tax instruments for redistribution: short-run versus long-run". *Journal of public economics*, 88(3-4):503–518.

Scheuer, F. (2014). "Entrepreneurial taxation with endogenous entry". *American Economic Journal: Economic Policy*, 6(2):126–63.

Scheuer, F. and Werning, I. (2016). "Mirrlees meets diamond-mirrlees". Technical report, National Bureau of Economic Research.

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

Slemrod, J. and Kopczuk, W. (2002). "The optimal elasticity of taxable income". *Journal of Public Economics*, 84(1):91–112.

Slemrod, J. and Yitzhaki, S. (2002). "Tax avoidance, evasion, and administration". In "Handbook of public economics", volume 3, pages 1423–1470. Elsevier.

Stantcheva, S. (2017). “Optimal taxation and human capital policies over the life cycle”. *Journal of Political Economy*, 125(6):1931–1990.

Stiglitz, J. E. (1982). “Self-selection and pareto efficient taxation”. *Journal of public economics*, 17(2):213–240.

Thuemmel, U. (2018). “Optimal taxation of robots”.

Appendices

A Proof of Proposition 1

By Assumption 2.1, we can differentiate the profit function (4) with respect to $x_d(i), x_r(i)$, and $n(i)$. Then, the optimal choices of N_n, $x_d(i)$ for $i \in [0,m]$, $n_i, x_r(i)$ for $i \in (m,1]$ require the following first-order conditions:

$$w_n = \frac{\alpha Y}{N_n}, \tag{30}$$

$$\begin{aligned}(1+\tau_x)\,\phi_d =&(1-\alpha)Y\left(\int_0^m x_d\left(s\right)^{\frac{\rho-1}{\rho}}ds + \left(\int_m^1 \left[x_r(s)^{\frac{q-1}{q}} + n(s)^{\frac{q-1}{q}}\right]^{\frac{\rho-1}{\rho}}ds\right)^{\frac{q}{q-1}}\right)^{-1} \\ & x_d(i)^{-\frac{1}{\rho}},\ \text{for } i \in [0,m],\end{aligned} \tag{31}$$

$$\begin{aligned}(1+\tau_x)\phi_r =&(1-\alpha)Y\left(\int_0^m x_d\left(s\right)^{\frac{\rho-1}{\rho}}ds + \left(\int_m^1 \left[\beta x_r(s)^{\frac{q-1}{q}} + (1-\beta)n(s)^{\frac{q-1}{q}}\right]^{\frac{\rho-1}{\rho}}ds\right)^{\frac{q}{q-1}}\right)^{-1} \\ & \left(\int_m^1 \left[\beta x_r(z)^{\frac{q-1}{q}} + (1-\beta)n(z)^{\frac{q-1}{q}}\right]^{\frac{\rho-1}{\rho}}dz\right)^{\frac{1}{q-1}} \beta x_r(i)^{-\frac{1}{q}},\ \text{for } i \in (m,1].\end{aligned} \tag{32}$$

$$\begin{aligned}w_r =&(1-\alpha)Y\left(\int_0^m x_d\left(s\right)^{\frac{\rho-1}{\rho}}ds + \left(\int_m^1 \left[\beta x_r(s)^{\frac{q-1}{q}} + (1-\beta)n(s)^{\frac{q-1}{q}}\right]^{\frac{\rho-1}{\rho}}ds\right)^{\frac{q}{q-1}}\right)^{-1} \\ & \left(\int_m^1 \left[\beta x_r(z)^{\frac{q-1}{q}} + (1-\beta)n(z)^{\frac{q-1}{q}}\right]^{\frac{\rho-1}{\rho}}dz\right)^{\frac{1}{q-1}} (1-\beta)n(i)^{-\frac{1}{q}},\ \text{for } i \in (m,1].\end{aligned} \tag{33}$$

From (31), it follows that in the m automated tasks, the optimal level of displacing robots is equal across tasks. From (32) and (33), it also follows that in the $1-m$ unautomated tasks, the optimal level of routine labor and cobots are also the same across tasks. Hitherto, whenever possible, we drop the task index i for simplicity. Note that the final producer will choose $m=1$ whenever the effective price of displacing capital is cheaper than that of the combination of reinstating capital and routine labor:

$$(1+\tau_x)\phi_d > \frac{w_r n + (1+\tau_x)\phi_r x_r}{\left(\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta) n(i)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}. \tag{34}$$

They will choose $m=0$ instead if the inequality above is the opposite[21]. We will observe $0 < m < 1$ if the inequality is equality. We are interested in this scenario of partial automation since otherwise, as we will show, the capital tax is always zero.

Then, as partial automation, we get $x_d(i) = \left[\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta) n(i)^{\frac{q-1}{q}}\right]$. Thus, at partial automation, we can drop the substitution parameters for tasks, ρ. Then, we get

$$Y = A\left[\int_0^m x_d(i)di + \left(\int_m^1 \left[\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta) n(i)^{\frac{q-1}{q}}\right] di\right)^{\frac{q}{q-1}}\right]^{(1-\alpha)} N_n^{\alpha},$$
$$\alpha \in (0,1), \rho \in [0,\infty). \tag{35}$$

Also, since the amount of $x_r(i)$ is the same across tasks, and since that of $n(i)$ is the same

[21]Note that the denominator in (34) is a certain output level $y(i)$

across tasks, we can write[22]

$$Y = A\left[\int_0^m x_d(i)di + \int_m^1 \left[\beta x_r(i)^{\frac{q-1}{q}} + (1-\beta)n(i)^{\frac{q-1}{q}}\right]^{\frac{q}{q-1}} di\right]^{(1-\alpha)} N_n^{\alpha},$$

$$\alpha \in (0,1)\,, \rho \in [0,\infty)\,. \tag{36}$$

What's more, at partial automation, the final good producer is indifferent in the level of automation m. In this case, we get

$$mx_d(i) = X_d, \text{ for } i \in [0,m], (1-m)x_r(i) = X_r, \text{ and } (1-m)n(i) = N_r,$$

$$, \text{ for } i \in (m,1], \tag{37}$$

where X_d is the total amount of displacing robots, X_r is that of reinstating robots, and N_r is the total employment of routine labor, as defined in the main text[23]. Note that by definition, we have $m = \frac{X_d}{X_d + \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}$ if $m > 0$. Then, we can rewrite the production function as

$$Y = A\left(X_d + \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)^{1-\alpha} N_n^{\alpha}$$

.

[22]It might be easier to see this with finite summation. If the sum is finite, then the integral is $\left(\sum_{i=m}^1 x_r(i)^{\frac{q-1}{q}} + \sum_{i=m}^1 n(i)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}$. Since $x_r(i) = x_r(i')$ and $n(i) = n(i')$ for $i \neq i'$, the sum equals to $\left([(1-m)x_r(i)]^{\frac{q-1}{q}} + [(1-m)n(i)]^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}} = (1-m)[x_r(i)^{\frac{q-1}{q}} + n(i)^{\frac{q-1}{q}}]^{\frac{q}{q-1}}$

[23]Note that $(1-m)x_r(i) = X_r$, and $(1-m)n(i) = N_r$ implies $(1-m)[x_r^{\frac{q-1}{q}} + n^{\frac{q-1}{q}}]^{\frac{q}{q-1}} = \left([(1-m)x_r]^{\frac{q-1}{q}} + [(1-m)n]^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}} = \left(X_r^{\frac{q-1}{q}} + N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}$.

B Proof of Lemma 1

A $q \longrightarrow 1$

Proof. The final good producer optimizes inputs and get

$$x_r(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)} \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{1}{\beta-1}}, & i \in (m,1] \\ 0, & \text{otherwise} \end{cases} \tag{38}$$

$$n(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)}, & i \in (m,1] \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

$$x_d(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)} \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}}, & i \in [0,m] \\ 0, & \text{otherwise} \end{cases} \tag{40}$$

$$m = \max\left\{1 - \left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}} \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \frac{\pi_r l_r}{\pi_n l_n}, 0\right\} \tag{41}$$

$$Y = A\left[\int_0^m x_d(i)^{\frac{\rho-1}{\rho}} di + \int_m^1 \left[x_r(i)^\beta n(i)^{1-\beta}\right]^{\frac{\rho-1}{\rho}} di\right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^\alpha, \quad \alpha \in (0,1), \rho \in [0,\infty). \tag{42}$$

At an interior solution, while w_n is the same in the main text,

$$w_r = (1-\beta)(1+\tau_x)\phi_d \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}}. \tag{43}$$

Combining these results, after some algebra, we can rewrite the resource constraint (15) as:

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \left(\frac{\tau_x + \alpha}{\alpha(1+\tau_x)} \right) + \frac{\pi_r w_r l_r}{1+\tau_x} \tag{44}$$

This captures all the equilibrium conditions of the production side.

Given that households cannot change their types, the optimality of hosuehold choice gives us

$$u\left(c_j, l_j\right) \geq u(c, l), \quad \forall (c, l) : c \leq w_j l - T\left(w_j l\right). \tag{45}$$

Then, the following IC constraints are necessary conditions for the optimality:

$$u\left(c_n, l_n\right) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right)$$
$$u\left(c_r, l_r\right) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right)$$

Notice that these two IC constraints are also sufficient conditions for the household side at an equilibrium since the government can freely adjust the tax schedule $T(.)$ so that for all $Y \notin \{Y_n, Y_r\}$, both household types receive worse allocations than their respective allocation. The government can achieve this by setting the tax schedule to be:

$$T(y) = y - \max \left\{ c \middle| u\left(c_i, l_i\right) \geq u\left(c, \frac{y}{w_i}\right), \text{ for } i = r, n \right\} \tag{46}$$

■

B $q \longrightarrow 0$

Proof. The final good producer optimizes inputs and get

$$x_r(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)}, & i \in (m, 1] \\ 0, & \text{otherwise} \end{cases} \tag{47}$$

$$n(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)}, & i \in (m, 1] \\ 0, & \text{otherwise} \end{cases} \tag{48}$$

$$x_d(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)}, & i \in [0, m] \\ 0, & \text{otherwise} \end{cases} \tag{49}$$

$$m = \max \left\{ 1 - \left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n}, 0 \right\} \tag{50}$$

$$Y = A \left[\int_0^m x_d(i)^{\frac{\rho-1}{\rho}} di + \int_m^1 \min[x_r(i), n(i)]^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^{\alpha}, \quad \alpha \in (0,1), \rho \in [0, \infty). \tag{51}$$

$$w_r = (1+\tau_x)(\phi_d - \phi_r) \tag{52}$$

Combining these results, again, we can rewrite the resource constraint (15) to

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \frac{\pi_r w_r l_r}{1+\tau_x} \tag{53}$$

This captures all the equilibrium conditions of the production side.

The rest of the proof is the same with $q \longrightarrow 1$. ∎

C $0 < q < 1$

Proof. The final good producer optimizes inputs and get

$$x_r(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)} \left(\frac{1-\beta}{\left(\frac{\phi_r}{\beta \phi_d}\right)^{q-1} - \beta} \right)^{\frac{q}{q-1}}, & i \in (m, 1] \\ 0, & \text{otherwise} \end{cases} \tag{54}$$

$$n(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)}, & i \in (m, 1] \\ 0, & \text{otherwise} \end{cases} \tag{55}$$

$$x_d(i) = \begin{cases} \frac{\pi_r l_r}{(1-m)} \left(\beta \left(\frac{1-\beta}{\left(\frac{\phi_r}{\beta \phi_d}\right)^{q-1} - \beta} \right) + (1-\beta) \right), & i \in [0, m] \\ 0, & \text{otherwise} \end{cases} \tag{56}$$

$$m = \max \left\{ 1 - \left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n}, 0 \right\} \tag{57}$$

$$Y = A \left[\int_0^m x_d(i)^{\frac{\rho-1}{\rho}} di + \int_m^1 \min[x_r(i), n(i)]^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^\alpha, \quad \alpha \in (0,1), \rho \in [0, \infty). \tag{58}$$

$$w_r = (1-\alpha)(1-m) \frac{Y}{\pi_r l_r}, \tag{59}$$

Combining these results, after a lot of tedious algebra, we can rewrite the resource constraint (15) to

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha\,(1+\tau_x)} + \frac{\pi_r w_r l_r}{1+\tau_x} \tag{60}$$

This captures all the equilibrium conditions of the production side.

The rest of the proof is the same with $q \longrightarrow 1$. This concludes the whole proof. ∎

C Proof of Proposition 2

Proof. Define $W(\tau_x) = \max \pi_r \omega_r u\,(c_r, l_r) + \pi_n \omega_n u\,(c_n, l_n) + v(G)$. Then, the social planner's optimization problem is

$$\underset{\tau_x}{\text{maximize}}\; W(\tau_x)$$

subject to

$$[\eta_r \pi_r] \quad u\,(c_r, l_r) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right),$$

$$[\eta_n \pi_n] \quad u\,(c_n, l_n) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right),$$

$$[\mu] \quad \pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \left(\frac{\tau_x + \alpha}{\alpha(1+\tau_x)}\right) + \frac{\pi_r w_r l_r}{(1+\tau_x)}.$$

Letting $\eta_r = 0$, we get

$$W'(\tau_x) = -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r} l_r\right)\frac{d\log(w_r/w_n)}{d\log(1+\tau_x)}\frac{1}{1+\tau_x}\frac{w_r l_r}{w_n}$$
$$+\mu\left[\pi_n w_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)^2}\left(\frac{d\log w_n}{d\log(1+\tau_x)} + \frac{1-\alpha}{(\tau_x+\alpha)}\right) + \frac{\pi_r w_r l_r}{(1+\tau_x)^2}\left(\frac{d\log w_r}{d\log(1+\tau_x)} - 1\right)\right]. \tag{61}$$

I separately prove the limit cases ($q \longrightarrow 0$ and $q \longleftrightarrow 1$) and interior case ($0 < q < 1$).

A $\quad q \longrightarrow 0$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1+\tau_x)(\phi_d - \phi_r) \Rightarrow \frac{d\log w_r}{d\log(1+\tau_x)} = 1$$

$$w_n = \alpha\frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_x)\phi_d]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_n}{d\log(1+\tau_x)} = -\frac{1-\alpha}{\alpha}$$

$$\frac{w_r}{w_n} = \frac{(1+\tau_x)^{\frac{1}{\alpha}}\left[\phi_d^{\frac{1}{\alpha}} - \phi_r\phi_d^{\frac{1-\alpha}{\alpha}}\right]}{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_r/w_n}{d\log(1+\tau_x)} = \frac{1}{\alpha}$$

Plugging these back into the envelope condition (61), we get:

$$W'(\tau_x) = -\eta_n\pi_n u_l\left(c_r, \frac{w_r}{w_r}l_r\right)\frac{1}{\alpha(1+\tau_x)}\frac{w_r l_r}{w_n} + \mu\pi_n w_n l_n\frac{\tau_x+\alpha}{\alpha(1+\tau_x)^2}\left[-\frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\tau_x+\alpha}\right]$$
$$= \frac{1}{\alpha(1+\tau_x)}\left[-\eta_n\pi_n u_l\left(c_r, \frac{w_r}{w_n}l_r\right)\frac{w_r l_r}{w_n} - \mu\pi_n w_n l_n\frac{\tau_x}{1+\tau_x}\frac{1-\alpha}{\alpha}\right]$$

Since $\mu > 0$, if $\tau_x \leq 0$, if $w_r \geq 0$—that is, $\phi_r \leq \phi_d$—, then we are guaranteed to have:

$$W'(\tau_x) > 0.$$

Thus, the social planner can always improve welfare by marginally increasing τ_x. Furthermore, the optimal level of τ_x implies $W'(\tau_x) = 0$, so we get:

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\eta_n \left(-u_l \left(c_r, \frac{w_r}{w_n} l_r\right) \frac{w_r l_r}{w_n}\right)}{\mu w_n l_n} \tag{62}$$

The first order condition with respect to l_r yields:

$$\begin{aligned} -\frac{\eta_n}{\mu} u_l \left(c_r, \frac{w_r}{w_r} l_r\right) \frac{w_r l_r}{w_n} &= -\left(\frac{\widetilde{\omega}_r \pi_r u_l \left(c_r, l_r\right) l_r + \frac{\pi_r w_r l_r}{1+\tau_x}}{\pi_n}\right) \\ &= \frac{\pi_r l_r}{\pi_n} (\phi_d - \phi_r) \left[\frac{\widetilde{\omega}_r \left(-u_l \left(c_r, l_r\right)\right)}{\phi_d - \phi_r} - 1\right] \end{aligned} \tag{63}$$

Thus, we get

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\pi_r l_r}{\pi_n w_n l_n} (\phi_d - \phi_r) \left[\frac{\widetilde{\omega}_r \left(-u_l \left(c_r, l_r\right)\right)}{\phi_d - \phi_r} - 1\right] \tag{64}$$

B $q \longrightarrow 1$ Case

Here, we analyze the Cobb-Douglas case. We have

$$X_r = \left(\frac{\phi_r}{\beta \phi_d}\right)^{\frac{1}{\beta-1}} \pi_r l_r. \tag{65}$$

Also,

$$
\begin{aligned}
m &= 1-\left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}} \frac{(\pi_r l_r)^{(1-\beta)}}{\pi_n l_n} X_r^{\beta} \\
&= 1-\left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}} \frac{(\pi_r l_r)^{(1-\beta)}}{\pi_n l_n} \left(\left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{1}{\beta-1}} \pi_r l_r\right)^{\beta} \\
&= 1-\left(\frac{(1+\tau_x)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n} \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}}. \qquad (66)
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
w_r &= (1-\beta)(1+\tau_x)\phi_d X_r^{\beta}(\pi_r l_r)^{-\beta} \\
&= (1-\beta)(1+\tau_x)\phi_d \left(\left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{1}{\beta-1}} \pi_r l_r\right)^{\beta} (\pi_r l_r)^{-\beta} \\
&= (1-\beta)(1+\tau_x)\phi_d \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}}. \qquad (67)
\end{aligned}
$$

With the equilibrium wages we have gotten above, we get:

$$
w_r = (1+\tau_x)(1-\beta)\phi_d \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \Rightarrow \frac{d\log w_r}{d\log(1+\tau_x)} = 1
$$

$$
w_n = \alpha\frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_x)\phi_d]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_n}{d\log(1+\tau_x)} = -\frac{1-\alpha}{\alpha}
$$

$$
\frac{w_r}{w_n} = \frac{[(1+\tau_x)\phi_d]^{\frac{1}{\alpha}}}{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}} \left(\frac{\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \Rightarrow \frac{d\log w_r/w_n}{d\log(1+\tau_x)} = \frac{1}{\alpha}
$$

Plugging these back into the envelope condition (61), we get:

$$\begin{aligned}
W'\left(\tau_x\right) &= -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r} l_r\right) \frac{1}{\alpha\left(1+\tau_x\right)} \frac{w_r l_r}{w_n} + \mu \pi_n w_n l_n \frac{\tau_x+\alpha}{\alpha\left(1+\tau_x\right)^2}\left[-\frac{1-\alpha}{\alpha}+\frac{1-\alpha}{\tau_x+\alpha}\right] \\
&= \frac{1}{\alpha\left(1+\tau_x\right)}\left[-\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_n} l_r\right) \frac{w_r l_r}{w_n} - \mu \pi_n w_n l_n \frac{\tau_x}{1+\tau_x} \frac{1-\alpha}{\alpha}\right]
\end{aligned}$$

Since $\mu > 0$, if $\tau_x \leq 0$, then we have:

$$W'(\tau_x) > 0.$$

Thus, the social planner can always improve welfare by marginally increasing τ_x. Furthermore, the optimal level of τ_x implies $W'(\tau_x) = 0$, so we get:

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\eta_n\left(-u_l\left(c_r, \frac{w_r}{w_n} l_r\right) \frac{w_r l_r}{w_n}\right)}{\mu w_n l_n} \tag{68}$$

The first order condition with respect to l_r yields:

$$\begin{aligned}
-\frac{\eta_n}{\mu} u_l\left(c_r, \frac{w_r}{w_r} l_r\right) \frac{w_r l_r}{w_n} &= -\left(\frac{\widetilde{\omega}_r \pi_r u_l\left(c_r, l_r\right) l_r + \frac{\pi_r w_r l_r}{1+\tau_x}}{\pi_n}\right) \\
&= \frac{\pi_r \phi_d l_r}{\pi_n}\left(\frac{\phi_r}{\beta \phi_d}\right)^{\frac{\beta}{\beta-1}} \phi_d(1-\beta)\left[\frac{\widetilde{\omega}_r\left(-u_l\left(c_r, l_r\right)\right)}{\left(\frac{\phi_r}{\beta \phi_d}\right)^{\frac{\beta}{\beta-1}} \phi_d(1-\beta)} - 1\right]
\end{aligned} \tag{69}$$

Thus, we get

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\pi_r l_r}{\pi_n}\left(\frac{\phi_r}{\beta \phi_d}\right)^{\frac{\beta}{\beta-1}} \phi_d(1-\beta)\left[\frac{\widetilde{\omega}_r\left(-u_l\left(c_r, l_r\right)\right)}{\left(\frac{\phi_r}{\beta \phi_d}\right)^{\frac{\beta}{\beta-1}} \phi_d(1-\beta)} - 1\right] \tag{70}$$

C $0 < q < 1$ case

We showed that the statement holds in the limit cases of $q \longrightarrow 0$ and $q \longrightarrow 1$. Thus, we are left with non-limit cases. The equilibrium conditions are stated in the main text for $0 < q < 1$ case. Although we cannot solve for the closed form solutions of endogenous variables w_r and m, we can deduce the sign of the comparative statics with respect to τ_x.

Recall that w_n is not a function of q. Recall that

$$w_r = (1-\beta)(1+\tau_x)\phi_d \left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta} \right)^{\frac{1}{q-1}} \tag{71}$$

Letting $\eta_r = 0$, we get

$$\begin{aligned} W'(\tau_x) = & -\eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r \right) \frac{d\log(w_r/w_n)}{d\log(1+\tau_x)} \frac{1}{1+\tau_x} \frac{w_r l_r}{w_n} \\ & + \mu \Bigg[\pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)^2} \left(\frac{d\log w_n}{d\log(1+\tau_x)} + \frac{1-\alpha}{(\tau_x+\alpha)} \right) + \\ & \frac{\pi_r w_r l_r}{(1+\tau_x)^2} \left(\frac{d\log w_r}{d\log(1+\tau_x)} - 1 \right) \Bigg]. \end{aligned} \tag{72}$$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1-\beta)(1+\tau_x)\phi_d \left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1} - \beta} \right)^{\frac{1}{q-1}} \Rightarrow \frac{d\log w_r}{d\log(1+\tau_x)} = 1$$

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[(1+\tau_x)\phi_d\right]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_n}{d\log(1+\tau_x)} = -\frac{1-\alpha}{\alpha}$$

$$\frac{d\log w_r/w_n}{d\log(1+\tau_x)} = \frac{1}{\alpha}$$

Plugging these back into the envelope condition (61), we get:

$$\begin{aligned}W'\left(\tau_x\right) &= -\eta_n\pi_n u_l\left(c_r, \frac{w_r}{w_r}l_r\right)\frac{1}{\alpha\left(1+\tau_x\right)}\frac{w_r l_r}{w_n} + \mu\pi_n w_n l_n\frac{\tau_x+\alpha}{\alpha\left(1+\tau_x\right)^2}\left[-\frac{1-\alpha}{\alpha}+\frac{1-\alpha}{\tau_x+\alpha}\right] \\ &= \frac{1}{\alpha\left(1+\tau_x\right)}\left[-\eta_n\pi_n u_l\left(c_r, \frac{w_r}{w_n}l_r\right)\frac{w_r l_r}{w_n} - \mu\pi_n w_n l_n\frac{\tau_x}{1+\tau_x}\frac{1-\alpha}{\alpha}\right]\end{aligned}$$

Since $\mu > 0$, if $\tau_x \leq 0$, if $w_r \geq 0$—that is, $\phi_r \leq \phi_d$—, then we are guaranteed to have:

$$W'(\tau_x) > 0.$$

Thus, the social planner can always improve welfare by marginally increasing τ_x.

Furthermore, the optimal level of τ_x implies $W'(\tau_x) = 0$, so we get:

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha}\frac{\eta_n\left(-u_l\left(c_r, \frac{w_r}{w_n}l_r\right)\frac{w_r l_r}{w_n}\right)}{\mu w_n l_n} \tag{73}$$

The first order condition with respect to l_r yields:

$$\begin{aligned}-\frac{\eta_n}{\mu}u_l\left(c_r, \frac{w_r}{w_r}l_r\right)\frac{w_r l_r}{w_n} &= -\left(\frac{\widetilde{\omega}_r\pi_r u_l\left(c_r, l_r\right)l_r + \frac{\pi_r w_r l_r}{1+\tau_x}}{\pi_n}\right) \\ &= \frac{\pi_r l_r}{\pi_n}(1-\beta)\phi_d\left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}-\beta}\right)^{\frac{1}{q-1}} \\ &\quad \left[\frac{\widetilde{\omega}_r\left(-u_l\left(c_r, l_r\right)\right)}{(1-\beta)\phi_d\left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}-\beta}\right)^{\frac{1}{q-1}}} - 1\right]\end{aligned}$$

Thus, we get

$$
\begin{aligned}
\frac{\tau_x}{1+\tau_x} =& \frac{\alpha}{1-\alpha}\frac{\pi_r l_r}{\pi_n}(1-\beta)\phi_d\left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}-\beta}\right)^{\frac{1}{q-1}} \\
&\left[\frac{\widetilde{\omega}_r\left(-u_l\left(c_r,l_r\right)\right)}{(1-\beta)\phi_d\left(\frac{(1-\beta)\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}}{\left(\frac{\phi_r}{\beta\phi_d}\right)^{q-1}-\beta}\right)^{\frac{1}{q-1}}}-1\right]
\end{aligned}
$$

■

D Proof of Proposition 3

Again, as in the proof of Proposition 2, I separately prove the limit cases ($q \longrightarrow 0$ and $q \longleftrightarrow 1$) and interior case ($0 < q < 1$).

A $q \longrightarrow 1$ case

The equilibrium levels of inputs will be

$$X_r = \left(\frac{(1+\tau_r)\phi_r}{(1+\tau_d)\beta\phi_d} \right)^{\frac{1}{\beta-1}} \pi_r l_r. \tag{74}$$

$$\begin{aligned} m &= 1 - \left(\frac{(1+\tau_d)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{(\pi_r l_r)^{(1-\beta)}}{\pi_n l_n} X_r^{\beta} \\ &= 1 - \left(\frac{(1+\tau_d)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{(\pi_r l_r)^{(1-\beta)}}{\pi_n l_n} \left(\left(\frac{(1+\tau_r)\phi_r}{(1+\tau_d)\beta\phi_d} \right)^{\frac{1}{\beta-1}} \pi_r l_r \right)^{\beta} \\ &= 1 - \left(\frac{(1+\tau_d)\phi_d}{(1-\alpha)A} \right)^{\frac{1}{\alpha}} \frac{\pi_r l_r}{\pi_n l_n} \left(\frac{(1+\tau_r)\phi_r}{(1+\tau_d)\beta\phi_d} \right)^{\frac{\beta}{\beta-1}}. \end{aligned} \tag{75}$$

Furthermore,

$$\begin{aligned} w_r &= (1-\beta)(1+\tau_d)\phi_d X_r^{\beta} (\pi_r l_r)^{-\beta} \\ &= (1-\beta)(1+\tau_d)\phi_d \left(\left(\frac{(1+\tau_r)\phi_r}{(1+\tau_d)\beta\phi_d} \right)^{\frac{1}{\beta-1}} \pi_r l_r \right)^{\beta} (\pi_r l_r)^{-\beta} \\ &= (1-\beta)\phi_d (1+\tau_d)^{\frac{1}{1-\beta}} \left(\frac{(1+\tau_r)\phi_r}{\beta\phi_d} \right)^{\frac{\beta}{\beta-1}}. \end{aligned} \tag{76}$$

Let $\tau_x = (\tau_d, \tau_r)$. Define $W(\tau_x) = \max \pi_r \omega_r u(c_r, l_r) + \pi_n \omega_n u(c_n, l_n) + v(G)$. Then, the social planner's optimization problem is

$$\underset{\tau_x}{\text{maximize}}\ W(\tau_x)$$

subject to

$$[\eta_r \pi_r] \quad u\left(c_r, l_r\right) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right),$$

$$[\eta_n \pi_n] \quad u\left(c_n, l_n\right) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right),$$

$$[\mu] \quad \pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \left(\frac{\tau_d + \alpha}{\alpha(1+\tau_d)}\right) + \pi_r w_r l_r \left(\frac{1+\tau_r - \beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)(1-\beta)}\right).$$

Letting $\eta_r = 0$, and taking the partial derivative of the welfare function with respect to τ_r, we get

$$\begin{aligned} \frac{\partial W(\tau_r)}{\partial \tau_r} = & -\eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r\right) \frac{d \log\left(w_r / w_n\right)}{d \log\left(1+\tau_r\right)} \frac{1}{1+\tau_r} \frac{w_r l_r}{w_n} \\ & + \mu \left[\pi_r w_r l_r \frac{1+\tau_r - \beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)^2(1-\beta)} \left(\frac{d \log w_r}{d \log(1+\tau_r)}\right) + \right. \\ & \left. + \pi_r w_r l_r \frac{\beta}{(1-\beta)(1+\tau_r)^2}\right]. \end{aligned} \tag{77}$$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1-\beta)\phi_d(1+\tau_d)^{\frac{1}{1-\beta}} \left(\frac{(1+\tau_r)\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \Rightarrow \frac{d \log w_r}{d \log\left(1+\tau_r\right)} = \frac{\beta}{\beta - 1}$$

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[\left(1+\tau_d\right)\phi_d\right]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d \log w_n}{d \log\left(1+\tau_r\right)} = 0$$

$$\frac{w_r}{w_n}=\frac{\left[(1+\tau_d)\,\phi_d\right]^{\frac{1-\beta+\alpha\beta}{\alpha(1-\beta)}}}{\alpha A^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}\left(\frac{(1+\tau_r)\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}}\Rightarrow\frac{d\log w_r/w_n}{d\log\left(1+\tau_r\right)}=\frac{\beta}{\beta-1}$$

Plugging these back into the envelope condition above, we get:

$$\begin{aligned}\frac{\partial W(\tau_r)}{\partial\tau_r}&=-\eta_n\pi_n u_l\left(c_r,\frac{w_r}{w_r}l_r\right)\frac{\beta}{\beta-1}\frac{1}{(1+\tau_r)}\frac{w_r l_r}{w_n}+\mu\pi_r w_r l_r\frac{\beta}{\beta-1}\frac{1+\tau_r-\beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)^2(1-\beta)}\\&\quad+\mu\pi_r w_r l_r\frac{\beta}{(1-\beta)(1+\tau_r)^2}\\&=\frac{1}{(1+\tau_r)}\frac{\beta}{\beta-1}\left[-\eta_n\pi_n u_l\left(c_r,\frac{w_r}{w_n}l_r\right)\frac{w_r l_r}{w_n}+\mu\pi_r w_r l_r\frac{\tau_r-\tau_d}{(1+\tau_d)(1+\tau_r)(1-\beta)}\right]\end{aligned}.$$

Notice that since $\eta>0$, the first term inside the closed bracket is positive, and since $\mu>0$, the second term inside the closed parenthesis is positive only if $\tau_r>\tau_d$. This holds true when $\tau_r=0$. Thus, when $\tau_r=0$ *and when* $\tau_d=0$, or when $\tau_r<0$ *and when* $\tau_d\leq\tau_r$, we are guaranteed to get

$$\frac{\partial W(\tau_r)}{\partial(-\tau_r)}>0. \tag{78}$$

Therefore, the social planner can achieve greater welfare by imposing a positive reinstating capital subsidy, $\tau_r<0$, when $\tau_d\leq\tau_r$.

Now, to get the optimal rate of τ_r, we equate the envelope condition to be zero and get

$$-\eta_n\pi_n u_l\left(c_r,\frac{w_r}{w_n}l_r\right)\frac{w_r l_r}{w_n}+\mu\pi_r w_r l_r\frac{\tau_r-\tau_d}{(1+\tau_d)(1+\tau_r)^2(1-\beta)}=0$$

The FOC of l_r implies

$$-\eta_n\pi_n u_l\left(c_r,\frac{w_r}{w_r}l_r\right)\frac{w_r l_r}{w_n}=-\left(\omega_r\pi_r u_l\left(c_r,l_r\right)l_r+\mu\pi_r w_r l_r\left(\frac{1+\tau_r-\beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)(1-\beta)}\right)\right). \tag{79}$$

Plugging this back into the condition, we get

$$\mu\pi_r w_r l_r \frac{\tau_r - \tau_d}{(1+\tau_d)(1+\tau_r)^2(1-\beta)} = \omega_r \pi_r u_l\left(c_r, l_r\right) l_r + \mu\pi_r w_r l_r \left(\frac{1+\tau_r-\beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)(1-\beta)}\right)$$

That is

$$\frac{\mu\pi_r w_r l_r}{1+\tau_r} = -\omega_r \pi_r u_l\left(c_r, l_r\right) l_r. \tag{80}$$

Next, taking the partial derivative of the welfare function with respect to τ_d, we get

$$\begin{aligned}\frac{\partial W(\tau_d)}{\partial \tau_d} =& -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r} l_r\right) \frac{d\log\left(w_r/w_n\right)}{d\log\left(1+\tau_d\right)} \frac{1}{1+\tau_d} \frac{w_r l_r}{w_n} \\ & + \mu\Bigg[\pi_n w_n l_n \frac{\tau_d+\alpha}{\alpha(1+\tau_d)^2}\left(\frac{d\log w_n}{d\log(1+\tau_d)} + \frac{1-\alpha}{(\tau_d+\alpha)}\right) + \\ & \pi_r w_r l_r \frac{1+\tau_r-\beta(1+\tau_d)}{(1+\tau_d)^2(1+\tau_r)(1-\beta)} \frac{d\log w_r}{d\log(1+\tau_d)} - \\ & \pi_r w_r l_r \frac{1}{(1+\tau_d)^2(1-\beta)}\Bigg].\end{aligned} \tag{81}$$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1-\beta)\phi_d(1+\tau_d)^{\frac{1}{1-\beta}}\left(\frac{(1+\tau_r)\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \Rightarrow \frac{d\log w_r}{d\log\left(1+\tau_d\right)} = \frac{1}{1-\beta}$$

$$w_n = \alpha\frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[\left(1+\tau_d\right)\phi\right]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_n}{d\log\left(1+\tau_d\right)} = -\frac{1-\alpha}{\alpha}$$

$$\frac{w_r}{w_n} = \frac{\left[\left(1+\tau_d\right)\phi_d\right]^{\frac{1-\beta+\alpha\beta}{\alpha(1-\beta)}}}{\alpha A^{\frac{1}{\alpha}}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}\left(\frac{(1+\tau_r)\phi_r}{\beta\phi_d}\right)^{\frac{\beta}{\beta-1}} \Rightarrow \frac{d\log w_r/w_n}{d\log\left(1+\tau_d\right)} = \frac{1-\beta+\alpha\beta}{\alpha(1-\beta)}$$

Plugging these back into the envelope condition above, we get:

$$\begin{aligned}\frac{\partial W(\tau_d)}{\partial \tau_d} &= -\eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r\right) \frac{1-\beta+\alpha\beta}{\alpha(1-\beta)} \frac{w_r l_r}{w_n(1+\tau_d)} + \\ &\quad \mu \left(\pi_n w_n l_n \frac{\tau_d+\alpha}{\alpha\left(1+\tau_d\right)^2}\left[-\frac{1-\alpha}{\alpha}+\frac{1-\alpha}{\tau_d+\alpha}\right] + \pi_r w_r l_r \left[\frac{\beta(\tau_r-\tau_d)}{(1-\beta)^2(1+\tau_d)^2(1+\tau_r)}\right]\right) \\ &= \frac{1}{\alpha\left(1+\tau_d\right)\left(1-\beta\right)}\Bigg[-\eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_n} l_r\right) \frac{w_r l_r}{w_n}(1-\beta+\alpha\beta) - \\ &\quad \mu \pi_n w_n l_n \frac{\tau_d}{1+\tau_d} \frac{(1-\alpha)(1-\beta)}{\alpha} + \mu \pi_r w_r l_r \frac{\alpha\beta(\tau_r-\tau_d)}{(1-\beta)(1+\tau_r)(1+\tau_d)}\Bigg]\end{aligned}$$

Since $\mu > 0$, if $\tau_d \leq 0$, and if $\tau_d \leq \tau_r$ then we have:

$$\frac{\partial W(\tau_d)}{\partial \tau_d} > 0.$$

Thus, the social planner can always improve welfare by marginally increasing τ_d. Furthermore, the optimal level of τ_d implies $\frac{\partial W(\tau_d)}{\partial \tau_d} = 0$ for a given level of τ_r.

The FOC of l_r implies

$$-\eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r\right) \frac{w_r l_r}{w_n} = -\left(\omega_r \pi_r u_l \left(c_r, l_r\right) l_r + \mu \pi_r w_r l_r \left(\frac{1+\tau_r-\beta(1+\tau_d)}{(1+\tau_d)(1+\tau_r)(1-\beta)}\right)\right). \tag{82}$$

Then, plugging this back into the envelope condition, we get

$$\begin{aligned}\pi_r \omega_r u_l(c_r, l_r) l_r = &- \mu \pi_n w_n l_n \frac{\tau_d}{1+\tau_d} \frac{(1-\alpha)(1-\beta)}{\alpha} + \qquad (83) \\ &\mu \pi_r w_r l_r \frac{\beta(2+\tau_d-\tau_r+\alpha(2\tau_r-2))-1-\tau_d+(\alpha-2)\beta^2(1+\tau_d)-\tau_r}{(1-\beta)(1+\tau_r)(1+\tau_d)}. \qquad (84)\end{aligned}$$

Then, we have two equations for two unknowns from (80) and (84).

B $q \longrightarrow 0$

Let $\tau_x = (\tau_d, \tau_r)$. Define $W(\tau_x) = \max \pi_r \omega_r u\left(c_r, l_r\right) + \pi_n \omega_n u\left(c_n, l_n\right) + v(G)$. Then, the social planner's optimization problem is

$$\underset{\tau_x}{\text{maximize}}\ W(\tau_x)$$

subject to

$$
\begin{aligned}
&[\eta_r \pi_r] \quad u\left(c_r, l_r\right) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right), \\
&[\eta_n \pi_n] \quad u\left(c_n, l_n\right) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right), \\
&[\mu] \quad \pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \left(\frac{\tau_d + \alpha}{\alpha(1+\tau_d)}\right) + \pi_r w_r l_r \frac{\phi_d - \phi_r}{(1+\tau_d)\phi_d - (1+\tau_r)\phi_r}.
\end{aligned}
$$

Letting $\eta_r = 0$, and taking the partial derivative of the welfare function with respect to τ_r, we get

$$\frac{\partial W(\tau_r)}{\partial \tau_r} = -\, \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_n} l_r\right) \frac{dw_r}{d\tau_r} \frac{l_r}{w_n}. \tag{85}$$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1+\tau_d)\phi_d - (1+\tau_r)\phi_r \Rightarrow \frac{dw_r}{d\tau_r} = -\phi_r$$

Plugging this back into the envelope condition above, we get:

$$\frac{\partial W(\tau_r)}{\partial \tau_r} = \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r \right) \frac{l_r}{w_n} \phi_r \quad .$$

Notice that since $\eta > 0$, this term is negative. This holds true at all τ_r. Therefore, we get

$$\frac{\partial W(\tau_r)}{\partial (-\tau_r)} > 0. \tag{86}$$

Therefore, the social planner can achieve greater welfare by imposing a positive reinstating capital subsidy, $\tau_r < 0$.

Furthermore, the condition above implies that the optimal $\tau_r = -1$. The intuition behind this result is discussed in the main text.

Next, taking the partial derivative of the welfare function with respect to τ_d, we get

$$\begin{aligned} \frac{\partial W(\tau_d)}{\partial \tau_d} = & - \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r \right) \frac{d \log (w_r / w_n)}{d \log (1+\tau_d)} \frac{1}{1+\tau_d} \frac{w_r l_r}{w_n} \\ & + \mu \left[\pi_n w_n l_n \frac{\tau_d + \alpha}{\alpha (1+\tau_d)^2} \left(\frac{d \log \left[w_n \left(\frac{\tau_d + \alpha}{\alpha(1+\tau_d)} \right) \right]}{d \log (1+\tau_d)} \right) \right]. \end{aligned} \tag{87}$$

With the equilibrium wages we have gotten above, we get:

$$w_r = (1+\tau_d)\phi_d - (1+\tau_r)\phi_r \Rightarrow \frac{d \log w_r}{d \log (1+\tau_d)} = \frac{\phi_d (1+\tau_d)}{(1+\tau_d)\phi_d - (1+\tau_r)\phi_r}$$

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[(1+\tau_d)\,\phi_d\right]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d \log w_n}{d \log (1+\tau_r)} = -\frac{1-\alpha}{\alpha}$$

$$\frac{w_r}{w_n} = \frac{(1+\tau_d)\phi_d - (1+\tau_r)\phi_r}{\alpha\frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_d)\phi_d]^{\frac{1-\alpha}{\alpha}}}} \Rightarrow \frac{d\log w_r/w_n}{d\log(1+\tau_r)} = \frac{\phi_d(1+\tau_d)}{(1+\tau_d)\phi_d - (1+\tau_r)\phi_r} + \frac{1-\alpha}{\alpha}$$

Plugging these back into the envelope condition above, we get:

$$\frac{\partial W(\tau_d)}{\partial \tau_d} = -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r}l_r\right)\frac{w_r l_r}{w_n}\left(\frac{\phi_d}{(1+\tau_d)\phi_d - (1+\tau_r)\phi_r} + \frac{1-\alpha}{\alpha(1+\tau_d)}\right) +$$
$$\mu\left(\pi_n w_n l_n \frac{\tau_d+\alpha}{\alpha(1+\tau_d)^2}\left[-\frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\tau_d+\alpha}\right]\right).$$

Given that the optimal $\tau_r = -1$, we get

$$\frac{\partial W(\tau_d)}{\partial \tau_d} = -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r}l_r\right)\frac{w_r l_r}{w_n}\frac{1}{\alpha(1+\tau_d)} +$$
$$\mu\left(\pi_n w_n l_n \frac{\tau_d+\alpha}{\alpha(1+\tau_d)^2}\left[-\frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\tau_d+\alpha}\right]\right).$$

Notice that the first and second terms are always positive. Note that the third term is positive when

$$\frac{\tau_d+\alpha}{\alpha(1+\tau_d)^2}\left[-\frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\tau_d+\alpha}\right] > 0$$

This is true when $\tau_d \leq 0$. Given this condition, we get

$$\frac{\partial W(\tau_d)}{\partial \tau_d} > 0.$$

Thus, the social planner can always improve welfare by marginally increasing τ_d. Notice that the form of the optimal τ_d is similar to the one in GRT, and we have

$$\frac{\tau_d}{1+\tau_d} = \frac{\alpha}{1-\alpha}\frac{\pi_r l_r}{\pi_n w_n l_n}(\phi_d - \phi_r)\left[\frac{\omega_r}{\mu}\frac{(-u_l(c_r, l_r))}{\phi_d - \phi_r} - 1\right]$$

C $0 < q < 1$

We showed that the statement holds in the limit cases of $q \longrightarrow 0$ and $q \longrightarrow 1$. Thus, we are left with non-limit cases. The equilibrium conditions are stated in the main text for $0 < q < 1$ case. Although we cannot solve for the closed form solutions of endogenous variables w_r and m, we can deduce the sign of the comparative statics with respect to τ_x.

Recall that w_n is not a function of q. It is useful to derive that at equilirbium, after some algebra, we get

$$Y = A\left(X_d + \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)^{1-\alpha}(\pi_n l_n)^{\alpha} \tag{88}$$

$$= \frac{w_n \pi_n l_n}{\alpha} \tag{89}$$

$$\int_0^1 (\phi_d x_d + \phi_r x_r)di = \phi_d \frac{m}{1-m}\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}} + \phi_r X_r \tag{90}$$

$$= \phi_d\left(\frac{1}{\left(\frac{(1+\tau_d)\phi_d}{(1-\alpha)A}\right)^{\frac{1}{\alpha}}\frac{1}{\pi_n l_n}} - \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right) + \phi_r X_r \tag{91}$$

$$X_r = \left(\frac{1-\beta}{\left(\frac{(1+\tau_r)\phi_r}{\beta(1+\tau_d)\phi_d}\right)^{q-1} - \beta}\right)^{\frac{q}{q-1}} \pi_r l_r \tag{92}$$

$$
\begin{aligned}
w_r &= (1-\beta)(1+\tau_d)\phi_d \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}} \right)^{\frac{1}{q-1}} (\pi_r l_r)^{-\frac{1}{q}} \\
&= (1-\beta)(1+\tau_d)\phi_d \left((1-\beta) \left(\frac{\beta}{\left(\frac{(1+\tau_r)\phi_r}{\beta(1+\tau_d)\phi_d}\right)^{q-1} - \beta} + 1 \right) \right)^{\frac{1}{q-1}} \qquad (93)
\end{aligned}
$$

Note that the government problem with the income taxation specified in Lemma 1 is the following:

$$
\underset{\tau_x}{\text{maximize}}\ W(\tau_x)
$$

subject to

$$
\begin{aligned}
&[\eta_r \pi_r] \quad u(c_r, l_r) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right), \\
&[\eta_n \pi_n] \quad u(c_n, l_n) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right), \\
&[\mu] \quad \pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n - \\
&\qquad \phi_d \pi_n l_n \left(\frac{(1-\alpha)A}{(1+\tau_d)\phi_d}\right)^{\frac{1}{\alpha}} + \phi_d \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta) N_r^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}} - \phi_r X_r,
\end{aligned}
$$

where $W(\tau_x) = \max \pi_r \omega_r u(c_r, l_r) + \pi_n \omega_n u(c_n, l_n) + v(G)$.

Then, for fixed τ_d, the envelope condition with respect to τ_r gives

$$\begin{aligned}W'(\tau_r) = &- \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r\right) \frac{dw_r}{d\left(1+\tau_r\right)} \frac{l_r}{w_n} \\ &+ \mu \Bigg[\phi_d \frac{d \log \left(\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)}{d \log(1+\tau_r)} \\ &\qquad \frac{\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}}{1+\tau_r} - \phi_r \frac{dX_r}{d(1+\tau_r)}\Bigg]. \end{aligned} \tag{94}$$

Note that

$$\frac{w_r}{d(1+\tau_r)} = \beta(1-\beta)(1+\tau_d)\phi_d(\pi_r l_r)^{-\frac{1}{q}} X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_r)} \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta) N_r^{\frac{q-1}{q}}\right)^{\frac{2-q}{q-1}},$$

$$\frac{d \log \left(\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)}{d \log(1+\tau_r)} = \beta X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_r)} \frac{1+\tau_r}{\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}},$$

and recall that from the first order condition of X_r on the production side equation (**??**), we have

$$(1+\tau_x)\phi_r = \beta(1+\tau_x)\phi_d \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{1}{q-1}} X_r^{-\frac{1}{q}}.$$

That is,

$$X_r^{-\frac{1}{q}} = \frac{\phi_r}{\beta\phi_d} \left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{-1}{q-1}}. \tag{95}$$

Then, we get

$$
\begin{aligned}
W'(\tau_r) = &- \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r \right) \beta(1-\beta)(1+\tau_d)\phi_d(\pi_r l_r)^{-\frac{1}{q}} X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_r)} \\
&\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta) N_r^{\frac{q-1}{q}} \right)^{\frac{2-q}{q-1}} \frac{l_r}{w_n}.
\end{aligned} \tag{96}
$$

Notice that since $\eta > 0$ and the derivative of X_r with respect to τ_r is negative, the whole term is negative. Therefore, we have

$$
\frac{dW(\tau_r)}{d(-\tau_r)} > 0
$$

Now, we fix τ_r and determine the sign of optimal τ_d.

$$
\begin{aligned}
W'(\tau_d) = &- \eta_n \pi_n u_l \left(c_r, \frac{w_r}{w_r} l_r \right) \frac{d \log(w_r/w_n)}{d \log(1+\tau_d)} \frac{1}{1+\tau_d} \frac{w_r l_r}{w_n} \\
&+ \mu \Bigg[\pi_n w_n l_n \frac{1}{\alpha(1+\tau_d)} \frac{d \log w_n}{d \log(1+\tau_d)} \\
&- \phi_d \left(\frac{(1-\alpha)A}{(1+\tau_d)\phi_d} \right)^{\frac{1}{\alpha}} \left[\frac{d}{d \log(1+\tau_d)} \log \left(\left(\frac{(1-\alpha)A}{(1+\tau_d)\phi_d} \right)^{\frac{1}{\alpha}} \right) \right] \frac{\pi_n l_n}{1+\tau_d} \\
&+ \phi_d \frac{d \log \left(\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}} \right)^{\frac{q}{q-1}} \right)}{d \log(1+\tau_d)} \\
&\frac{\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}} \right)^{\frac{q}{q-1}}}{1+\tau_d} - \phi_r \frac{dX_r}{d(1+\tau_d)} \Bigg].
\end{aligned} \tag{97}
$$

Note that

$$\frac{d\log w_r}{d\log(1+\tau_d)} = 1 + \beta X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_d)} \frac{1+\tau_d}{\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}}$$

$$\frac{d\log \frac{w_r}{w_n}}{d\log(1+\tau_d)} = 1 + \beta X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_d)} \frac{1+\tau_d}{\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}} + \frac{1-\alpha}{\alpha}$$

$$\frac{d\log\left(\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{q}{q-1}}\right)}{d\log(1+\tau_d)} = \beta X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_d)} \frac{1+\tau_d}{\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}},$$

and again recall that

$$X_r^{-\frac{1}{q}} = \frac{\phi_r}{\beta\phi_d}\left(\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}\right)^{\frac{-1}{q-1}}. \tag{98}$$

Then, we get

$$\begin{aligned} W'(\tau_d) = & -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r} l_r\right)\left(1 + X_r^{-\frac{1}{q}} \frac{dX_r}{d(1+\tau_d)} \frac{1+\tau_d}{\beta X_r^{\frac{q-1}{q}} + (1-\beta)(\pi_r l_r)^{\frac{q-1}{q}}} + \frac{1-\alpha}{\alpha}\right) \\ & \frac{1}{1+\tau_d} \frac{w_r l_r}{w_n} + \mu\left[\frac{\alpha-1}{\alpha^2} \frac{\pi_n w_n l_n}{1+\tau_d} + \phi_d\left(\frac{(1-\alpha)A}{(1+\tau_d)\phi_d}\right)^{\frac{1}{\alpha}} \frac{\pi_n l_n}{1+\tau_d}\right] \end{aligned} \tag{99}$$

Notice that the derivative of X_r with respect to τ_d is positive, so the first term is guaranteed to be positive. The second term is always negative (since $0 < \alpha < 1$), while the third term is positive. Then, notice

$$\frac{\alpha-1}{\alpha^2} \frac{\pi_n w_n l_n}{1+\tau_d} + \phi_d\left(\frac{(1-\alpha)A}{(1+\tau_d)\phi_d}\right)^{\frac{1}{\alpha}} \frac{\pi_n l_n}{1+\tau_d} \geq 0$$

when

$$-(1+\tau_d)\phi_d + \phi_d \geq 0.$$

Thus, when $\tau_d \leq 0$, the inequality holds true. Therefore, when $\tau_d \leq 0$, $W'(\tau_d) > 0$.

E Dynamic Model

In this section, I shall extend the previous results to a dynamic model in which capital is now an investment good. Note that without weak separability between consumption and leisure, the uniform taxation result in Atkinson and Stiglitz (1976) fails. Thus, to demonstrate that the channel of my non-uniform tax result below does not come from the non-separability, I add an additive separability assumption for consumption and labor into the household utility function. Following Guerreiro et al. (2019), I focus on a case where households can only change the hours worked but cannot alter their occupation that they are born into.

.1 households

The utility function of household type $j \in \{r, n\}$ is now defined as

$$U_j = \sum_{t=0}^{\infty} \beta^t \left[u\left(c_{j,t}\right) - h\left(l_{j,t}\right) + v\left(G_t\right) \right]. \tag{100}$$

Let $u'(c) > 0, h'(l) > 0, v'(G) > 0,$ and $u''(c) < 0, h''(l) > 0, v''(G) < 0$ denote the first and second derivatives of each function, respectively.

.2 Technology

Production of the final good is expressed as

$$Y_t = F\left(N_{n,t}, N_{r,t}, X_{d,t}, X_{r,t}\right), \tag{101}$$

where Y_t indicates total production at time t, $N_{n,t}$ and $N_{r,t}$ denote total mass of non-routine and routine labor, respectively, and $X_{d,t}$ and $X_{r,t}$ represent total units of displacing and reinstating capital used, respectively. I assume that $F(.)$ satisfies the constant returns to scale property in its inputs. The planner can observe a household's effective contribution to production, $y_{j,t}$. A household's effective contribution equals the household's marginal productivity multiplied by the total number of hours worked, $y_{j,t} = F_{N_j}(t)l_{j,t}$ for $j = r, n,$ where $F_{N_j}(t) = \partial F\left(N_{n,t}, N_{r,t}, X_{d,t}, X_{r,t}\right)/\partial N_{j,t}$. We notate the marginal productivity of each capital by $F_{X_k}(t) = \partial F\left(N_{n,t}, N_{r,t}, X_{d,t}, X_{r,t}\right)/\partial X_{k,t}$ for $k \in d, r$.

I also assume that displacing capital has a higher degree of complementarity to non-routine labor than routine labor, while reinstating capital has a greater degree of complementarity to routine labor than non-routine labor. That is, I assume

$$\varepsilon_{F_{N_r}/F_{N_n},X_d}(t) = \frac{d\log\left(\frac{F_{N_r}(t)}{F_{N_n}(t)}\right)}{d\log X_{d,t}} < 0, \tag{102}$$

$$\varepsilon_{F_{N_r}/F_{N_n},X_r}(t) = \frac{d\log\left(\frac{F_{N_r}(t)}{F_{N_n}(t)}\right)}{d\log X_{r,t}} > 0. \tag{103}$$

A unit cost of displacing capital is $\phi_{d,t}$ units of output, while that of reinstating capital is $\phi_{r,t}$ units of output. The stock of capital evolves according to the following rule:

$$X_{k,t+1} = \left(1 - \delta_X\right)X_{k,t} + i_{k,t}/\phi_{k,t}, \quad \text{for } k \in \{d, r\}, \tag{104}$$

where δ_X denotes the depreciation rate of both types of capital. Note that I assume the same depreciation rate for the two types of capital. With this, I define the resource constraint in period t to be:

$$\pi_r c_{r,t} + \pi_n c_{n,t} + G_t \leq F\left(\pi_n l_{n,t}, \pi_r l_{r,t}, X_{d,t}, X_{r,t}\right) - \\ \phi_{d,t}\left[X_{d,t+1} - \left(1-\delta_X\right)X_{d,t}\right] - \phi_{r,t}\left[X_{r,t+1} - \left(1-\delta_X\right)X_{r,t}\right]. \tag{105}$$

.3 First-best Allocation

Pareto weights for routine and non-routine households can take any non-negative value, $\omega_r, \omega_n \geq 0$ as long as normalized such that $\omega_r\pi_r + \omega_n\pi_n = 1$. Then, the planner's objective function is

$$\sum_{t=0}^{\infty}\beta^t\left\{\sum_{j=n,r}\omega_j\pi_j\left[u\left(c_{j,t}\right) - h\left(l_{j,t}\right)\right] + v\left(G_t\right)\right\}. \tag{106}$$

The first-best allocation optimizes the above function subject to the above resource constraint (89). The first order conditions for consumption and labor at time t together imply

$$\frac{h'\left(l_{j,t}\right)}{u'\left(c_{j,t}\right)} = F_{N_j}(t). \tag{107}$$

Furthermore, the first order conditions for consumption at t and $t+1$ together with the first order condition for capital at $t+1$ imply the intertemporal trade-offs

$$\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} = \frac{F_{X_k}(t+1) + \phi_{k,t+1}\left(1-\delta_x\right)}{\phi_{k,t}} \quad \text{for } k \in \{d,r\},\ j \in \{n,r\}. \tag{108}$$

Moreover, the first-best allocation requires

$$\omega_r u'\left(c_{r,t}\right) = \omega_n u'\left(c_{n,t}\right). \tag{109}$$

Note that these conditions together with the resource constraint with equality characterize the Pareto frontier, as we adjust the Pareto weights.

.4 Second Best Planning Problem

In this section, I consider a case where the planner cannot discriminate/distinguish between two types of labor but can distinguish between two types of capital. In this case, we have the following incentive constraints for workers:

$$\sum_{t=0}^{\infty} \beta^t \left[u\left(c_{r,t}\right) - h\left(l_{r,t}\right)\right] \geq \sum_{t=0}^{\infty} \beta^t \left[u\left(c_{n,t}\right) - h\left(\frac{F_{N_n}(t)}{F_{N_r}(t)} l_{n,t}\right)\right], \tag{110}$$

$$\sum_{t=0}^{\infty} \beta^t \left[u\left(c_{n,t}\right) - h\left(l_{n,t}\right)\right] \geq \sum_{t=0}^{\infty} \beta^t \left[u\left(c_{r,t}\right) - h\left(\frac{F_{N_r}(t)}{F_{N_n}(t)} l_{r,t}\right)\right]. \tag{111}$$

The planning problem is to maximize the utilitarian welfare function subject to the above incentive constraints (110) and (111) together with the resource constraints (105). As in the previous sections, I focus on the optimal allocations in which the incentive constraint of the non-routine worker binds and the incentive constraint of the routine worker slacks. For the next proposition, I introduce the following household-specific intratemporal and intertemporal wedges

$$\tau_{j,t}^{n} \equiv 1 - \frac{h'\left(l_{j,t}\right)}{u'\left(c_{j,t}\right)} \frac{1}{F_{N_j}(t)}, \quad \text{for } j \in \{n, r\} \tag{112}$$

and

$$\tau_{j,t+1}^{X_k} \equiv 1 - \frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} \frac{\phi_{k,t}}{F_{X_k}(t+1) + \phi_{k,t+1}\left(1 - \delta_x\right)}, \quad \text{for } j \in \{n, r\},\ k \in \{d, r\}. \tag{113}$$

The following proposition provides the dynamic counterpart of the static results with the second-best differential capital taxation.

Proposition 4. *Suppose that the incentive constraint for routine labor does not bind. Then, the optimal plan equates the intertemporal wedge of routine worker to that of non-routine worker,* $\tau_{r,t+1}^{X_k} = \tau_{n,t+1}^{X_k} = \tau_{t+1}^{X_k}$ *for* $k \in \{d, r\}$*. If routine labor hours are strictly positive* $(l_{r,t+1} > 0)$*, and if the planner can differentiate between two types of capital, then the intertemporal wedge is strictly positive for displacing capital,* $\tau_{t+1}^{X_d} > 0$ *and strictly negative for reinstating capital,* $\tau_{t+1}^{X_r} < 0$*.*

.5 Third Best Planning Problem

In this section, I consider a case where the planner cannot discriminate/distinguish between neither labor types nor capital types. Then, we have the following additional constraint in the planner's problem:

$$\tau_{r,t+1} = \tau_{d,t+1}. \tag{114}$$

Equivalently,

$$\frac{\phi_{r,t}}{F_{X_r}(t+1) + \phi_{r,t+1}\,(1-\delta_x)} = \frac{\phi_{d,t}}{F_{X_d}(t+1) + \phi_{d,t+1}\,(1-\delta_x)}. \tag{115}$$

With this, we still get a strictly positive uniform capital tax, as long as the production distortion by capital tax is not large enough. It is useful to define the following term:

$$M(t+1) = \frac{F_{X_dX_d}(t+1)[F_{X_r}(t+1) + (1-\delta_X)\phi_{d,t+1}] - F_{X_dX_r}(t+1)[F_{X_d}(t+1) + (1-\delta_X)\phi_{d,t+1}]}{F_{X_dX_r}(t+1)[F_{X_d}(t+1) + (1-\delta_X)\phi_{d,t+1}] - F_{X_rX_r}(t+1)[F_{X_d}(t+1) + (1-\delta_X)\phi_{d,t+1}]}\mu_{t+1},$$

where $F_{X_k X_k}$ indicates the second derivative of the production function with respect to capital type k, μ_{t+1} is the Lagrange multiplier for the resource constraint of period $t+1$. This is a term to capture the production distortion by a uniform capital taxation. Note that the cross-partial derivate captures the tax effects on production through complementarity/substitutability between two types of capital. Also, it is useful to define the following term for each capital type k:

$$H_k = \frac{\pi_n \eta_n}{\mu_{t+1}\phi_{d,t} X_{k,t+1}} h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1}\right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n}, X_k}(t+1),$$

where $\beta\eta_n$ is the Lagrangian multiplier for the incentive constraint of non-routine households. This captures the welfare gain/loss from the informational gain/loss by the uniform capital taxation. It is welfare gain for displacing capital while loss for reinstating capital.

Proposition 5. *Suppose that the incentive constraint for routine labor does not bind. And suppose $\tau_{t+1}^{X_r} = \tau_{t+1}^{X_d}$. Finally, suppose X_d is a perfect substitute to X_r. Then, $\tau_{t+1} > 0$ if and only if the following sum is strictly negative.*

$$\underbrace{H_d}_{\textit{Informational gain}} + \overbrace{M(t+1)}^{\textit{Production distortion}} \left(\underbrace{\frac{u'(c_{j,t})}{\beta u'(c_{j,t+1})} - \frac{F_{X_r}(t+1) + \phi_{r,t+1}(1-\delta_x)}{\phi_{r,t}}}_{\textit{Intertemporal wedge in } X_r} - \overbrace{H_r}^{\textit{Informational loss}} \right)$$

In particular, the following condition is a sufficient condition for a strictly positive uniform capital tax rate:

$$\frac{F_{X_d}(t+1)}{F_{X_r}(t+1)} \leq \frac{\phi_{d,t}}{\phi_{r,t}} \tag{116}$$

Notice that compared to the static case, we have additional intertemporal trade-off terms.

The sufficient condition above means that as long as the marginal rate of technological substitution next period is less than or equal to the price ratio today, the optimal uniform capital tax rate is strictly positive. In other words, bang tomorrow for the buck investment today

F Proof for Proposition 4

Proof. Recall that the planner's problem is to maximize

$$\sum_{t=0}^{\infty} \beta^t \left\{ \sum_{j=n,r} \omega_j \pi_j \left[u\left(c_{j,t}\right) - h\left(l_{j,t}\right) \right] + v\left(G_t\right) \right\} \tag{117}$$

subject to

$$\begin{aligned}
[\beta^t \mu_t] \quad \pi_r c_{r,t} + \pi_n c_{n,t} + G_t \leq & F\left(\pi_n l_{n,t}, \pi_r l_{r,t}, X_{d,t}, X_{r,t}\right) - \\
& \phi_{d,t}\left[X_{d,t+1} - \left(1-\delta_X\right) X_{d,t}\right] - \\
& \phi_{r,t}\left[X_{r,t+1} - \left(1-\delta_X\right) X_{r,t}\right], \ \forall\, t
\end{aligned} \tag{118}$$

$$[\eta_n] \quad \sum_{t=0}^{\infty} \beta^t \left[u\left(c_{n,t}\right) - h\left(l_{n,t}\right) \right] \geq \sum_{t=0}^{\infty} \beta^t \left[u\left(c_{r,t}\right) - h\left(\frac{F_{N_r}(t)}{F_{N_n}(t)} l_{r,t} \right) \right]. \tag{119}$$

Then, the first order conditions for $c_{n,t}$ and $c_{r,t}$ imply

$$\beta^t \omega_n u'\left(c_{n,t}\right) \left(1 + \frac{\eta_n}{\pi_n}\right) = \beta^t \mu_t, \tag{120}$$

$$\beta^t \omega_r u'\left(c_{r,t}\right) \left(1 - \frac{\eta_n}{\pi_r}\right) = \beta^t \mu_t. \tag{121}$$

Then, we get

$$\frac{u'(c_{n,t})}{\beta u'(c_{n,t+1})} = \frac{u'(c_{r,t})}{\beta u'(c_{r,t+1})} = \frac{\mu_t}{\beta \mu_{t+1}} \tag{122}$$

Note that the first order condition for $X_{d,t+1}$ is

$$\begin{aligned} \mu_t \phi_{d,t} =& \beta \mu_{t+1} \left[F_{X_d}(t+1) + \phi_{d,t+1}(1-\delta_x) \right] \\ & + \beta \frac{\pi_n \eta_n}{X_{d,t+1}} h' \left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1) \end{aligned} \tag{123}$$

Therefore,

$$\begin{aligned} \frac{u'(c_{n,t})}{\beta u'(c_{n,t+1})} =& \frac{u'(c_{r,t})}{\beta u'(c_{r,t+1})} = \frac{F_{X_d}(t+1) + \phi_{d,t+1}(1-\delta_x)}{\phi_{d,t}} + \\ & \frac{\pi_n \eta_n}{\mu_{t+1} \phi_{d,t} X_{d,t+1}} h' \left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1) \end{aligned} \tag{124}$$

Since $\mu_t > 0, \eta_n > 0$, and $\varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1) < 0$, as long as $l_{r,t+1} > 0$

$$\frac{u'(c_{r,t})}{\beta u'(c_{r,t+1})} = \frac{u'(c_{n,t})}{\beta u'(c_{n,t+1})} < \frac{F_{X_d}(t+1) + \phi_{d,t+1}(1-\delta_x)}{\phi_{d,t}} \tag{125}$$

On the other hand, the first order condition for $X_{r,t+1}$ is

$$\begin{aligned} \mu_t \phi_{r,t} =& \beta \mu_{t+1} \left[F_{X_r}(t+1) + \phi_{r,t+1}(1-\delta_x) \right] \\ & + \beta \frac{\pi_n \eta_n}{X_{r,t+1}} h' \left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1) \end{aligned} \tag{126}$$

Thus,

$$\frac{u'\left(c_{n,t}\right)}{\beta u'\left(c_{n,t+1}\right)} = \frac{u'\left(c_{r,t}\right)}{\beta u'\left(c_{r,t+1}\right)} = \frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}} + \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1) \tag{127}$$

Since $\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1) > 0$, as long as $l_{r,t+1} > 0$

$$\frac{u'\left(c_{r,t}\right)}{\beta u'\left(c_{r,t+1}\right)} = \frac{u'\left(c_{n,t}\right)}{\beta u'\left(c_{n,t+1}\right)} > \frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}. \tag{128}$$

So, if the planner can differentially tax the two types of capital, then the planner will tax displacing capital and subsidize reinstating capital in dynamic settings as well. ■

G Proof of Proposition 5

Proof. We know that due to the asymmetric information on capital types, we have the following constraint

$$\tau_{t+1}^{X_r} = \tau_{t+1}^{X_d}, \tag{129}$$

which we rewrite as

$$\frac{\phi_{d,t}}{\phi_{r,t}} = \frac{F_{X_d}(t+1) + \phi_{d,t+1}\left(1-\delta_x\right)}{F_{X_r}(t+1) + \phi_{r,t+1}\left(1-\delta_x\right)}. \tag{130}$$

I use $\beta^t \Omega_{t+1}$ as the Lagrange multiplier for this constraint. Recall from the proof for Propositio 4 that the first-order condition for $X_{d,t+1}$ without the uniform capital tax constraint (130) is that for $j \in \{n, r\}$,

$$\begin{aligned} \frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} =& \frac{F_{X_d}(t+1) + \phi_{d,t+1}\left(1-\delta_x\right)}{\phi_{d,t}} + \\ & \frac{\pi_n \eta_n}{\mu_{t+1}\phi_{d,t} X_{d,t+1}} h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1}\right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n}, X_d}(t+1) \end{aligned} \tag{131}$$

Then, the FOC with the constraint is

$$\begin{aligned} \frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} =& \frac{F_{X_d}(t+1) + \phi_{d,t+1}\left(1-\delta_x\right)}{\phi_{d,t}} + \\ & \frac{\pi_n \eta_n}{\mu_{t+1}\phi_{d,t} X_{d,t+1}} h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1}\right) \frac{F_{N_r}(t+1)}{F_{N_n}(t+1)} l_{r,t+1} \varepsilon_{F_{N_r}/F_{N_n}, X_d}(t+1) \\ & - \frac{\Omega_{t+1}}{\mu_{t+1}} \left(\frac{F_{X_d X_d}(t+1)[F_{X_r}(t+1) + (1-\delta_X)] - F_{X_d X_r}(t+1)[F_{X_d}(t+1) + (1-\delta_X)\phi_{d,t+1}]}{(F_{X_r}(t+1) + (1-\delta_X)\phi_{r,t+1})^2}\right) \end{aligned},$$

where Ω is the Lagrange multiplier for the uniform capital tax constraint (130). On the other hand, the first order condition for $X_{r,t+1}$ without the uniform tax constraint is

$$\begin{aligned}\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)}=&\frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}+\\&\frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)\end{aligned}\tag{132}$$

Similarly, the first-order condition for $X_{r,t+1}$ with the uniform tax constraint is

$$\begin{aligned}\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)}=&\frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}+\\&\frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)\\&-\frac{\Omega_{t+1}}{\mu_{t+1}}\left(-\frac{F_{X_rX_r}(t+1)[F_{X_d}(t+1)+(1-\delta_X)\phi_{d,t+1}]}{(F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1})^2}+\frac{F_{X_dX_r}(t+1)}{F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1}}\right).\end{aligned}$$

Then, solving for Ω, we get

$$\begin{aligned}\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)}-\frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}-&\\\frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)=&\\-\frac{\Omega_{t+1}}{\mu_{t+1}}\left(\frac{F_{X_dX_r}(t+1)[F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1}]-F_{X_rX_r}(t+1)[F_{X_d}(t+1)+(1-\delta_X)\phi_{d,t+1}]}{(F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1})^2}\right)&\end{aligned}$$

That is,

$$\Omega_{t+1}=\frac{\frac{u'(c_{j,t})}{\beta u'(c_{j,t+1})}-\frac{F_{X_r}(t+1)+\phi_{r,t+1}(1-\delta_x)}{\phi_{r,t}}-\frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)}{-\frac{1}{\mu_{t+1}}\left(\frac{F_{X_dX_r}(t+1)[F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1}]-F_{X_rX_r}(t+1)[F_{X_d}(t+1)+(1-\delta_X)\phi_{d,t+1}]}{(F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1})^2}\right)}.\tag{133}$$

Let

$$M(t+1)=\frac{F_{X_dX_d}(t+1)[F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1}]-F_{X_dX_r}(t+1)[F_{X_d}(t+1)+(1-\delta_X)\phi_{d,t+1}]}{F_{X_dX_r}(t+1)[F_{X_r}(t+1)+(1-\delta_X)\phi_{r,t+1}]-F_{X_rX_r}(t+1)[F_{X_d}(t+1)+(1-\delta_X)\phi_{d,t+1}]}.$$

Notice that the sign of the numerator is always opposite to that of the denominator unless it is zero. Thus, $M(t+1)$ is non-positive. Plug this back into the FOC of $X_{d,t+1}$, we get

$$\begin{aligned}\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} =& \frac{F_{X_d}(t+1)+\phi_{d,t+1}\left(1-\delta_x\right)}{\phi_{d,t}}+\\ & \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{d,t}X_{d,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1)\\ & +M(t+1)\Bigg(\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)}-\frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}-\\ & \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)\Bigg)\end{aligned}.$$

After some manipulation, we get

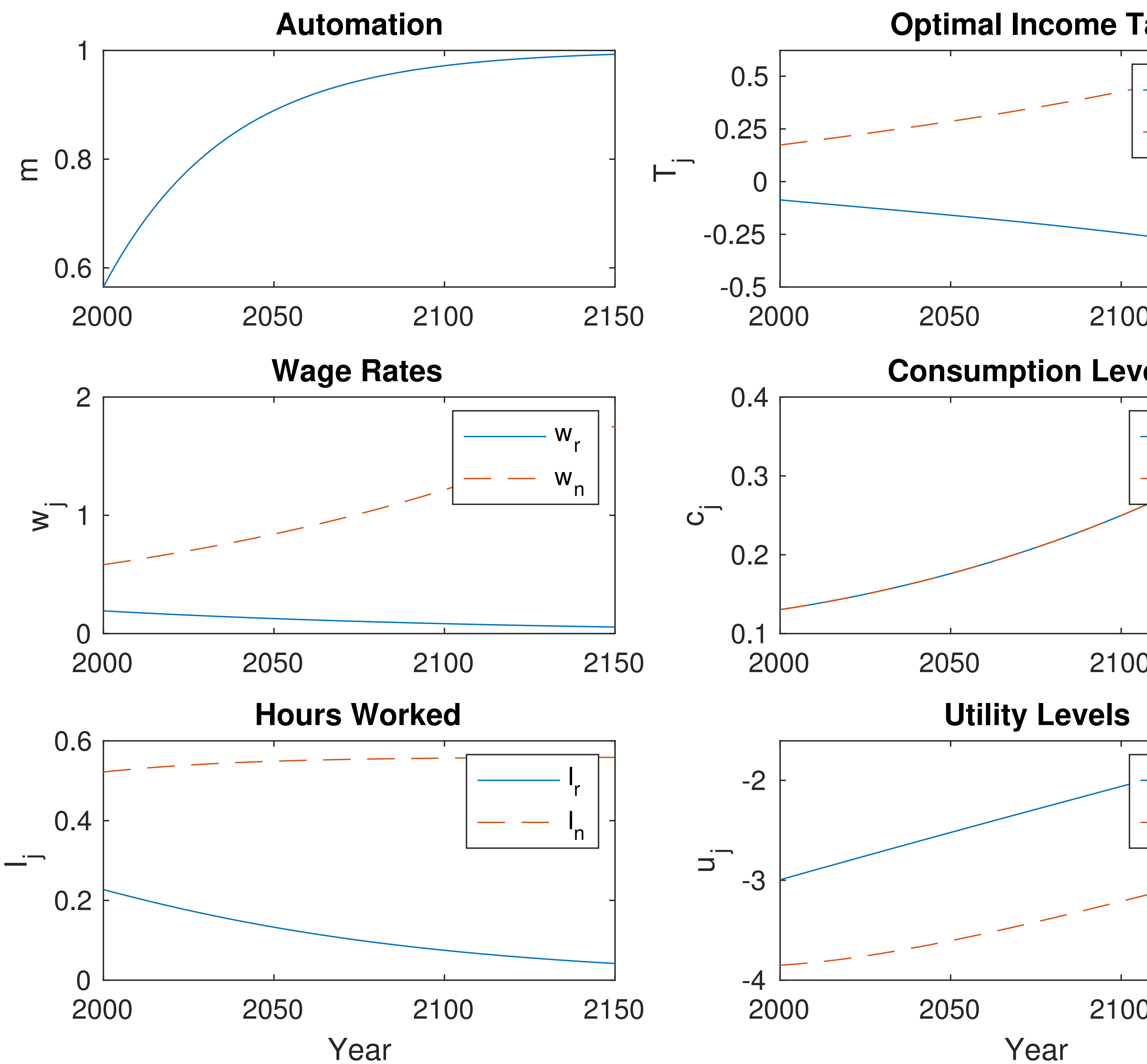
$$\begin{aligned}\frac{u'\left(c_{j,t}\right)}{\beta u'\left(c_{j,t+1}\right)} =& \frac{F_{X_d}(t+1)+\phi_{d,t+1}\left(1-\delta_x\right)}{\phi_{d,t}}+\\ & \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{d,t}X_{d,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1)\\ & +\frac{M(t+1)}{1-M(t+1)}\Bigg(\frac{F_{X_d}(t+1)+\phi_{d,t+1}\left(1-\delta_x\right)}{\phi_{d,t}}+\\ & \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{d,t}X_{d,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_d}(t+1)\\ & -\frac{F_{X_r}(t+1)+\phi_{r,t+1}\left(1-\delta_x\right)}{\phi_{r,t}}-\\ & \frac{\pi_n\eta_n}{\mu_{t+1}\phi_{r,t}X_{r,t+1}}h'\left(\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\right)\frac{F_{N_r}(t+1)}{F_{N_n}(t+1)}l_{r,t+1}\varepsilon_{F_{N_r}/F_{N_n},X_r}(t+1)\Bigg)\end{aligned}.$$

Notice that $-1 \leq \frac{M(t+1)}{1-M(t+1)} \leq 0$. This implies that the following condition is a sufficient condition for a strictly positive uniform capital tax rate:

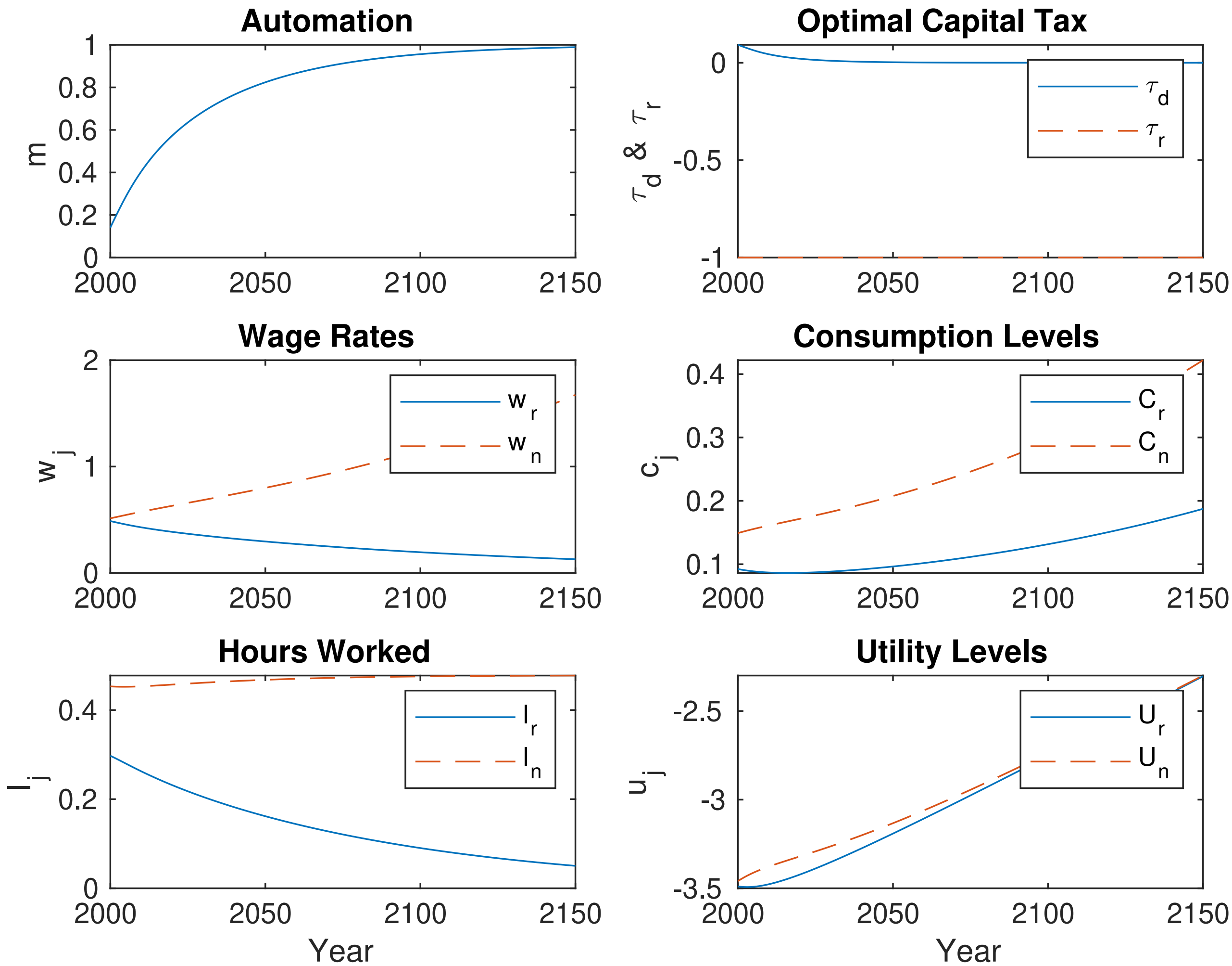
$$\frac{F_{X_d}(t+1)}{F_{X_r}(t+1)} \leq \frac{\phi_{d,t}}{\phi_{r,t}} \tag{134}$$

H Figures

A Figure 1: First-best Outcome

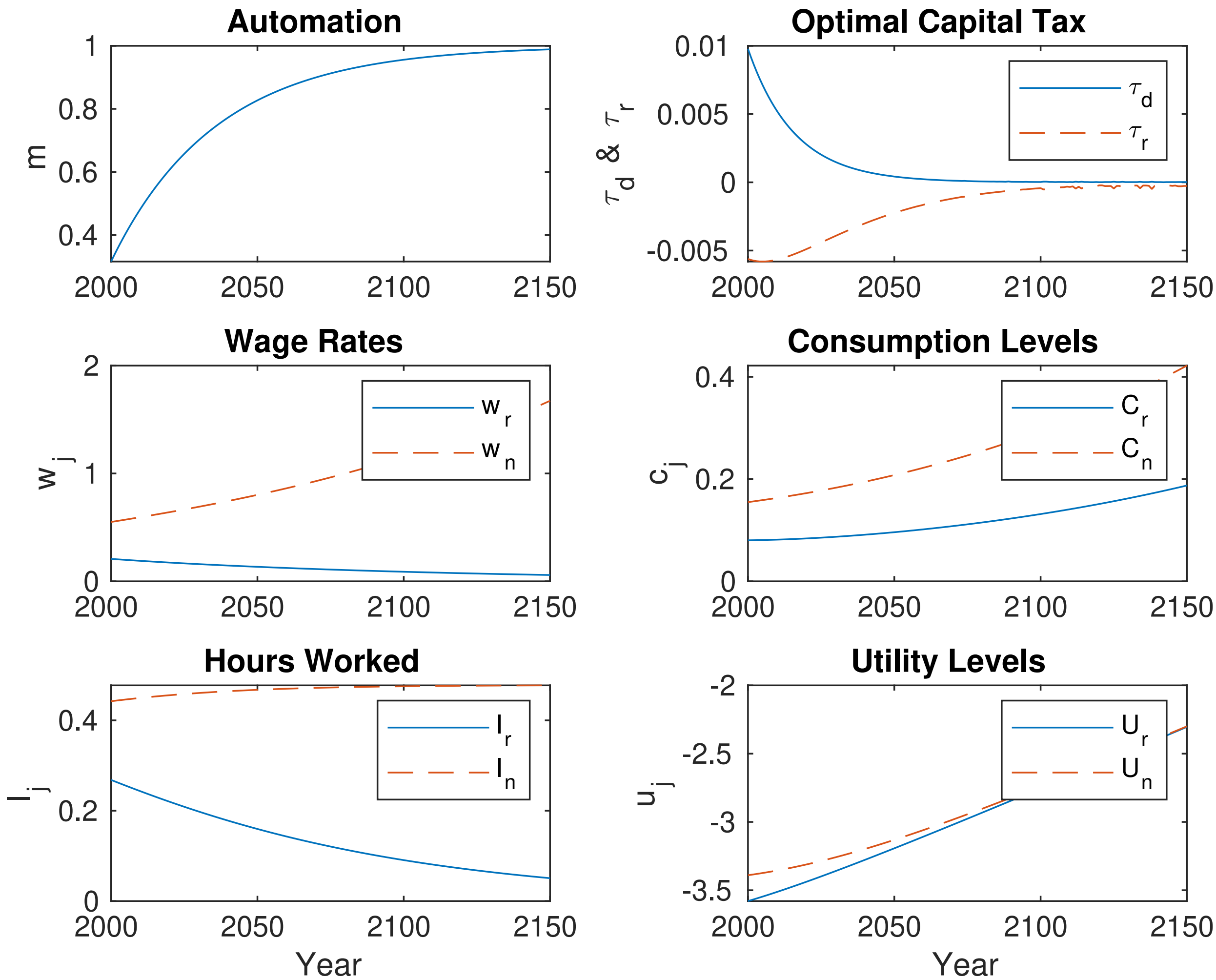


B Figure 2: Second-best Outcome—Leontief

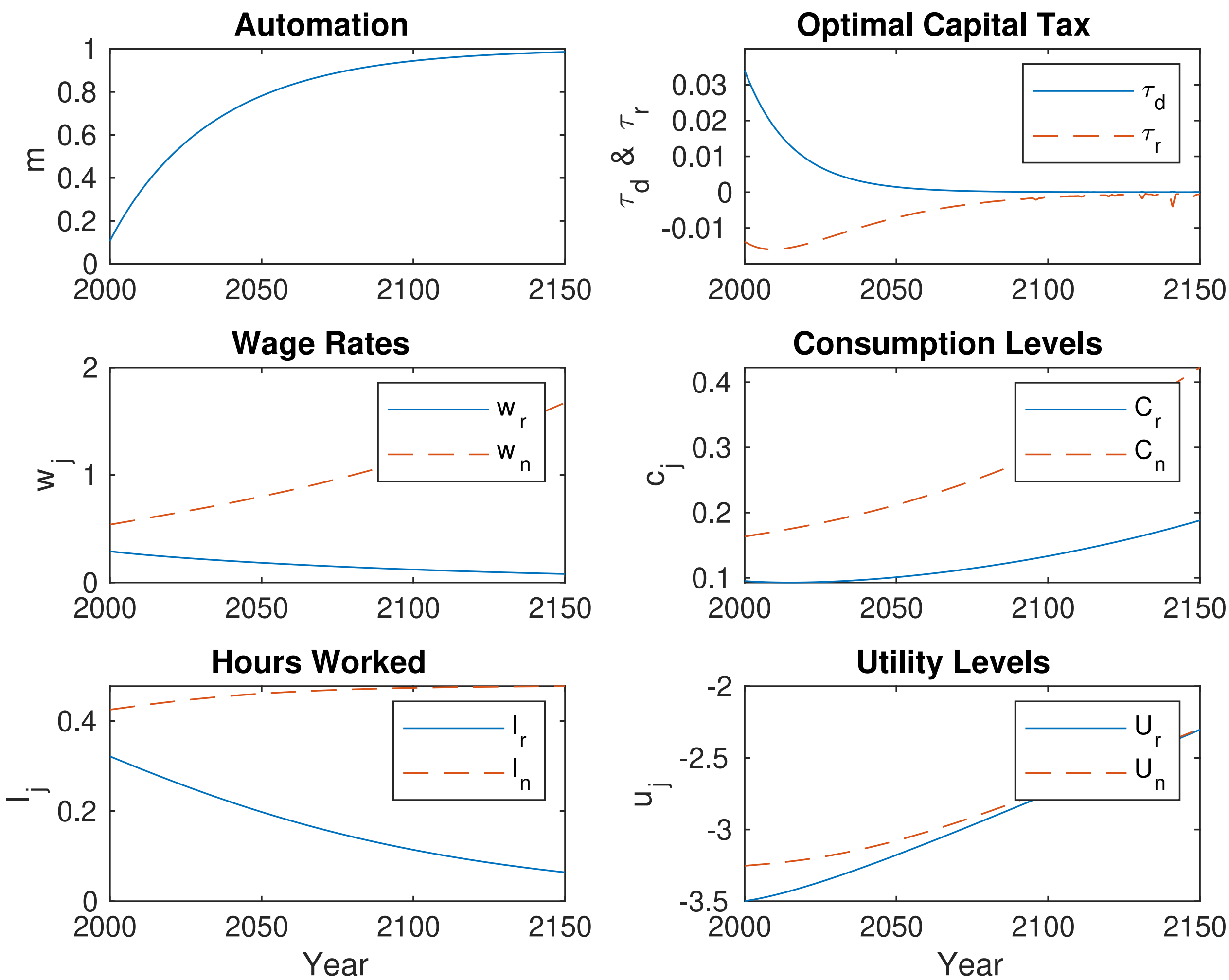


C Figure 3x: Second-best Outcome—Cobb-Douglas

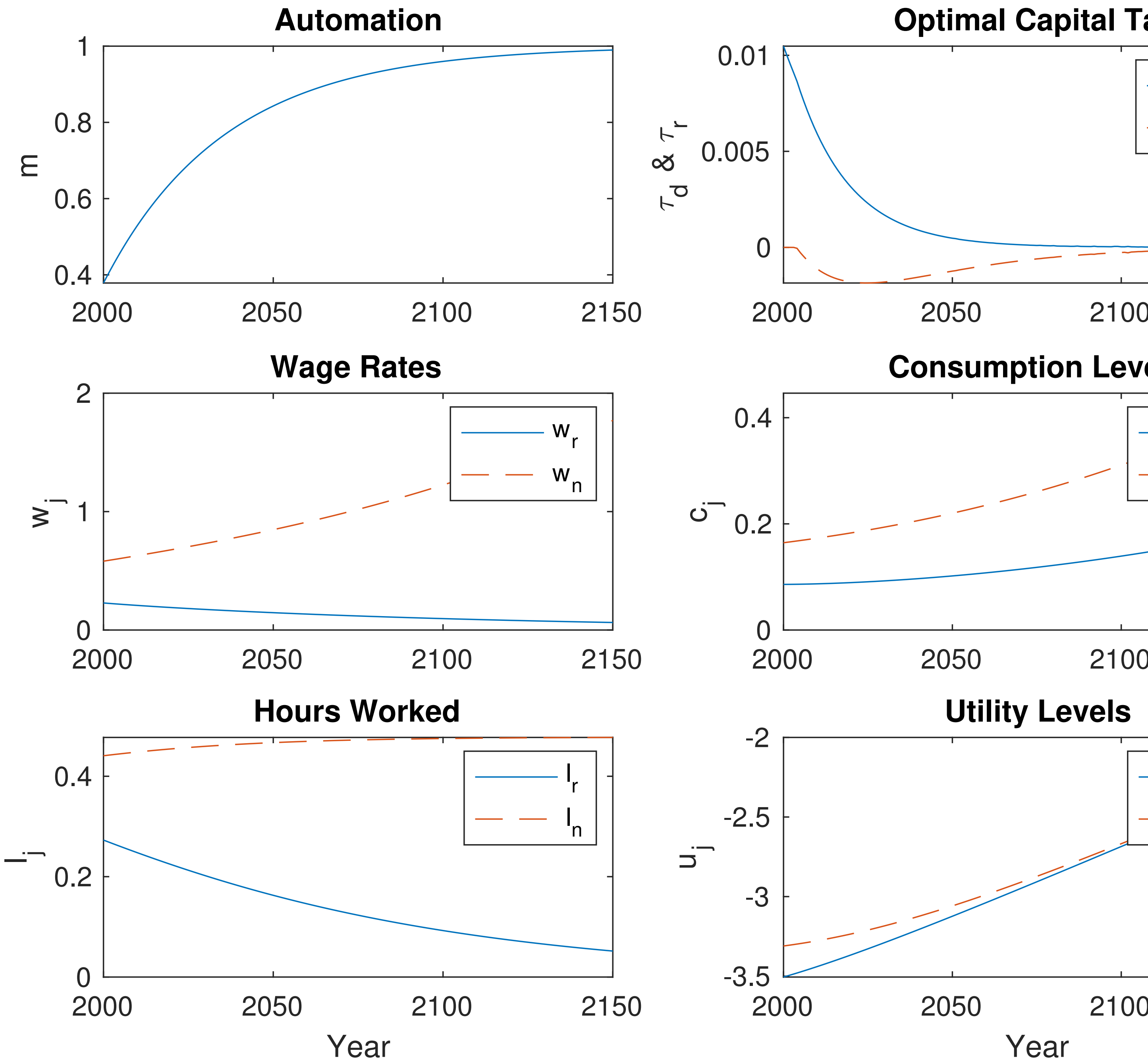
C.1 Figure 3a: $\beta = 0.5$



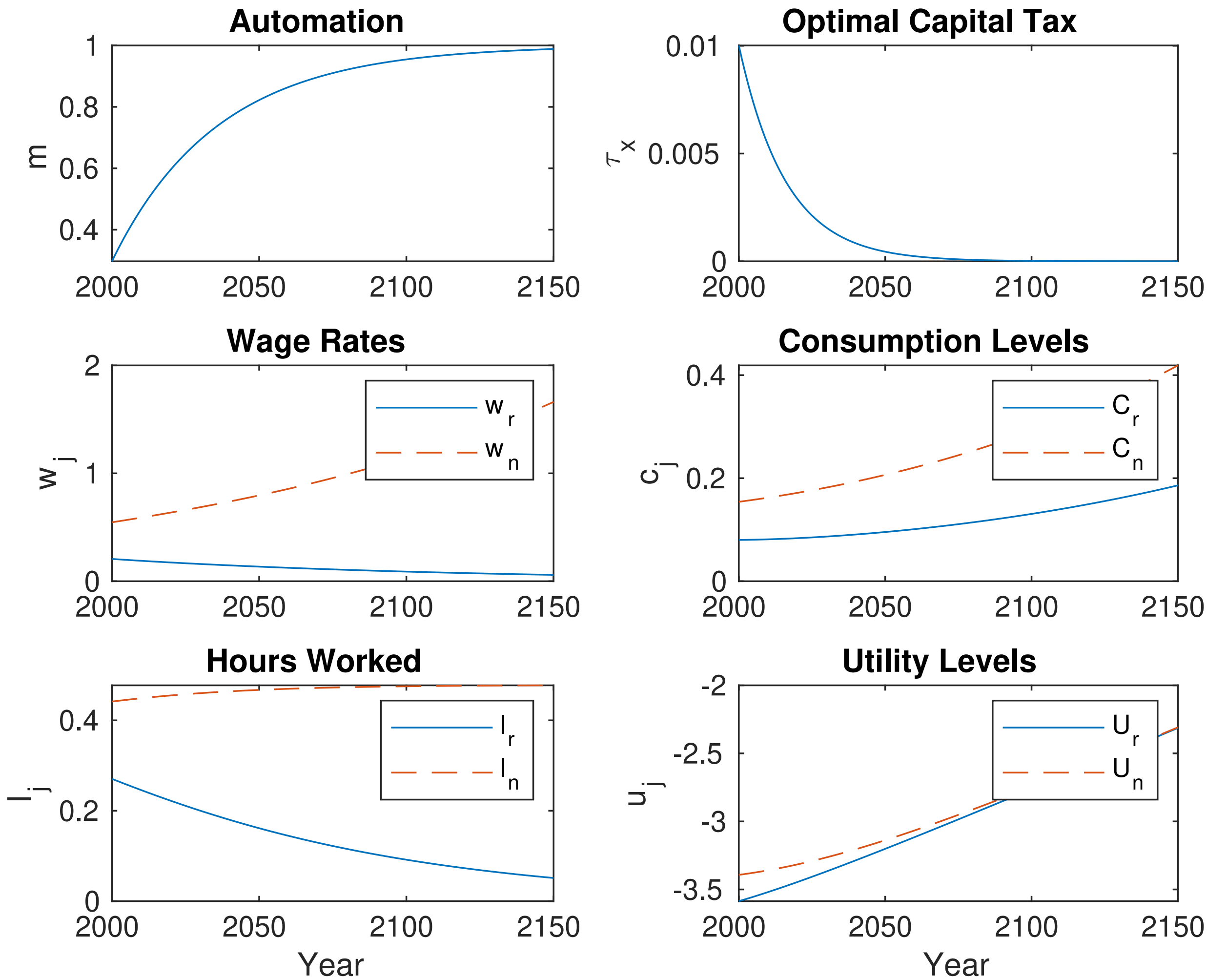
C.2 Figure 3b: $\beta = 0.2$



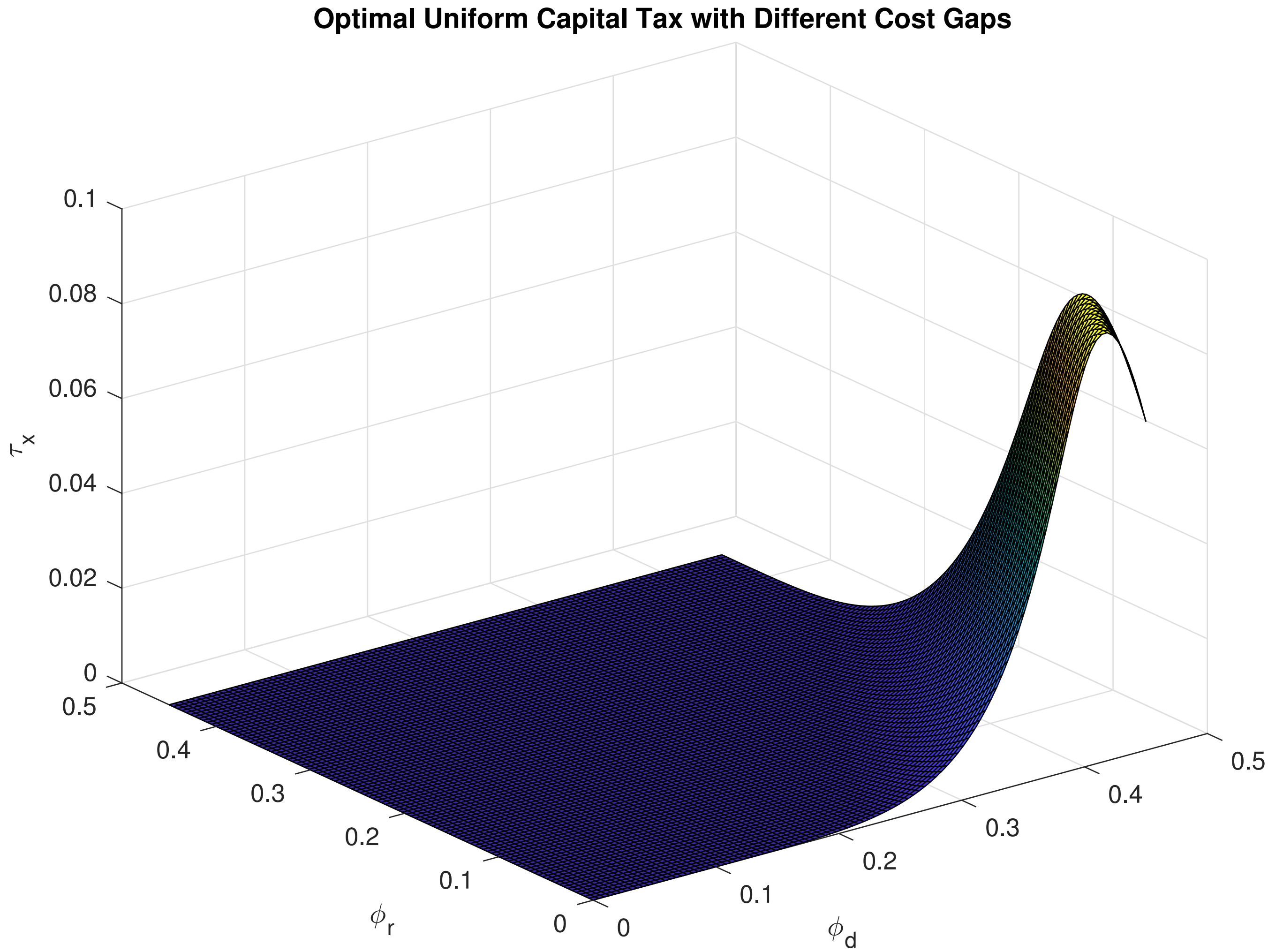
C.3 Figure 3c: $\beta = 0.7$



D Figure 4: Third-best Outcome–Leontief

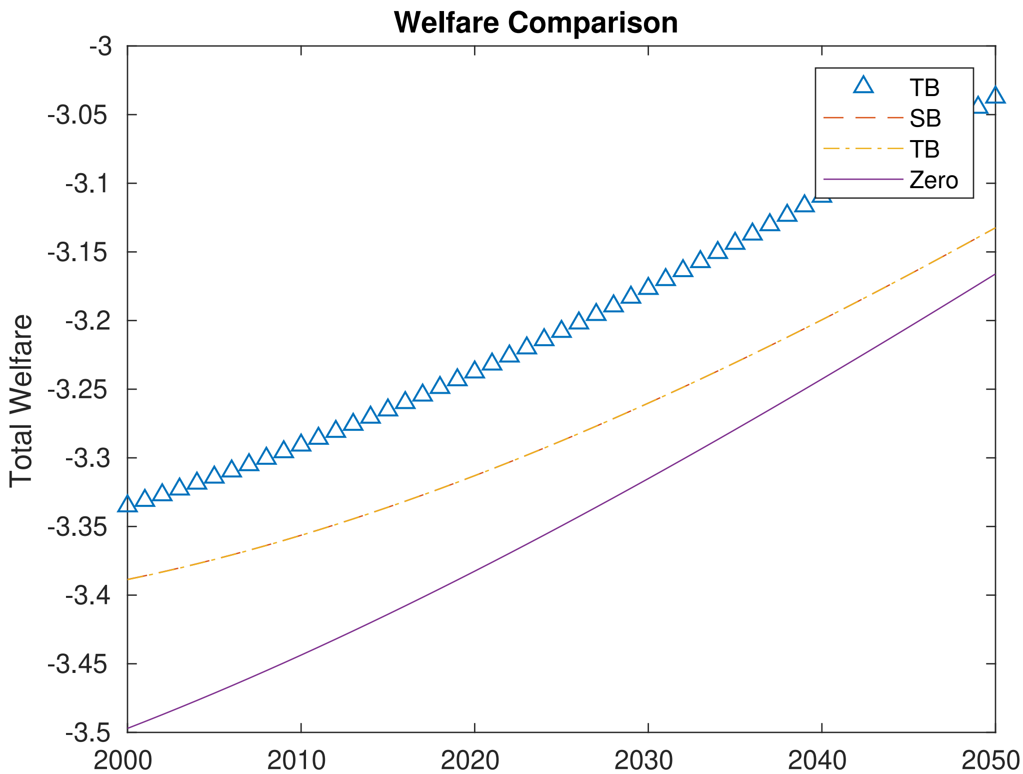


E Figure 5: Third-best Outcome–Relation between τ_x and Capital Cost Difference

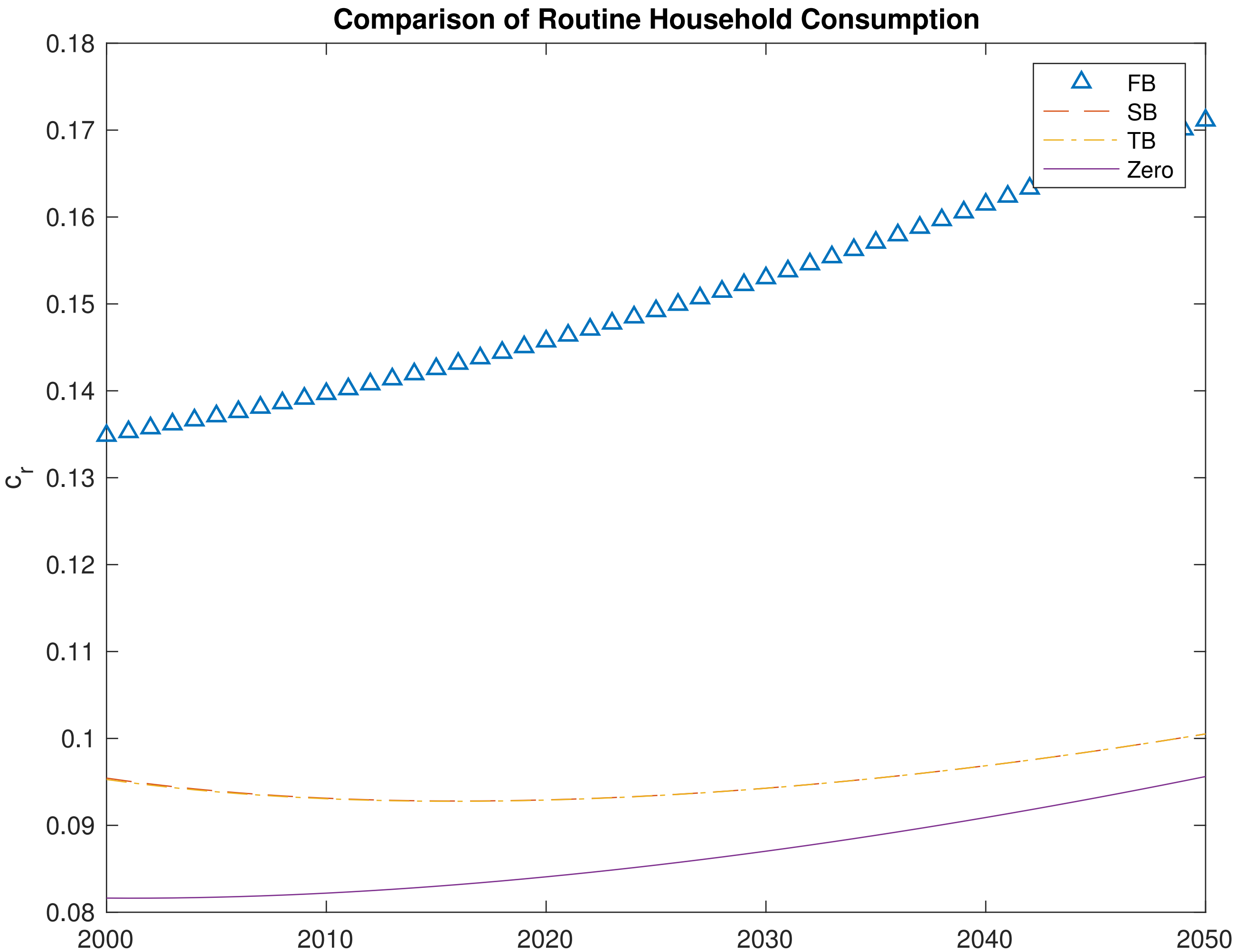


F Figure 6x: Welfare and Consumption Amount Comparison of First-best, Second-best, and Third-best Outcomes

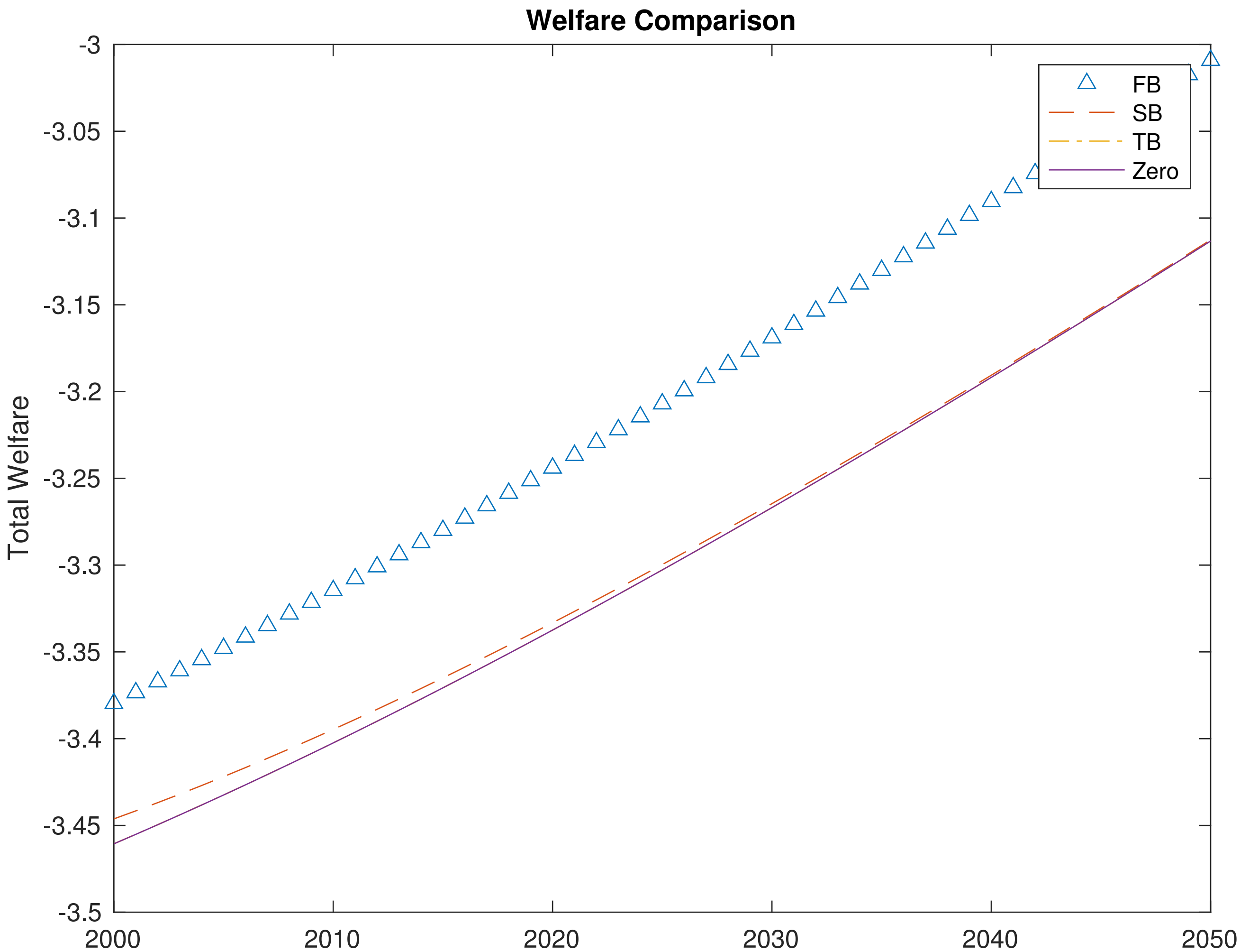
F.1 Figure 6a: Cobb-Douglas ($\beta = 0.2$)



F.2 Figure 6b: Routine Household Consumption—Cobb-Douglas ($\beta = 0.2$)



F.3 Figure 6c: Leontief



F.4 Figure 6c: Routine Household Consumption—Leontief