AUTOMATION, WAGE INEQUALITY AND OPTIMAL INCOME TAXATION *

Xiaoyong Cui   Guojun Lu   Wenjian Li

First Version: May, 2021
Last Update: February 28, 2024

Abstract

This paper explores the incidence of capital and labor income taxes and empirical statistics-based optimal taxation in a general equilibrium framework with endogenous automation. It first identifies two channels that contribute to wage inequality : the substitution effect and the automation effect. While capital deepening can partially alleviate wage inequality through the substitution effect, it simultaneously exacerbates wage inequality through the automation effect. Our theoretical analysis indicates that these two effects revise the conventional Mirrleesian optimal tax formula. The quantitative analysis demonstrates that the substitution effect leads to a more progressively optimal tax system, whereas the automation effect operates in the opposite direction. In the end, both the optimal capital and labor income tax rate are inverted U-shaped as a function of income.

Keywords. Automation technology, wage inequality, tax incidence, optimal income taxation.

*Xiaoyong Cui, Peking University, School of Economics, No.5 Yiheyuan Road, Haidian District, Beijing, P.R. China, cuixiaoyong@pku.edu.cn. Guojun Lu (corresponding author), Peking University, School of Economics, No.5 Yiheyuan Road, Haidian District, Beijing 100871, P.R. China, luguojun@pku.edu.cn. Wenjian Li, Zhejiang University, School of Economics, NO.866 Yuhangtang Road, Hangzhou 310058, P.R. China, wenjian_li@zju.edu.cn

1 Introduction

Over the past four decades, automation technology has been widely recognized as a key contributor to the increasing wage inequality in the United States (Acemoglu and Restrepo, 2022). In this scenario, the question of how to design an optimal redistributive policy becomes crucial from both practical and intellectual perspectives. The automation process involves that tasks originally produced by labor are displaced by capital. However, it remains unclear how to tax capital and labor income when automation technology is endogenously determined.

By integrating the insights from Mirrlees (1971) and Acemoglu and Restrepo (2022), we address the aforementioned question in a Mirrleesian economy that has been augmented to include changes in endogenously automated technologies. Specifically, we develop a tractable general equilibrium model that allows for the analysis of arbitrarily nonlinear income taxes. Following the techniques pioneered by Sachs et al. (2020), we study the incidence of capital and labor income tax reforms and derive a parsimonious characterization of optimal income taxation. The key and counter-intuitive finding of this paper shows that although automation technology exacerbates wage inequality, it also diminishes the progressivity of optimal income taxation.

We embed the endogenous assignment model(Costinot and Vogel, 2010; Ales et al., 2015) with the task-based framework (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018a) to microfound our production function. Concretely, there is a continuum of occupations to which skills are endogenously assigned, and each occupation involves a continuum of tasks that requires either capital or labor to be employed in order to complete it. In our baseline framework, both wages and automation technology are endogenously determined. Nevertheless, our model is tractable enough to decompose the different effects that contribute to equilibrium wages.

To emphasize the significance of automation technology, we initially disentangle two channels that demonstrate how factor inputs influence wages in equilibrium: the substitution effect and the automation effect. The substitution effect captures the impact of the substitutional or complementary relationship between factor inputs on factor prices while holding automation technology constant. In a model featuring two types of factor inputs, namely capital and labor, this effect encompasses both capital-skill complementarity(Krusell et al., 2000; Cui et al., 2021) and imperfect substitution between skills (Stiglitz, 1982; Sachs et al., 2020). By incorporating these two factors, our model facilitates the simultaneous analysis of optimal capital and labor income taxation within a general equilibrium framework.

Unlike the substitution effect, the automation effect implies that factor inputs lead to endogenous technological change in automation, which subsequently affects factor prices indirectly. Recent studies have examined the impact of automation on wages from both theoretical and empirical perspectives(Acemoglu and Loebbing, 2022; Acemoglu and Restrepo, 2022; Moll et al., 2022), which bolsters our confidence in distinguishing this effect from the others. By defining the demand-side elasticities, we find that capital deepening can increase wage levels through the substitution effect, which results from the enhancement of marginal productivity of labor. On the contrary, it can decrease wage levels through the automation effect, by reducing the demand for labor.

As the comparative advantage between capital and labor may vary across different occupations and tasks, it is important to note that changes in wage levels can conceal significant heterogeneity. We delve deeper into the implications of these two effects on the distribution of wages. Although capital deepening raises wages for all skill types, low-skilled workers benefit disproportionately due to the substitution effect. Conversely, capital deepening diminishes wages for all skill types through the automation effect, with low-skilled occupations bearing the brunt of the impact. Our findings emphasize that capital deepening can reduce wage inequality through the substitution effect while simultaneously exacerbating it through the automation effect.

Armed with the analyses above, we introduce tax reform for the discussion of tax incidence, in the same vein of Sachs et al. (2020). However, our work differs from theirs by implementing tax perturbations within the context of multidimensional taxation and endogenous automation. In our model, both capital and labor income tax reforms have significant implications for the relative supply of capital and labor, the adoption of automation technology, the distribution of wages, and ultimately the social welfare. We show that discussions on tax incidence would be incomplete without considering the impact of automated technological change.

Next, we extend the variational approach developed by Sachs et al. (2020) to solve optimal multidimensional taxation, i.e., optimal nonlinear capital and labor income taxes, in terms of sufficient statistics. In the appendix F, we also demonstrate the equivalence between mechanism design and variational approach in solving optimal multidimensional taxation. Compared to the existing literature that examines the optimal labor income tax problem using a task-to-talent assignment model(Ales et al., 2015; Sachs et al., 2020; Loebbing, 2020), our occupation-to-talent assignment model incorporates the production of tasks by capital. While automation commonly refers to the process in which capital replaces labor in the production of tasks, this framework enables us to analyze the implications of automated technological change for both optimal capital and labor income taxes.

Finally, we conduct a numerical analysis of optimal taxation and discover that both the substitution effect and automation effect make significant contributions to the optimal tax schemes. To begin, we calibrate the 2019 US economy using the Distributional National Accounts (DINAs) constructed by Piketty et al. (2018). Our findings reveal that occupations with higher labor income exhibit a lower level of automation. Subsequently, we conduct simulations to determine the optimal taxation under various scenarios, including the tax system with separable nonlinear labor and capital income taxation(NLIT-NCIT system), as well as scenarios where the capital income taxation is constrained to be linear (NLIT-LCIT system). Our analysis reveals that both optimal nonlinear capital and labor income taxes exhibit an inverted U-shaped pattern. Relative to the benchmark with exogenous wages(Mirrlees, 1971; Saez, 2001), the substitution effect contributes to a more progressive optimal tax system, while the automation effect counteracts this progressivity.

Our work is related to several streams of literature, one of which concerns is about automation technology. While some studies model automation using a task-based framework(Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018a), several others investigate the implications of automation technology for growth(Acemoglu and Restrepo, 2018b), employment (Acemoglu and Restrepo, 2020), labor share(Hémous and Olsen, 2022; Bergholt et al., 2022; Hubmer and Restrepo, 2021), and inequality (Moll et al., 2022; Acemoglu and Restrepo, 2022). However, few studies have explored the optimal government's policy response to this technology, except some recent works on robot tax (Costinot and Werning, 2018; Guerreiro et al., 2022; Thuemmel, 2023). While taxing the users of capital, such as robots or equipment, could regulate automation technology directly, the question remains as whether taxing the owners of capital, such as wealth or capital income, would be effective. This paper explores the design of optimal capital income tax in the context of automation technology.

Another relevant stream of literature pertains to capital taxation. Since the influential Chamley-Judd result that suggests capital should not be taxed in the long run (Judd, 1985; Chamley, 1986), the debate regarding the optimal design of capital taxes has persisted. Some studies introduce equipment-skill complementarity to argue that the optimal tax rate is not zero even in the steady state (Slavik and Yazici, 2014; Cui et al., 2021). Saez and Stantcheva (2018) argue that taxing capital income may be desirable if individuals derive utility from wealth. While we incorporate heterogeneity in wealth endowments to examine the motivations redistribution behind both linear and nonlinear capital income taxes, other studies consider various forms of heterogeneity to study savings or capital taxation, such as heterogeneous rates of return (Ferey et al., 2021; Gerritsen et al., 2020). Recent studies have explored the optimal taxation of multiple incomes, including the taxation of couples or different sources of income(Jacquet and Lehmann, 2021; Spiritus et al., 2022; Golosov and Krasikov, 2023). Our study serves as a valuable addition to this line of research.

The last related stream of literature is on tax incidence (Harberger, 1962) and optimal income taxation (Mirrlees, 1971). Extensive research on optimal income taxation has started since the seminal work of Mirrlees (1971). Some studies explore this basic theory methodologically in formulating policy recommendations (Saez, 2001; Diamond, 1998; Diamond and Saez, 2011). Others examine examine optimal income taxation in general equilibrium framework, such as taking into account the occupation choice (Rothschild and Scheuer, 2013), or introducing biased technical change(Ales et al., 2015; Loebbing, 2020). Our study aligns closely with Sachs et al. (2020), who have developed a comprehensive framework for analyzing tax incidence and optimal taxation. Building upon their work, we contribute to the existing literature by examining tax incidence and optimal taxation for both capital and labor income taxes within a general equilibrium framework.

We summarize the contributions of this paper as follows. Firstly, we build upon the research of Moll et al. (2022) and Acemoglu and Restrepo (2022) by advancing the discussion on optimal redistributive policy design. Specifically, we investigate the optimal combination of capital and labor income taxation, recognizing the implications of automation for wage or income inequality. Moreover, we calibrate the US economy to the year 2019 and undertake simulations to analyze the optimal dual-tax system across various scenarios. Through this analysis, we gain valuable insights into how tax policies should adapt and respond to the impact of automation technology. From this perspective, our research carries significant policy implications as it provides policymakers with valuable guidance on designing effective tax policies that address challenges arising from automation.

By incorporating the interaction between capital and labor in the production function within a general equilibrium framework, our study extends the simple theory of optimal capital taxation developed by Saez and Stantcheva (2018), who has derived optimal tax formulas in a partial equilibrium setting. Our expression for optimal capital income taxation incorporates an additional general equilibrium term that is linked to the optimal labor income taxation. This term arises due to the incidence of capital income tax reform on the distribution of wages. Cui et al. (2021) has captured this general equilibrium term by introducing capital-skill complementarity, while our study complements their work by demonstrating that the replacement of labor by capital in production tasks, specifically through automation technology, also contributes to the general equilibrium term. Moreover, our results from numerical simulations provide strong evidence that the automation effect is highly consequential and should not be overlooked in the analysis of optimal income taxation.

Finally, we contributes to the insightful work of Sachs et al. (2020) in two key aspects. First, we generalize their approach by extending their unidimensional model to a multidimensional one. Considering that individuals receive both capital income and labor income concurrently, our analysis of tax incidence and optimal taxation becomes a multidimensional problem. Secondly, we decompose the general equilibrium effect in their study into substitution effect and automation effect. From this perspective, our research can be viewed as an application of their theoretical framework to the scenario of automation technology.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 focuses on the impact of automation technology on wages. Tax incidence analysis is implemented in Section 4. Section 5 lays out the optimal income tax system. Finally, Section 6 presents the quantitative results and Section 7 concludes.

2 The Model

2.1 Economic Environment

Individuals.–The economy is populated by a unit mass of individuals differing in their skills, indexed by n, and wealth, indexed by q. Accordingly, individuals have a two-dimensional type $(n, q) \in N \times Q$ with $N = [\underline{n}, \bar{n}]$ and $Q = [\underline{q}, \bar{q}]$, following a joint distribution $F(n, q)$ with density $f(n, q)$. Individuals live for two periods. In the first period, he decides to save a_q unit assets out of his initial wealth y_q, and consume the rest to derive utility $u(y_q - a_q)$. In the next period, he supplies l_n unit labor with wage w_n, thus obtaining labor income $z_n = w_n l_n$. Moreover, his assets can be rented in capital market with rental rate R and he can receive capital income $x_q = R a_q$. In this period, he consumes all his after-tax income. Individuals with a given type (n, q) maximize their utility subject to the budget constraint:

$$\begin{aligned} \max_{a_q, l_n} \quad & U(n, q) \equiv u(y_q - a_q) + c(n, q) - v(l_n) \\ \text{s.t.} \quad & c(n, q) = w_n l_n + (1 + R) a_q - T(w_n l_n, R a_q) \end{aligned} \tag{1}$$

Where $c(n, q)$ denotes his consumption in the second period. $T(\cdot)$ is a twice continuously differentiable income tax function implemented by the government. $-v(l_n)$ is the disutility of labor. We assume that there is no depreciation in assets a_q. Moreover, $u(\cdot)$ and $v(\cdot)$ is twice continuously differentiable with $u'(\cdot) > 0, u''(\cdot) < 0$ and $v'(\cdot) > 0, v''(\cdot) > 0$.

Technology.– There is one final good, which is produced by a continuum occupational output Y_θ, where $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$ denotes the occupation. Output of the final good is given by the following constant elasticity of substitution (CES) production function:

$$Y = \left\{ \int_{\underline{\theta}}^{\bar{\theta}} \beta_\theta Y_\theta^\rho d\theta \right\}^{1/\rho}.$$

Where β_θ is a distributional parameter. In each occupation, there is a unit continuum of tasks $i \in [0,1]$, some of which may be automated by capital.[1] The occupational output is produced according to a Cobb-Douglas production function:

$$\ln Y_\theta = \int_0^1 \ln Y_\theta(i) di,$$

where tasks can be produced using different-skill labor or capital as follows:

$$Y_\theta(i) = \psi^k(\theta, i) K(\theta, i) + \int_{\underline{n}}^{\bar{n}} \psi^l(n, \theta, i) L(n, \theta, i) dn$$

where $\psi^k(\theta, i)$ and $\psi^l(n, \theta, i)$ denote, respectively, the productivity of capital and n-type labor in task i of occupation θ. We assume that $\psi^l(n, \theta, i)$ is twice differentiable and strictly log-supermodular:

$$\psi^l(n', \theta', i)\psi^l(n, \theta, i) > \psi^l(n, \theta', i)\psi(n', \theta, i) \quad \forall n' > n, \theta' > \theta, i \in [0,1].$$

Following Costinot and Vogel (2010), this property ensures that there exists a continuous and strictly increasing matching function $\theta(n)$ such that n can be mapped to θ. Intuitively, individuals with skill n choose occupation based on their comparative advantage. Moreover, we order $\psi^l(n, \theta, i)/\psi^k(\theta, i)$ is strictly increasing with i to ensure there exists a threshold task α_θ such that tasks in $[0, \alpha_\theta]$ are produced with capital and tasks in $(\alpha_\theta, 1]$ are produced with labor. Intuitively, labor has more comparative advantage over capital in more complicated tasks (higher index i). Given these assumptions on factor's productivity, the production of tasks could be reduced as follows:

$$Y_{\theta(n)}(i) = \begin{cases} \psi^k(\theta(n), i) K(\theta(n), i) & \text{if } i \in \left[0, \alpha_{\theta(n)}\right] \\ \psi^l(n, \theta(n), i) L(n, \theta(n), i) & \text{if } i \in \left(\alpha_{\theta(n)}, 1\right] \end{cases}$$

[1]Taking blue-collar workers and teachers for example, worker's tasks involve carrying, building, driving, assembling etc, while the duty of teacher is teaching, writhing and coming up with new idea etc. Intuitively, tasks of the latter are less likely to be automated by capital.

In what follows, we do not distinguish skill type n and occupation type θ due to their monotone mapping relationship. To make our notations clear and brief, we reduce the index of occupation θ and hold the following equations through out the paper,

$$\psi_n^l(i) \equiv \psi^l(n, \theta(n), i), \quad \psi_n^k(i) \equiv \psi^k(\theta(n), i), \quad K_n(i) \equiv K(\theta(n), i), \quad L_n(i) \equiv L(n, \theta(n), i).$$

Then, the output and automation technology in occupation $\theta(n)$, i.e., $Y_{\theta(n)}$ and $\alpha_{\theta(n)}$ can be reduced to Y_n and α_n naturally.

Government.–The government wishes to raise a given amount of revenue for government expenditure B, but only has access to the instruments of capital and labor income taxes on individuals.

$$B = \int_{\underline{n}}^{\bar{n}} \int_{\underline{q}}^{\bar{q}} T(z_n, x_q) f(n, q) dq dn. \tag{2}$$

Equilibrium.–An equilibrium of the model is given by a tax function T, a collection of quantities and prices, a set of automation technology, aggregate output and capital, such that:

- Capital and labor, $\{K_n(i), L_n(i)\}_{n\in N, i\in[0,1]}$, automation technology, $\{\alpha_n\}_{n\in N}$, are allocated in a profit maximizing way to produce output Y given factor price $\{w_n\}_{n\in N}$, R.
- Capital and labor supply, $\{a_q\}_{q\in Q}$, $\{l_n\}_{n\in N}$, maximize individuals' utility $\{U(n,q)\}_{n\in N, q\in Q}$.
- Labor and capital markets clear[2]:

$$l_n = L_n = \int_{\alpha_n}^{1} L_n(i) di \quad \forall n \in N, \quad K = \int_{n\in N} \int_0^{\alpha_n} K_n(i) di dn = \int_{n\in N} \int_{q\in Q} a_q f(n, q) dq dn$$

2.2 Macroeconomic Aggregates

To characterize the equilibrium variables in the way of a parsimonious set of equations, we first introduce the aggregate results.

Lemma 1 ***(Equilibrium output)*** *Suppose $\psi_n^l(i)/\psi_n^k(i)$ is strictly increasing with i for all $n \in N$, and the interior point solution for the level of automation $\{\alpha_n\}_{n\in N}$ exists, then the skill output could be reduced as a Cobb-Douglas production function:*

$$Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n} \tag{3}$$

[2] Good markets is naturally clearing, $\int_{n\in N} \int_{q\in Q} c(n,q) f(n,q) dq dn + B = K + Y$, due to the Walras's law.

where

$$A_n(\alpha_n) = \frac{e^{\int_0^{\alpha_n} \ln \psi_n^k(i)di + \int_{\alpha_n}^1 \ln \psi_n^l(i)di}}{\alpha_n^{\alpha_n}(1-\alpha_n)^{1-\alpha_n}}$$

denotes the total factor productivity in the workplace of skill-type n. Denote $\mathcal{L} \equiv \{L_n\}_{n\in N}$ *and* $\boldsymbol{\alpha} \equiv \{\alpha_n\}_{n\in N}$*, the aggregate output is given by a CES production function:*

$$Y \equiv F(K, \mathcal{L}; \boldsymbol{\alpha}) = \left\{ \int_{\underline{n}}^{\bar{n}} \beta_n \left[\tilde{A}_n(\alpha_n) K^{\alpha_n} L_n^{1-\alpha_n} \right]^\rho dn \right\}^{1/\rho} \tag{4}$$

where $\tilde{A}_n(\alpha_n) = A_n(\alpha_n)\phi_n^{\alpha_n}(\alpha_n)$*, and* $\phi_n(\alpha_n)$ *the share of capital allocated to skill-type n out of the total capital.*

Proof. See Appendix A.1. ■

There are two notes about Lemma 1. First, As the occupational output is in the form of CD production function, the degree of automation coincides with the capital share. However, both of them are no longer exogenous, but adjust with the automated technological change. The intuition behind is that, for the workplace where capital gains more importance (larger α_n), more tasks of individuals with skill-type n are automated. In addition, reassigning tasks between capital and labor could also promote productivity, captured by the improvement of $A_n(\alpha_n)$. Second, with no distinction on capital that cooperated with different skill labor, the aggregate output Y can be expressed as a function of aggregate capital K, a set of labor input $\mathcal{L}$ and automation technology $\boldsymbol{\alpha}$. As we will see, this parsimonious form is convenient for our tax analysis.

Definition 1 *Denote* $\mu_n(i) = \frac{\psi_n^l(i)}{\psi_n^k(i)} = \delta_n \cdot i^\eta$ *as the comparative advantage of labor with skill-type n over capital in task i, where* $\dot{\delta}_n > 0, \eta > 0$*.*

In the context of automation technology, defining the comparative advantages of capital and labor in tasks is necessary. The definition above shows that the more complex the task (higher i), the more comparative advantage of producing with labor than capital, i.e., $\mu_n(i)$ increases with i for any skill-type n. In addition, relative to low-skilled labor, high-skilled individuals have more comparative advantages over capital for a given task, i.e., $\mu_n(i)$ increases with n for any tasks i. We will calibrate parameters δ_n and η in the quantitative analysis. Armed with this preliminary work, now we turn to the equilibrium factor prices and degree of automation.

Lemma 2 ***(Factor prices and automation)*** *With the price of aggregate output normalized to one, in* *equilibrium, wages and rental rate can be given as follows:*

$$w_n \equiv w_n(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{(1-\alpha_n)\gamma_n Y}{L_n}, \quad R \equiv R(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{\alpha Y}{K}, \quad \forall n \in N. \tag{5}$$

where $\gamma_n = p_n Y_n / Y$ *denotes the the share of output value produced by skill-type* n *in the total output value, and with* $\int_{\underline{n}}^{\bar{n}} \gamma_n dn = 1$.
Denote $\alpha = \int_{\underline{n}}^{\bar{n}} \gamma_n \alpha_n dn$ *as the average degree of automation in the economy, the equilibrium automation technology is the solution of the following equations:*

$$\alpha_n \equiv \alpha_n(K, \mathcal{L}) = 1 - \frac{1}{\gamma_n} \frac{\mu_n(\alpha_n) L_n}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n) L_n dn}, \quad \alpha \equiv \alpha(K, \mathcal{L}) = \frac{K}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n) L_n dn}, \quad \forall n \in N. \tag{6}$$

Proof. See Appendix A.2. ■

Equation (5) in Lemma 2 discloses two channels that affect equilibrium factor prices. The first is capital-skill complementarity and imperfect substitution between skills, which has been discussed by Slavik and Yazici (2014) and Sachs et al. (2020) respectively. In this paper, we generalize them as being the *substitution effect*, for it captures the substitution relationship between all factor inputs. The second channel, the adjustment of automation technology, called *automation effect*, is captured by the new arguments $\boldsymbol{\alpha}$.[3] As we will see, these two effects compose the workhorse of this paper. The adoption of automation technology is captured by equation (6), which is endogenous with respect to capital and labor input $\{K, \mathcal{L}\}$. Intuitively, the more capital in the economy, the higher extent of automation. While there is no difference in the using of capital or labor in task α_n, i.e., $\mu_n(\alpha_n) = \psi_n^l(\alpha_n)/\psi_n^k(\alpha_n) = w_n/R$, the equilibrium degree of automation in skill-type n, α_n, coincides with the capital share RK_n/p_nY_n, and the average degree of automation in economy, α, coincides with the aggregate capital share RK/Y. Thus, the rise in automation technology goes hand in hand with the decline in labor share. Recent studies have found that the decline in aggregate labor share could mainly be attributed to the process of automation(Hémous and Olsen, 2022; Bergholt et al., 2022).

2.3 Definition of Elasticities

Before we move on to the analysis in the following sections, two bunches of elasticities are defined here for convenience. As will be shown later, these elasticities play a central role in our

[3] Acemoglu and Restrepo (2018a) isolates displace effect, captured by $1 - \alpha_n$, and productivity effect, captured by $\gamma_n Y$, that automated technological change has impact on wage. In this paper, we do not distinguish these two effects, but make automation technology endogenized.

decomposition of mechanisms.

The first set of elasticities is about *supply-side elasticities*. These elasticities are defined to capture individuals' behavior of factor supply. To simplify our analysis, we restrict the tax function to be separable, i.e., $T(z_n, x_q) = T_z(z_n) + T_x(x_q)$, [4] which means that the government levies nonlinear tax on labor and capital income separately. Following the first-order conditions of individuals, the behavior of labor supply depends on marginal labor income retention rate and wage, $l_n(1 - T_z'(z_n), w_n)$. Symmetrically, the behavior of capital supply depends on marginal capital income retention rate and rental rate, $a_q(1 - T_x'(x_q), R)$. Following the standard definition of elasticity, we denote the elasticity of labor supply with respect to wage and the elasticity of capital supply with respect to rental rate as follows,

$$\epsilon_{l_n,w_n} = \frac{w_n}{l_n}\frac{dl_n}{dw_n}, \quad \epsilon_{a_q,R} = \frac{R}{a_q}\frac{da_q}{dR}.$$

To define the behavioral elasticity for nonlinear tax, we consider a tax perturbation in the same vein of Gerritsen (2016),

$$\tilde{T}_i = T_i + \kappa_i \tau_i \quad \text{with} \quad i \in [z, x]. \tag{7}$$

where κ_i denotes the reform parameter, and τ_i is the reform function of income tax. τ_i is assumed to be twice differentiable in taxable income. A marginal reform of the income tax can be studies by considering a change $d\kappa_i$ as $\kappa_i \rightarrow 0$. For a given taxable income $i \in [z, x]$, such reform can both raise the tax burden by $\tau_i d\kappa_i$ and the marginal tax rate by $\tau_i' d\kappa_i$. We define the elasticity of labor supply with respect to the marginal retention rate as

$$\epsilon_{l_n,1-T_z'} = -\frac{1 - T_z'(z_n)}{l_n}\frac{dl_n}{\tau_z'(z_n)d\kappa_z}|_{\kappa_z=0}$$

which is the relative change in labor supply, dl_n/l_n, due to a relative change in the marginal retention rate, $-\tau_z'(z_n)d\kappa_z/(1 - T_z(z_n))$. Similarly, the elasticity of capital supply with respect to the marginal retention rate can be defined as

$$\epsilon_{a_q,1-T_x'} = -\frac{1 - T_x'(x_q)}{a_q}\frac{da_q}{\tau_x'(x_q)d\kappa_x}|_{\kappa_x=0}.$$

We summarize these supply-side elasticities (behavioral elasticities) in Table 1.

[4] This kind of tax system has been discussed by Gerritsen et al. (2020),Jacquet and Lehmann (2021),Ferey et al. (2021). With a comprehensive income tax system $T(z_n, x_q)$, we just need to define some cross-elasticity, e.g., labor supply w.r.t capital tax rate, this is beyond the scope of this article.

Table 1: Supply-side elasticities

Marginal retention rate elasticity of labor supply	$\epsilon_{l_n,1-T'_z}$	$\frac{[1-T'_z(z_n)]e_{l_n,1-t_z}}{1-T'_z(z_n)+e_{l_n,1-t_z}T''_z(z_n)z_n}$
Marignal retention rate elasticity of capital supply	$\epsilon_{a_q,1-T'_x}$	$\frac{\left[1-T'_x(x_q)\right]e_{a_q,1-t_x}}{1-T'_x(x_q)+e_{a_q,1-t_x}T''_x(x_q)x_q}$
Wage elasticity of labor supply	ϵ_{l_n,w_n}	$\frac{[1-T'_z(z_n)-T''_z(z_n)z_n]e_{l_n,1-t_z}}{1-T'_z(z_n)+e_{l_n,1-t_z}T''_z(z_n)z_n}$
Rental rate elasticity of capital supply	$\epsilon_{a_q,R}$	$\frac{\left[1-T'_x(x_q)-T''_x(x_q)x_q\right]e_{a_q,1-t_x}}{1-T'_x(x_q)+e_{a_q,1-t_x}T''_x(x_q)x_q}$

Proof. See Appendix B.1. ■

Where $e_{l_n,1-t_z} \equiv -\frac{1-t_z}{l_n}\frac{dl_n}{dt_z} = \frac{v'(l_n)}{l_n v''(l_n)}$ and $e_{a_q,1-t_x} \equiv -\frac{1-t_x}{a_q}\frac{da_q}{dt_x} = -\frac{u'(y_q-a_q)-1}{a_q u''(y_q-a_q)}$ represent the marginal retention rate elasticities of labor and capital respectively, in the context of linear labor and capital income taxes t_z and t_x. When $T''_z(z_n) = T''_x(x_q) = 0$, $\epsilon_{l_n,1-T'_z}$ and $\epsilon_{a_q,1-T'_x}$ collapse to $e_{l_n,1-t_z}$ and $e_{a_q,1-t_x}$, moreover, we have $\epsilon_{l_n,1-T'_z} = \epsilon_{l_n,w_n}$ and $\epsilon_{a_q,1-T'_x} = \epsilon_{a_q,R}$.

Now we turn to the *demand-side elasticities*, which is defined according to the first-order conditions of producer. From Lemma 2, we know that equilibrium factor prices are determined by factor inputs and the equilibrium automation technology,

$$w_n \equiv w_n(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{(1-\alpha_n)\gamma_n Y}{L_n}, \quad R \equiv R(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{\alpha Y}{K}.$$

Note that $Y \equiv F(K, \mathcal{L}; \boldsymbol{\alpha})$. It is clear that when the automation effect is shut down, i.e., holding $\boldsymbol{\alpha}$ unchanged, factor inputs can still affect equilibrium price via substitution effect, which can be captured by the following definition of elasticity,

$$\epsilon_{w_n,L_{n'}} = \frac{L_{n'}}{w_n}\frac{dw_n}{dL_{n'}}, \quad \epsilon^D_{w_n,L_n} = \frac{L_n}{w_n}\frac{dw_n}{dL_n}, \quad \epsilon_{w_n,K} = \frac{K}{w_n}\frac{dw_n}{dK}, \forall n, n' \in N. \tag{8}$$

Readers should be careful about the notations of $\epsilon^D_{w_n,L_n}$, which we distinguish from ϵ_{w_n,L_n}. For a given skill-type n, the labor input L_n can affect wage w_n directly holding Y unchanged, we denote this channel by $\epsilon^D_{w_n,L_n}$. It can also affect wage through equilibrium effect indirectly, where Y changes, captured by ϵ_{w_n,L_n}. Moreover, $\epsilon_{w_n,L_{n'}}, n \neq n'$ capture the spillover effects of labor input $L_{n'}$ on wage w_n due to imperfect substitution between skills. $\epsilon_{w_n,K}$ captures the capital-skill complementarity. Since we restrict rental rate R to be uniform, the definition of elasticity for substitution effect is more straightforward,

$$\epsilon_{R,L_n} = \frac{L_n}{R}\frac{dR}{dL_n}, \quad \epsilon_{R,K} = \frac{K}{R}\frac{dR}{dK}, \quad \forall n \in N.$$

Next, it is time to turn on the automation effect. Armed with the equilibrium factor prices, we first define the elasticity of them with respect to automation technology,

$$\epsilon_{w_n,\alpha_n} = \frac{\alpha_n}{w_n}\frac{dw_n}{d\alpha_n}, \quad \epsilon_{R,\alpha} = \frac{\alpha}{R}\frac{dR}{d\alpha}, \quad \forall n \in N.$$

The intuition behind that is the automation can induce the change of demand structure for labor and capital, consequently impacting factor prices.[5] Furthermore, Lemma 2 shows that the equilibrium automation technology is determined by factor inputs,

$$\alpha_n \equiv \alpha_n(K,\mathcal{L}) = 1 - \frac{1}{\gamma_n}\frac{\mu_n(\alpha_n)L_n}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn}, \quad \alpha \equiv \alpha(K,\mathcal{L}) = \frac{K}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn}, \quad \forall n \in N.$$

Following the same logic, we define the elasticity of automation technology with respect to factor inputs,

$$\epsilon_{\alpha_n,L_{n'}} = \frac{L_{n'}}{\alpha_n}\frac{d\alpha_n}{dL_{n'}}, \quad \epsilon^D_{\alpha_n,L_n} = \frac{L_n}{\alpha_n}\frac{d\alpha_n}{dL_n}, \quad \epsilon_{\alpha_n,K} = \frac{K}{\alpha_n}\frac{d\alpha_n}{dK}, \quad \forall n \in N,$$

$$\epsilon_{\alpha,L_n} = \frac{L_n}{\alpha}\frac{d\alpha}{dL_n}, \quad \epsilon_{\alpha,K} = \frac{K}{\alpha}\frac{d\alpha}{dK}, \quad \forall n \in N.$$

We summarize these demand-side elasticities in Table 2.

Table 2: Demand-side elasticities

Indirect wage elasticity of labor	$\epsilon_{w_n,L_{n'}}$	$(1-\rho)(1-\alpha_{n'})\gamma_{n'}$
Direct wage elasticity of labor	$\epsilon^D_{w_n,L_n}$	$\rho(1-\alpha_n)-1$
Wage elasticity of capital	$\epsilon_{w_n,K}$	$(1-\rho)\alpha+\rho\alpha_n$
Rental rate elasticity of labor	ϵ_{R,L_n}	$(1-\alpha_n)\gamma_n$
Rental rate elasticity of capital	$\epsilon_{R,K}$	$\alpha-1$
Wage elasticity of automation	ϵ_{w_n,α_n}	$-\frac{\alpha_n}{1-\alpha_n}$
Rental rate elasticity of average automation	$\epsilon_{R,\alpha}$	1
Indirect automation elasticity of labor	$\epsilon_{\alpha_n,L_{n'}}$	$\frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'}+\int(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,L_{n'}}dn}{\epsilon_{\mu_n(\alpha_n),\alpha_n}+\alpha_n/(1-\alpha_n)}$
Direct automation elasticity of labor	$\epsilon^D_{\alpha_n,L_n}$	$\frac{\rho(1-\alpha_n)-1}{\alpha_n/(1-\alpha_n)}$
Automation elasticity of capital	$\epsilon_{\alpha_n,K}$	$\frac{(1-\rho)\alpha+\rho\alpha_n+\int(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn}{\epsilon_{\mu_n(\alpha_n),\alpha_n}+\alpha_n/(1-\alpha_n)}$
Average automation elasticity of labor	ϵ_{α,L_n}	$-(1-\alpha_n)\gamma_n-\int(1-\alpha_{n'})\gamma_{n'}\epsilon_{\mu_{n'}(\alpha_{n'}),\alpha_{n'}}\epsilon_{\alpha_{n'},L_n}dn'$
Average automation elasticity of capital	$\epsilon_{\alpha,K}$	$1-\alpha-\int(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn$

[5] Note that $\epsilon_{w_n,\alpha_{n'}} = 0$ for $n \neq n'$ due to the Envelope Theorem, and with no heterogeneous in rental rate, the definition of $\epsilon_{R,\alpha_n}, \forall n \in N$ is redundant, which can be understood by $R = \alpha Y/K$.

Proof. See Appendix B.2. ■

Following the Chain rule, it is clear for us to characterize the automation effect of factor inputs on factor prices. For instance, capital supply can lead to the adjustment of automation technology, i.e., $\epsilon_{\alpha_n,K}$, which will further induce the change of wage, i.e., ϵ_{w_n,α_n}. Readers can try to assemble these elasticities by themselves.

3 Automated Technical Change

The relationship between factor prices and factor inputs is an important building block of our tax incidence analysis below. In this section, we take labor and capital inputs as exogenous for the moment and consider the impacts of them on wages and rental rate. As already alluded to, these impacts could be decomposed into *automation effect* and *substitution effect*. In the following analysis, we focus on the implications of automated technical change for factor prices and wage distribution.

3.1 Automation and Factor Prices

Automation Effect–The key ingredients of automation effect is the adjustment of automation technology corresponding to factor inputs, and further, the implications of automation technology for factor prices. Using the demand-side elasticities defined in table 2, the rates of change in wage and rental rate induced by change in factor inputs through automation effect can be given as follows,

$$\frac{dw_n}{w_n}|_{AE} = \epsilon_{w_n,\alpha_n}\left[\epsilon^D_{\alpha_n,L_n}\frac{dL_n}{L_n} + \int_{\underline{n}}^{\bar{n}}\epsilon_{\alpha_n,L_{n'}}\frac{dL_{n'}}{L_{n'}}dn' + \epsilon_{\alpha_n,K}\frac{dK}{K}\right], \quad \forall n \tag{9}$$

$$\frac{dR}{R}|_{AE} = \epsilon_{R,\alpha}\left[\int_{\underline{n}}^{\bar{n}}\epsilon_{\alpha,L_n}\frac{dL_n}{L_n}dn + \epsilon_{\alpha,K}\frac{dK}{K}\right] \tag{10}$$

Equation (9) shows that automation effect on the rate of change in wage w_n can be expressed in three terms. Labor inputs $\mathcal{L}$ lead to automated technical change directly and indirectly, which are captured by the first two terms. Capital input K can also induce the adjustment of automation technology, which is captured by the third term, the adjustment of automation technology can then have implications for wage. As for rental rate, analogously, both labor inputs $\mathcal{L}$ and capital input K can induce automated technical change, then changes in rental rate, which are characterized by the first and second term in equation (10) respectively. The signs of these terms are of the most interest to us. Assuming $0 < \rho < 1$, we have,[6]

$$\epsilon_{R,\alpha}; \epsilon_{\alpha_n,L_{n'}}; \epsilon_{\alpha_n,K}; \epsilon_{\alpha,K} > 0, \quad \epsilon_{w_n,\alpha_n}; \epsilon^D_{\alpha_n,L_n}; \epsilon_{\alpha,L_n} < 0.$$

To elucidate the intuition behind, we first argue that $\epsilon_{R,\alpha} > 0$ and $\epsilon_{w_n,\alpha_n} < 0$ actually reflect the change in structure of demand for factor inputs as the automation technology changes. The higher the degree of automation, the more the demand for capital (the less the demand for labor), the higher the rental rate (the lower the wage), and vice versa. Then let us turn to the adjustment of automation. The direct effect of labor input L_n may take over the task produced by capital, then lead to the decrease in automation, i.e., $\epsilon^D_{\alpha_n,L_n} < 0$. But for the indirect effect of labor input $\mathcal{L}$, it generates positive spillover effect on automation technology in a given occupation, and thus $\epsilon_{\alpha_n,L_{n'}} > 0$. As capital in the economy becomes more abundant, the degree of occupational and average automation technology will increase, i.e., $\epsilon_{\alpha_n,K}$ and $\epsilon_{\alpha,K}$ are positive. Follow the same intuition, the effect of labor input on the degree of average automation technology is negative $\epsilon_{\alpha,L_n} < 0$. Thus, for the wage in a given skill-type n, the automation effect of labor inputs may increase the wage directly and decrease the wage indirectly, through the first and second terms in (9), respectively. As for the third term, capital input leads to a decreasing in wages through automation effect. Analogously, the automation effect of labor inputs on rental rate is negative, but this effect of capital input is positive.

Substitution Effect–Now we turn to the direct impact of factor inputs on wage and rental rate, holding automation technology fixed, that is *substitution effect*. Similarly, using demand-side elasticities defined in table 2, the rate of change in wage and rental rate induced by substitution effect can be given as follows,

$$\frac{dw_n}{w_n}|_{SE} = \epsilon^D_{w_n,L_n}\frac{dL_n}{L_n} + \int_{\underline{n}}^{\bar{n}} \epsilon_{w_n,L_{n'}}\frac{dL_{n'}}{L_{n'}}dn' + \epsilon_{w_n,K}\frac{dK}{K} \quad \forall n \tag{11}$$

$$\frac{dR}{R}|_{SE} = \int_{\underline{n}}^{\bar{n}} \epsilon_{R,L_n}\frac{dL_n}{L_n}dn + \epsilon_{R,K}\frac{dK}{K}, \tag{12}$$

Here, the effects are broken down in the same way as the discussion of automation effect. Following the conditions in Appendix C.1, we have $\epsilon^D_{w_n,L_n} < 0$, which means an increasing in labor inputs will reduce his own wage directly, and $\epsilon_{w_n,L_{n'}} > 0$, which captures the indirectly positive spillover effect of labor input on wage. Capital input can also increase the level of wage due to capital-skill complementarity, i.e., $\epsilon_{w_n,K} > 0$. As for rental rate, i.e., the price of capital, an increase of labor inputs will increase rental rate while the increase of capital input will decrease

[6] For brevity, we leave the corresponding prove in Appendix C.1.

it. The intuition behind is that factor inputs are *q-complements* in our setting. [7]

By now, attentive readers may have noticed that the automation effect and the substitution effect are indeed opposite to each other. For instance, while capital accumulation ($dK/K > 0$) can improve the level of wages through substitution effect, it can also lead to a scenario where most tasks are taken over by capital, i.e., automation effect, when there is little need for labor, wages will decrease. Labor accumulation follows the similar intuition. To visually display the analysis above, Figure 1 depicts quantitative results of the demand-side elasticities at different labor income quantiles.

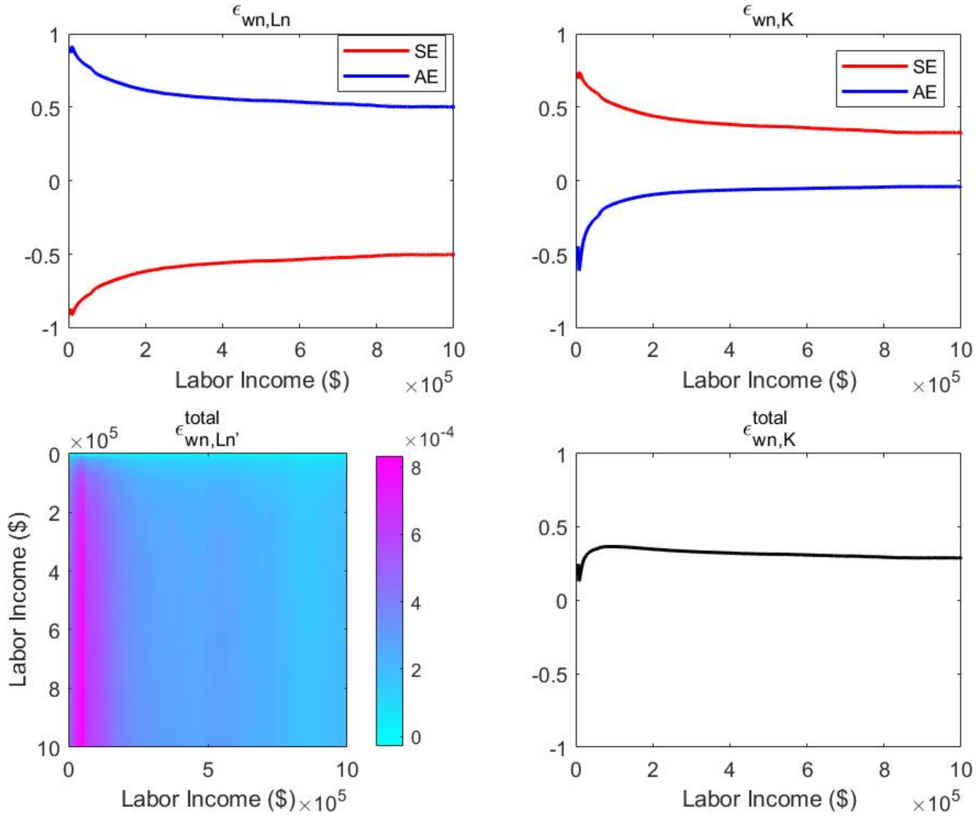

Figure 1: Elasticity of wage with respect to labor and capital inputs

The left upper panel of Figure 1 shows how labor input affect his own wage directly, which can be decomposed into automation effect, i.e., $\epsilon_{w_n,\alpha_n}(\epsilon^D_{\alpha_n,L_n} + \epsilon_{\alpha_n,L_n})$, and substitution effect, i.e. $\epsilon^D_{w_n,L_n} + \epsilon_{w_n,L_n}$. While occupational labor input reduces his own wage through substitution effect, it can also increase the own-wage through automation. The left bottom panel displays

[7] Kim (2000),p4: Inputs i and j are net q - complements (substitutes) in the production of a fixed output, so that an increase in the quantity of the jth input increases (decreases) the price, or the marginal valuation.

the matrix of total elasticity of wage with respect labor input, both own-occupational and cross-occupational, the value on the diagonal corresponds to the effect of own-occupational labor input on wages. It is a novel finding that the values on the matrix are numerically insignificant, which is slightly positive, implying there exists counteracting forces between substitution effect and automation effect. Due to automated technical change, wages are less sensitive to labor inputs. The intuition behind is similar to Clemens et al. (2018) and Blundell et al. (2022), in which a relative decrease of agricultural workers or a relative increase of educated people will induce directed technical change, which counteracts with the effects of change in labor inputs itself, thus wages do not change dramatically. Another interesting finding is that, relative to the labor input of high income individual, wages are more sensitive to the labor input of middle or bottom income individual, but is still far below the sensitivity of wage with respect to capital input.

The right upper panel depicts how wage reacts to capital input through automation effect and substitution effect, $\epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,K}$ and $\epsilon_{w_n,K}$. As anticipated, capital inputs increase wages through substitution but decrease them through automation effect. Moreover, we sum the two effects and display the comprehensive effects in the right bottom panel. We find that the substitution effect of capital input on wage dominates the automation effect in our baseline calibration, stimulating capital inputs may benefit the wage income of all individual. However, as we can see, high income individuals benefits more.

There are two takeaways from the figure above. First, either labor input or capital input can lead to the change in wages through automation effect and substitution effect in opposite directions. However, wages are more sensitive to capital input relative to labor income when taking account the automation technology. Second, as Acemoglu and Autor (2011) have mentioned, the U.S. labor market has witnessed *stagnation* in the real wage of bottom and middle income populations. Our model sheds light on this phenomenon, since the right bottom panel shows that wage stagnation in low skill occupation can go hand in hand with capital accumulation.

Proposition 1 *Assume $0 < \rho < 1$ and the equilibrium automation technology $\boldsymbol{\alpha}$ exists at all wage quantiles, then the automation effect and substitution effect of factor inputs on factor prices are always in the opposite directions. Moreover, when $\eta = 0$, these two effects are totally counteracted.*[8]

[8]From Table 2, let $\epsilon_{\mu_n(\alpha_n),\alpha_n} = \eta = 0$, one can find that $\epsilon_{w_n,L_{n'}} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,L_{n'}} = 0$, and $\epsilon_{w_n,K} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,K} = 0$, and so on.

3.2 Automation and Wage Distribution

We are not only concerned about the impact of factor inputs on the level of wages, but also care about the impact of factor inputs on wage distribution, which plays a central role in the design of optimal income taxation. Given any two individuals with different skill types, $n, \tilde{n} \in N$ and $n > \tilde{n}$. One could regard n as skilled college graduates and $\tilde{n}$ as unskilled high school graduates, then the implications of factor inputs for wage premium can be viewed through the lens of the following decomposition.

Automation effect–We substitute the expressions of elasticities into equation (9), the automation effect on wage premium can be given as follows,

$$\begin{aligned}\frac{dw_n}{w_n}|_{AE} - \frac{dw_{\tilde{n}}}{w_{\tilde{n}}}|_{AE} = &\left[1-\rho(1-\alpha_n)\right]\frac{dL_n}{L_n} - \left[1-\rho(1-\alpha_{\tilde{n}})\right]\frac{dL_{\tilde{n}}}{L_{\tilde{n}}} \\ &+ \left[\frac{1-\rho}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}}+1} - \frac{1-\rho}{\eta(1-\alpha_n)/\alpha_n+1}\right]\frac{1}{1-\tilde{\eta}}\int(1-\alpha_{n'})\gamma_{n'}\frac{dL_{n'}}{L_{n'}}dn' \\ &+ \left[\frac{\rho(1-\tilde{\eta})\alpha_{\tilde{n}}+\varphi}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}}+1} - \frac{\rho(1-\tilde{\eta})\alpha_n+\varphi}{\eta(1-\alpha_n)/\alpha_n+1}\right]\frac{1}{1-\tilde{\eta}}\frac{dK}{K}, \quad \forall n > \tilde{n}.\end{aligned} \tag{13}$$

Where $\varphi = (1-\rho)\alpha + \rho\int\frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)}dn > 0$, and $\tilde{\eta} = \int\frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)}dn \in (0,1)$. Note that we have rewritten the expressions of $\epsilon_{\alpha_n, L_{n'}}$ and $\epsilon_{\alpha_n, K}$ in Appendix C.1.

Substitution effect–As for the substitution effect on wage premium, we substitute the expressions of elasticities into equation (11), then obtain,

$$\frac{dw_n}{w_n}|_{SE} - \frac{dw_{\tilde{n}}}{w_{\tilde{n}}}|_{SE} = \left[\rho(1-\alpha_n)-1\right]\frac{dL_n}{L_n} - \left[\rho(1-\alpha_{\tilde{n}})-1\right]\frac{dL_{\tilde{n}}}{L_{\tilde{n}}} + \rho(\alpha_n - \alpha_{\tilde{n}})\frac{dK}{K} \quad \forall n > \tilde{n}. \tag{14}$$

Proposition 2 *Suppose* $0 < \rho < 1$ *and* $\dot{\alpha}_n < 0$ *are satisfied* [9], *given a change in the structure of factor inputs,* $dL_n/L_n = dL_{\tilde{n}}/L_{\tilde{n}} = dL/L > 0, \forall n, \tilde{n} \in N$ *and* $dK/K > 0$, *then labor inputs can directly decrease wage premium through automation effect while increase it through substitution effect, moreover, labor inputs can also indirectly increase wage premium through automation effect due their spillover. Capital accumulation can increase wage premium through automation effect while decrease it through substitution effect.*

Proof. See Appendix C.2. ■

Proposition 2 shows the implications of automation effect and substitution effect for wage distribution. To understand the intuition behind, we turn to figure 1 again. Taking capital accumulation for instance, the SE-line in the right upper panel shows that capital input increases

[9] The intuition behind α_n decreases with n is that in the occupation of high skill-type, less task could be automated. We check it in the quantitative analysis.

wages at all skill types, but it benefits low skilled workers more due to the substitution effect, thus induces to a decrease in wage premium. On the contrary, the AE-line shows that capital input reduce wages of all skill types through automation effect, but wage of low skill occupation bear most of the brunt, leading to an increase in wage premium. A similar is for the labor inputs. In a word, the automation effect and substitution effect of factor inputs on wage distribution also stand on the opposite. As we shall see, these two effects play a central role in the design of tax system. Moreover, in consideration of a continuum of wages, our model may shed light on *wage staganation*, due to these heterogeneous effects across wages. Acemoglu and Autor (2011) prove that a task-based framework could interpret several central trends in the US labor market, which bolster our confidence that tax policy should make some responses to automated technical change. As automation induced by capital accumulation accounts for part of the increasing in wage inequality, taxing capital may be desirable if redistribution is valuable. However, the cost of taxing capital is that it curbs capital accumulation, leading to decreasing in wages, then government revenue raised from labor income may incur losses. We turn to tax incidence analysis to shed light on the trade-off between efficiency and equity in the following section.

4 Tax Incidence Analysis

So far, we have clarified how factor inputs impact their prices through the automation effect and substitution effect. In this section, we introduce tax reforms to initiate the analysis of tax incidence. Drawing upon the central theoretical contribution made by Sachs et al. (2020), we derive closed-form formulas for the first-order effects of tax reforms on individual labor and capital supplies, wages and rental rate, government revenue, and then social welfare. As we shall see, our extension supplements the work of Sachs et al. (2020) by introducing capital-skill complementarity and automation technology, which can be viewed through the lens of a revised integral equation system.

4.1 Effects on Factor Supplies

Given a nonlinear tax perturbation described in preceding, i.e., $\tilde{T}_i = T_i + \kappa_i \tau_i$ with $i \in [z, x]$, we denote dl_n/l_n as the percentage change in labor supply induced by the tax reform as $\kappa_i \to 0$, analogously, the change in capital supply da_q/a_q, the change in wage dw_n/w_n, and the change in rental rate dR/R.

With a separable tax function, we know that factor supplies depend on their prices and marginal retention rate, that is

$$\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T_z'} \frac{\tau_z'(z_n)}{1-T_z'(z_n)} + \epsilon_{l_n,w_n} \frac{dw_n}{w_n}, \quad \frac{da_q}{a_q} = -\epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} + \epsilon_{a_q,R} \frac{dR}{R}. \tag{15}$$

Where the first term captures the direct effect of tax reform and the the second term accounts for the general-equilibrium feedback via factor prices. Bear in mind that the implications of factor inputs for factor prices have been decomposed into substitution and automation effect, armed with the foregoing analysis, we have the following lemma that

Lemma 3 *Given any tax perturbation τ_i, and $\kappa_i \to 0$ for $i \in [z, x]$, the incidence of the tax reforms on factor supplies can be depicted by the following integral equation system,*

$$\begin{aligned} \frac{dl_n}{l_n} = & -\epsilon_{l_n,1-T_z'} \frac{\tau_z'(z_n)}{1-T_z'(z_n)} + \underbrace{\epsilon_{l_n,w_n} \left[\int \epsilon_{w_n,L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{w_n,K} \frac{dK}{K} \right]}_{\text{Substitution Effect}} \\ & + \underbrace{\epsilon_{l_n,w_n} \epsilon_{w_n,\alpha_n} \left[\int \epsilon_{\alpha_n,L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{\alpha_n,K} \frac{dK}{K} \right]}_{\text{Automation Effect}} \end{aligned} \tag{16}$$

$$\begin{aligned} \frac{da_q}{a_q} = & -\epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} + \underbrace{\epsilon_{a_q,R} \left[\int \epsilon_{R,L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{R,K} \frac{dK}{K} \right]}_{\text{Substitution Effect}} \\ & + \underbrace{\epsilon_{a_q,R} \epsilon_{R,\alpha} \left[\int \epsilon_{\alpha,L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{\alpha,K} \frac{dK}{K} \right]}_{\text{Automation Effect}} \end{aligned} \tag{17}$$

Note that $\epsilon^D_{w_n,L_n} + \epsilon_{w_n,\alpha_n} \epsilon^D_{\alpha_n,L_n} = 0$. Equation (16) counterparts to the integral equation proposed by Sachs et al. (2020), but ours is distinct in two ways: First, they do not take into account the capital-skill complementarity, which is captured by the second term in the second square bracket, i.e., $\epsilon_{l_n,w_n} \epsilon_{w_n,K} \frac{dK}{K}$. Second, while their general framework does not specify technology change, we take automated technical change into consideration, which has been regarded to have a profound impact on labor market, especially on wage inequality. This channel is captured by the third term of (16). The first term of (16) appears in many Mirrleesian taxation literature where wage is exogenous, which captures the direct effect of tax reform on labor supply. Parallel research has examined the implications of directed technical change for tax design (Loebbing, 2020), However, their work does not consider the interaction between capital and labor supply through the equilibrium effect, represented by the second term in the square brackets.

Analogously, equation (17) characterizes the effect of tax reform on capital supply. The first term captures the direct effect of tax reform, this term appears in Saez and Stantcheva (2018) where the net return on capital is exogenous. The second and third terms capture the general equilibrium feedback due to the endogenous rental rate, via substitution effect and automation effect, respectively. To the best of our knowledge, there is limited literature that examines the incidence of capital income tax reform in a general equilibrium framework, which could have implications for the design of capital income taxation.

To make the incidence of tax reform on factor supplies more clear, we solve the simultaneous equations (16) and (17) and obtain the following proposition.

Proposition 3 *Denote $\epsilon^{total}_{w_n,i} = \epsilon_{w_n,i} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,i}$ and $\epsilon^{total}_{R,i} = \epsilon_{R,i} + \epsilon_{R,\alpha}\epsilon_{\alpha,i}$ for $i \in \{\mathcal{L}, K\}$, the integral equation system can be reduced to*

$$\frac{dl_n}{l_n} = \underbrace{-\epsilon_{l_n,1-T'_z}\frac{\tau'_z(z_n)}{1-T'_z(z_n)}}_{DE}\underbrace{-\epsilon_{l_n,w_n}\epsilon^{total}_{w_n,K}\chi\int\omega_q\epsilon_{a_q,1-T'_x}\frac{\tau'_x(x_q)}{1-T'_x(x_q)}dq}_{CE} \\ \underbrace{+\epsilon_{l_n,w_n}\int\left[\epsilon^{total}_{w_n,L_{n'}} + \epsilon^{total}_{w_n,K}\chi\bar{\epsilon}_{K,R}\epsilon^{total}_{R,L_{n'}}\right]\frac{dl_{n'}}{l_{n'}}dn'}_{GE} \tag{18}$$

$$\frac{dK}{K} = -\chi\int\omega_q\epsilon_{a_q,1-T'_x}\frac{\tau'_x(x_q)}{1-T'_x(x_q)}dq + \chi\bar{\epsilon}_{K,R}\int\epsilon^{total}_{R,L_{n'}}\frac{dl_{n'}}{l_{n'}}dn'. \tag{19}$$

Where $\chi = \frac{1}{1-\bar{\epsilon}_{K,R}\epsilon^{total}_{R,K}}$, $\bar{\epsilon}_{K,R} = \int\omega_q\epsilon_{a_q,R}dq$, and $\omega_q = \frac{a_q f_q(q)}{K}$.

Moreover, denote $\zeta_{w_n,L_{n'}} = \epsilon^{total}_{w_n,L_{n'}} + \epsilon^{total}_{w_n,K}\chi\bar{\epsilon}_{K,R}\epsilon^{total}_{R,L_{n'}}$ and assume $\int\int(\epsilon_{l_n,w_n}\zeta_{w_n,l_{n'}})^2dn'dn < 1$,[10] *the integral equation for labor supply can be solved as follows*

$$\frac{dl_n}{l_n} = \sum_{t=0}^{\infty}\left(\frac{dl_n}{l_n}\right)^t, \quad \forall n \in N. \tag{20}$$

where

$$\left(\frac{dl_n}{l_n}\right)^0 = -\epsilon_{l_n,1-T'_z}\frac{\tau'_z(z_n)}{1-T'_z(z_n)} - \epsilon_{l_n,w_n}\epsilon^{total}_{w_n,K}\chi\int\omega_q\epsilon_{a_q,1-T'_x}\frac{\tau'_x(x_q)}{1-T'_x(x_q)}dq$$

[10] This technical condition ensures that the series in (20) converges, and the technique details can be found in Sachs et al. (2020). We verify it in the quantitative analysis.

$$\left(\frac{dl_n}{l_n}\right)^t = \int \zeta_{w_n,L_{n'}} \left(\frac{dl_n}{l_n}\right)^{t-1} dn', \forall t > 0.$$

Proof. See Appendix D.1 ■

Equation (18) in Proposition 3 is a revised integral equation that counterparts to the one in Lemma 1 of Sachs et al. (2020). The first term captures the incidence of a reform in labor income tax on labor supply, namely the *direct effect*. In addition, reform in capital income tax would also has impact on capital supply, then on wages, and thus the labor supply. We call this channel the *cross effect*. Though the tax function is restricted to be separable, the *cross effect* appears due to capital-skill complementarity. In our baseline environment, we have $\epsilon^{total}_{w_n,K} > 0, \forall n$ (See Figure 1), thus it is easy to know that a more aggressive tax reform, i.e., $\tau_z^{'}(z_n), \tau_x^{'}(x_n) > 0$, would compress individual labor supply via DE and CE. The third term characterizes the general equilibrium feedback, which is a standard term in general equilibrium framework. However, we isolate two different channels, imperfect substitution between skills and capital-skill complementarity, captured by the first and second term in the square bracket. Turn to equation (19), while it is intuitive that capital income tax reform impacts capital supply directly, symmetrically, labor supplies induced by tax reform can also have implications for capital supply due to capital-skill complementarity.

4.2 Effects on Factor Prices

Following our preceding analysis in Section 3, we know that the comprehensive effects on factor prices can be given by

$$\frac{dw_n}{w_n} = \frac{dw_n}{w_n}|_{AE} + \frac{dw_n}{w_n}|_{SE}, \quad \frac{dR}{R} = \frac{dR}{R}|_{AE} + \frac{dR}{R}|_{SE}.$$

Plugging equation (19) into equation (9) - (12), it is easily to obtain the incidence of arbitrary tax reform (τ_z, τ_x) on equilibrium factor prices.

Corollary 1 *The incidence of tax reform on factor prices is given by*

$$\frac{dw_n}{w_n} = -\epsilon^{total}_{w_n,K}\chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1-T_x^{'}(x_q)} dq + \int \left[\epsilon^{total}_{w_n,L_{n'}} + \epsilon^{total}_{w_n,K}\chi\bar{\epsilon}_{K,R}\epsilon^{total}_{R,L_{n'}}\right] \frac{dl_{n'}}{l_{n'}} dn' \tag{21}$$

$$\frac{dR}{R} = -\epsilon^{total}_{R,K}\chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1-T_x^{'}(x_q)} dq + \int \left[\epsilon^{total}_{R,L_{n'}} + \epsilon^{total}_{R,K}\chi\bar{\epsilon}_{K,R}\epsilon^{total}_{R,L_{n'}}\right] \frac{dl_{n'}}{l_{n'}} dn' \tag{22}$$

Corollary 1 shows that the adjustment of factor prices can be given as a function of labor supply changes characterized by Proposition 3. Since $\epsilon^{total}_{w_n,K} > 0$ and $\epsilon^{total}_{R,K} < 0$, [11] we know that the incidence of an aggressive capital income tax reform may be negative on wages but positive on rental rate, which is captured by the first term in equation (21) and (22). While the US tax code has gone through a reduction in capital tax between 2000-2018(Acemoglu et al., 2020), to some extent, our model might lend support to the decreasing in return to capital, and increasing in average wage. Moreover, as tax reforms (τ_z, τ_x) lead to the changes in labor supply, which can further induce factor price changes, i.e., the last term in equation (21) and (22).

4.3 Effects on Government Revenue

Now we move on to the incidence of tax reforms on government revenue. As depicted, the government obtains revenue by levying tax on capital income and labor income separately. For convenience, we denote $f_q(q) = \int_N f(n,q)dn$ as the density of individuals with wealth type-q, and $f_n(n) = \int_Q f(n,q)dq$ the density of individuals with skill type-n, thus government revenue is given by

$$B = \int_Q T_x(Ra_q) f_q(q)dq + \int_N T_z(w_n l_n) f_n(n)dn.$$

Note that $x_q = Ra_q$ and $z_n = w_n l_n$. Given the tax reform (τ_x, τ_z), it is easily to derive the change in government revenue as follow

$$\begin{aligned} dB = & \underbrace{\int_Q \tau_x(x_q) f_q(q)dq + \int_N \tau_z(z_n) f_n(z_n)dn}_{\text{Mechanical Effect}} \\ & + \underbrace{\int_Q T'_x(x_q)\left[\frac{da_q}{a_q} + \frac{dR}{R}\right] x_q f_q(q)dq + \int_N T'_z(z_n)\left[\frac{dl_n}{l_n} + \frac{dw_n}{w_n}\right] z_n f_n(n)dn}_{\text{Behavioral Effect}} . \end{aligned} \tag{23}$$

The first term on the right hand side of (23) is the *mechanical effect* of tax reform on government revenue, holding factor supplies and factor prices unchanged. As individual labor and capital supply respond to tax reform, which can further lead to the change in wages and rental rate, thus resulting in government revenue changing with tax base, i.e., the *behavioral effect*, captured by the second term. In fact, equation (23) is an extension of equation (15) in Sachs et al. (2020). For a CES production function, the Euler's homogeneous function theorem still stands but revised as in Appendix E.

However, by taking into account the automated technical change, the general-equilibrium

[11] From Table 2, it is easily to obtain $\epsilon^{total}_{R,K} = -\int (1-\alpha_n)\gamma_n \eta \epsilon_{\alpha,K} dn < 0$, and $\epsilon^{total}_{w_n,K} > 0$ can be seen in Figure 1 directly.

contribution of tax reform on government revenue distinct from the one in Sachs et al. (2020). Firstly, neglecting the automation effect may result in an overestimation of the "trickle-down" effects in general equilibrium, as the automation effect consistently operates in the opposite direction to the substitution effect. Moreover, we illustrate the existence of additional "trickle-down" mechanisms in the context of multi-base incomes. Considering a tax reform that decreases the marginal tax rate on capital income, the government loses some revenue collected from individual capital income. However, lower capital tax rate stimulates investment, then increases wages because of $\epsilon_{w_n,K}^{total} > 0$. As the labor share increases, raising the marginal tax rates on high incomes will benefit government revenue collected from individual labor income. In a word, relative to the exogenous-wage setting, tax reform that reduces marginal tax rate on capital income while raising the marginal tax rate on labor income, has a general-equilibrium contribution on government revenue.

4.4 Effects on Social Welfare

Tax instruments are not only used to raise government revenue, but play a central role in income redistribution. As tax reforms lead to the change in individual utilities, it would also have implications for social welfare since the government values individual utilities in different weights. We define the social welfare function as follow,

$$W = \frac{1}{\lambda} G\left(\{V(n,q)\}_{n\times q\in N\times Q}\right) + B \tag{24}$$

where $V(n,q)$ denotes the indirect utility function after solving the individual utility maximization problem as given in (1). The function $G(\cdot)$ is defined over individual utilities, and is assumed continuously differentiable, increasing, and concave. λ is the marginal value of public funds (MVPF). To characterize the redistributive motivation of the government, we further define the social welfare weight assigned to individuals as

$$g(n,q) = \frac{1}{\lambda}\frac{dG}{dV(n,q)}, \quad \forall n \in N, q \in Q.$$

Thus for a government that favors the low-skilled individuals endowed with little wealth, we have $g(n,q) > g(n',q')$ for $n < n'$ and $q < q'$. Without loss of generality, we normalize $g(n,q)$ such that $\int_N \int_Q g(n,q)f(n,q)dndq = 1$. Moreover, we denote $g_n(n) = \int_Q g(n,q)f(n,q)dq/f_n(n)$ and $g_q(q) = \int_N g(n,q)f(n,q)dn/f_q(q)$ as the social welfare weight of individuals with skill type-n and wealth type-q, respectively. Now we are allowed to investigate how social welfare responses to tax reforms.

Proposition 4 *The incidence of tax reforms (τ_z, τ_x) on social welfare is given by*

$$
\begin{aligned}
dW = & \underbrace{\int (1-g_n(n))\tau_z(z_n) f_n(n) dn + \int (1-g_q(q))\tau_x(x_q) f_q(q) dq}_{ME} \\
& \underbrace{-\int \left[T_z'(z_n) z_n \frac{\epsilon_{l_n,1-T_z'} \tau_z'}{1-T_z'(z_n)} \right] f_n(n) dn - \int \left[T_x'(x_q) x_q \frac{\epsilon_{a_q,1-T_x'} \tau_x'}{1-T_x'(x_q)} \right] f_q(q) dq}_{BE} \\
& \underbrace{+\int \left[g_n(n)(1-T_z'(z_n)) + T_z'(z_n)(1+\epsilon_{l_n,w_n}) \right] z_n \frac{dw_n}{w_n}|_{SE} f_n(n) dn + \int \left[g_q(q)(1-T_x'(x_q)) + T_x'(x_q)(1+\epsilon_{a_q,R}) \right] x_q \frac{dR}{R}|_{SE} f_q(q) dq}_{SE} \\
& \underbrace{+\int \left[g_n(n)(1-T_z'(z_n)) + T_z'(z_n)(1+\epsilon_{l_n,w_n}) \right] z_n \frac{dw_n}{w_n}|_{AE} f_n(n) dn + \int \left[g_q(q)(1-T_x'(x_q)) + T_x'(x_q)(1+\epsilon_{a_q,R}) \right] x_q \frac{dR}{R}|_{AE} f_q(q) dq}_{AE}
\end{aligned} \tag{25}
$$

Proof. See Appendix D.2. ■

The first term on the right hand of equation (25) is *mechanical effect*, assuming that individual is irresponsive to tax reform. It captures the redistributive gains of tax reform since revenue is transferred between the government and individuals. The second term, known as *behavior effect*, captures individual behavior in response to the tax perturbation. These two effects coincide with the classical intuition of equity-efficiency trade-off, which can be dated back to Mirrlees (1971) and Saez (2001). However, in the context of multiple factor incomes, one should take into account the incidence of both labor income tax reform and capital income tax reform simultaneously. In addition, as depicted in Corollary 1, tax reforms impact factor prices through two channels: *substitution effect* and *automation effect*, these adjustments of factor prices in general equilibrium further lead to the change in social welfare, which are captured by the third and fourth term on the right hand of equation (25). The incidence of tax reform on social welfare differs from the one in Sachs et al. (2020) along two key dimensions: our analysis takes into account multiple factor incomes and automated technical change. Because of these two elements, the second integral in the third term and the two integrals in the fourth term are new channels of the general equilibrium effect.

To elucidate the general equilibrium effect, we take one of these integrals for instance,

$$\int \left[g_n(n)(1-T_z'(z_n)) + T_z'(z_n)(1+\epsilon_{l_n,w_n}) \right] z_n \frac{dw_n}{w_n}|_{SE} f_n(n) dn \tag{26}$$

Suppose that tax reforms increase wage through substitution effect, $\frac{dw_n}{w_n}|_{SE}>0$, it could be a reduction in the marginal capital income tax rate or a raise in the marginal labor income tax rate, so that his labor income increases $(1+\epsilon_{l_n,w_n}) z_n \frac{dw_n}{w_n}|_{SE}$, leading to an increase in government revenue $T_z'(1+\epsilon_{l_n,w_n}) z_n \frac{dw_n}{w_n}|_{SE}$, the second term in the square bracket. Moreover, an increase in his after-tax labor income would contribute to the social welfare by $g_n(n)(1-T_n')z_n\frac{dw_n}{w_n}|_{SE}$, the first term in the square bracket. Note that ϵ_{l_n,w_n} does not appear here due the envelope theorem. Finally, aggregating these effects across individuals weighted by their density $f_n(n)$ leads to (26). The analysis is similar for the rest of integrals in the third and fourth terms in equation (25).

Next, we show that the incidence of tax reforms on social welfare may shed light on the design of optimal (welfare-maximizing) tax system. Considering a given tax reform that enhances the progressivity of labor income tax combined with a reduction of capital income tax rate, through substitution effect, this kind of tax reform may raise government revenue obtained from individual labor income as well as improve individual utility. The cost of this tax reform is compressing individual labor supply as well as reducing capital share in the economy, thus a loss of government revenue obtained from individual capital income. As long as this "trickle down" forces contributes to the social welfare, the tax reform is desirable relative to the exogenous price setting. Moreover, taking into account the automated technical change, i.e., automation effect, the implications for tax design would be revised. Bear in mind that automation effect and substitution effect always stand on the opposite, it is intuitive that the optimal labor income tax should be less progressive relative to the exogenous automation technology setting. Since as the "trickle down" forces weakens, the marginal welfare gain from enhancing the progressivity of labor income tax is reduced. We characterize the optimal tax system theoretically and quantitatively in the following analysis.

5 Optimal Taxation

In this section, we extend the variational techniques proposed by Sachs et al. (2020) to the case with multiple factor incomes. Following the dual approach, we derive the optimal (welfare-maximizing) tax system expressed with sufficient statistics. In Appendix F, we also provide nonlinear income tax formulas using primal approach (mechanism design). As anticipated, tax formulas derived by the two approaches are indeed equivalent.

5.1 Nonlinear Labor Income Taxation

Part of the optimal tax system is a nonlinear labor income taxation (NLIT), extensive literature has studied it since the seminal work of Mirrlees (1971). In the context of general equilibrium, Sachs et al. (2020) has proposed several equivalent ways to deduce the optimal tax formula, one of which is a novel tax reform approach, as known as the dual approach. As we shall see, this basic approach can be applied to multiple factor incomes. Consider a multiple tax perturbations (τ_z, τ_x), from equation (17) and (18), we know that the incidence of the given tax reform on factor supplies can be given by

$$\frac{da_q}{a_q} = -\epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1-T_x^{'}(x_q)} + \epsilon_{a_q,R} \int \epsilon_{R,L_{n'}}^{total} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{a_q,R} \epsilon_{R,K}^{total} \frac{dK}{K} \tag{27}$$

$$\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T_z'} \frac{\tau_z^{'}(z_n)}{1-T_z^{'}(z_n)} - \epsilon_{l_n,w_n} \epsilon_{w_n,K}^{total} \chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1-T_x^{'}(x_q)} dq + \epsilon_{l_n,w_n} \int \zeta_{w_n,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn' \tag{28}$$

Now, by specifying an elementary labor income tax reform τ_z^1 at income $z = z^*$, we are aimed to find the counteracting perturbation τ_z^2 and τ_x, such that equation (27) and (28) can be reduced to

$$\frac{da_q}{a_q} = 0, \quad \frac{dl_n}{dl_n} = -\epsilon_{l_n,1-T_z'} \frac{\tau_z^{1'}(z_n)}{1-T_z'(z_n)}.$$

Lemma 4 *Given a Dirac labor income tax perturbation* τ_z^1, *i.e.,* $\tau_z^1 = \mathbf{I}_{z \geq z^*}$ *and* $\tau_z^{1'}(z) = \delta_{z^*}(z)$, *the counteracting tax perturbation* τ_z^2 *and* τ_x *should be*

$$\tau_z^2(z_n) = -\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})} \left(\frac{dz_{n^*}}{dn^*}\right)^{-1} \int_0^{z_n} \epsilon_{w,L_{n^*}}^{total} [1 - T_z'(z) - T_z''(z)z] dz \tag{29}$$

$$\tau_x(x_q) = (1 - T_x^{'}(x_q)) x_q \int \epsilon_{R,L_{n'}}^{total} \frac{dl_{n'}}{l_{n'}} dn', \tag{30}$$

Moreover, the incidence of tax reform on wage and rental rate can be given as follows

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,L_{n^*}}^{total} \frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})} \left(\frac{dz_{n^*}}{dn^*}\right)^{-1}, \quad \frac{dR}{R} = -\epsilon_{R,L_{n^*}}^{total} \frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})} \left(\frac{dz_{n^*}}{dn^*}\right)^{-1}. \tag{31}$$

Proof. See Appendix E.1 ■

The intuition behind Lemma 4 is that we use tax perturbation τ_z^2 to counteract the adjustment of wages, and τ_x is used to counteract the adjustment of rental rate. As factor supplies are given by $a_q(1 - T_x', R)$ and $l_n(1 - T_z', w_n)$, it is easily obtained that $\frac{da_q}{a_q} = 0$ and $\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T_z'} \frac{\tau_z^{1'}}{1-T_z'}$. While the capital supply has been fixed, the impact of tax reform on factor prices can only through the adjustment of labor supply, i.e., equation (31). In this way, we are allowed to avoid involving integral terms, and the deduction of optimal nonlinear labor income taxation becomes more tractable.

The dual approach implies that, for any tax reform, there is no marginal improvement on social welfare (24) when tax system is at optimal. In other words, the incidence of tax reform on social welfare must equal zero while solving the optimal tax formula, i.e.,

$$dW = \int\int g(n,q)dV(n,q)f(n,q)dndq + dB = 0. \tag{32}$$

For convenience, we define the average weight above a given skill-type n as

$$\bar{g}_{z_n} \equiv \bar{g}_n = \frac{\int_{n'>n}\int_Q g(n',q)f(n',q)dqdn'}{\int_{n'>n}\int_Q f(n',q)dqdn'} = \frac{\int_{n'>n} g_{n'}(n')f_{n'}(n')dn'}{1-F_n(n)}.$$

Armed with the incidence analysis in subsection 4.4, the optimal NLIT can be deduced by specifying the tax reform and the corresponding incidence on factor supplies and factor prices in Lemma 4.

Proposition 5 *Consider a given tax reform in Lemma 4, with $dW = 0$ satisfied, the function of optimal nonlinear labor income tax can be given as follow*[12]*,*

$$\begin{aligned}
\frac{T_z'(z_{n^*})}{1-T_z'(z_{n^*})} = {} & \frac{1}{\epsilon_{l_{n^*},1-T_z'}}(1-\bar{g}_{z_{n^*}})\frac{1-F_z(z_{n^*})}{z_{n^*}f_z(z_{n^*})} \\
& \underbrace{-\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\left[(1-\bar{g}_{z_n})\left(\frac{1-T_z'(z_n)}{1-T_z'(z_{n^*})}\right)\left(\frac{1-F_z(z_n)}{z_{n^*}f_z(z_{n^*})}\right)z_n\right]'\epsilon^{SE}_{w_n,L_{n^*}}dz_n}_{\text{Substitution Effect}} \\
& \underbrace{-\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\left[(1-\bar{g}_{z_n})\left(\frac{1-T_z'(z_n)}{1-T_z'(z_{n^*})}\right)\left(\frac{1-F_z(z_n)}{z_{n^*}f_z(z_{n^*})}\right)z_n\right]'\epsilon^{AE}_{w_n,L_{n^*}}dz_n}_{\text{Automation Effect}}
\end{aligned} \tag{33}$$

where $\epsilon^{SE}_{w_n,L_{n^*}} = \epsilon^{D}_{w_{n^*},L_{n^*}}\Delta_{n^*}(n) + \epsilon_{w_n,L_{n^*}}$, $\epsilon^{AE}_{w_n,L_{n^*}} = \epsilon_{w_{n^*},\alpha_{n^*}}\epsilon^{D}_{\alpha_{n^*},L_{n^*}}\Delta_{n^*}(n) + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,L_{n^*}}$.

Proof. See Appendix E.2. ∎

Note that $\Delta_{n^*}(n) = 1$ if $n = n^*$, else, $\Delta_{n^*}(n) = 0$, so that $\epsilon^{SE}_{w_n,L_{n^*}}$ and $\epsilon^{AE}_{w_n,L_{n^*}}$ denote the channels of substitution effect and automation effect, respectively. The first term on the right hand of equation (33) in Proposition 5 is known colloquially as ABC rule of optimal income tax (Diamond, 1998; Saez, 2001). There are three elements in this term that determine optimal income tax rates: the behavior elasticity $\epsilon_{l_n,1-T_z'}$, the redistributive motivation of the government

[12] Loebbing (2020) derived the optimal nonlinear labor income tax formula considering directed technical change, however, he adopted mechanism design method, i.e., primal approach. In appendix F, we shows that primal approach and dual approach are indeed equivalent in deducing optimal tax formula.

$1 - \bar{g}_{z_n}$, and the income (or skill) distribution $\frac{1-F_z(z_{n^*})}{z_{n^*} f_z(z_{n^*})}$. As the skill distribution is bounded, i.e., $\bar{g}_{\underline{n}} = F_n(\bar{n}) = 1$, it leads to a well known result that the marginal tax rate should be zero at the top or bottom income level (Seade, 1977).

The second term on the right hand of equation (33) captures implications of imperfect substitution between skills for optimal income tax rates, which can be dated back to Stiglitz (1982). In their two type individuals model, the general equilibrium forces lead to a lower (resp., higher) top (resp., bottom) marginal tax rate. Sachs et al. (2020) has generalized Stiglitz (1982) to a setting with a continuum of skills, and show that if the tax system is initially suboptimal and progressive, the general-equilibrium "trickle-down" forces may raise the benefits of increasing the marginal tax rates on high incomes. In this paper, as $\epsilon^{SE}_{w_n, L_{n^*}} < 0$ for $n = n^*$, and $\epsilon^{SE}_{w_n, L_{n^*}} > 0$ for $n \neq n^*$, the second term coincides with the conventional "trickle down" forces. Raising the marginal tax rate on high income individuals may raise government revenue due to this general equilibrium effect.

The third term is our novel finding since it accounts for the implications of automated technical change for optimal marginal income tax rate. As has been discussed in Section 3, the automation effect is always contrary to the substitution effect, it is intuitive that optimal marginal tax rate on the high income individuals should be lower relative to exogenous-automation setting.

5.2 Capital Income Taxation

Our model can also be used to the discussion of optimal capital income taxation, which composes an important part of the tax system. In this subsection, we derive formulas for optimal nonlinear and linear capital income taxation in the context of automated technical change.[13]

5.2.1 Nonlinear Capital Income Taxation

The deduction of optimal nonlinear capital income taxation is analogously to the one of labor income taxation. We aim to specify a tax perturbation τ_x^1, and find another counteracting tax perturbation τ_x^2 and τ_z, such that the general equilibrium feedback can be counteracted. Symmetrically, consider a tax reform $\tau_x = \tau_x^1 + \tau_x^2$ and τ_z, such that

$$\frac{da_q}{a_q} = -\epsilon_{a_q, 1-T'_x} \frac{\tau_x^{1'}(x_q)}{1 - T'_x(x_q)}, \quad \frac{dl_n}{l_n} = 0.$$

[13] Saez and Stantcheva (2018) derive formulas for optimal nonlinear and linear capital income taxation in a partial-equilibrium setting, we extend it to a general-equilibrium setting and take into account automation technology.

Armed with the integral equation system given by (27) and (28), it is easy to derive that

Lemma 5 *Given a Dirac capital income tax perturbation τ_x^1, i.e., $\tau_x^1 = \mathbf{I}_{x \geq x^*}$ and $\tau_x^{1'}(x) = \delta_{x^*}(x)$, the counteracting tax perturbation τ_x^2 and τ_z are given by*

$$\tau_x^2(x_q) = -\epsilon_{R,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1} (1-T_x'(x_q))x_q$$

$$\tau_z(z_n) = -\frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1} \int_0^{z_n} \epsilon_{w,K}^{total}[1-T_z'(z)-T_z''(z)z]dz$$

Moreover, the incidence of above tax reform on wage and rental rate can be given by

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1}, \quad \frac{dR}{R} = -\epsilon_{R,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1}.$$

Proof. See Appendix E.3. ∎

The intuition behind Lemma 5 is that we use τ_x^2 to counteract the adjustment of rental rate, such that capital supply is only determined by tax reform τ_x^1. As for labor supply, we use τ_z to counteract dw_n/w_n, such that the incidence of the comprehensive tax reform on labor supply is shut down, i.e., $dl_n/l_n = 0$. This preliminary work is use to avoid the involving of integral term, thus making the derivation of the optimal formula more tractable.

Proposition 6 *Consider the tax reform given in Lemma 5, with $dW = 0$ satisfied, the optimal nonlinear capital income tax formula can be given by*

$$\begin{aligned}\frac{T_x'(x_{q^*})}{1-T_x'(x_{q^*})} = & \frac{1}{\epsilon_{a_{q^*},1-T_x'}}(1-\bar{g}_{x_{q^*}})\frac{1-F_x(x_{q^*})}{x_{q^*}f_x(x_{q^*})} \\ & \underbrace{-\frac{1}{RK(1-T_x'(x_{q^*}))}\int\left[(1-\bar{g}_{z_n})\left(1-T_z'(z_n)\right)(1-F_z(z_n))z_n\right]'\epsilon_{w_n,K}^{SE}dz_n}_{\text{Substitution Effect}} \\ & \underbrace{-\frac{1}{RK(1-T_x'(x_{q^*}))}\int\left[(1-\bar{g}_{z_n})\left(1-T_z'(z_n)\right)(1-F_z(z_n))z_n\right]'\epsilon_{w_n,K}^{AE}dz_n}_{\text{Automation Effect}}\end{aligned} \tag{34}$$

Proof. See Appendix E.4. ∎

Proposition 6 shows that the optimal nonlinear capital income tax formula is determined by three terms. The first term on the right hand of equation (34), captures the ABC rule for capital income tax, which first appears in Saez and Stantcheva (2018). In a partial-equilibrium setting, there are three elements that govern the optimal marginal tax rate: the behavior elasticity of capital supply, the social welfare weights and the capital income distribution. The more unequal the distribution, i.e., higher $\frac{1-F_x(x_{q^*})}{x_{q^*} f_x(x_{q^*})}$, the higher the marginal tax rate at the level of capital income x_{q^*}.

The second term on the right hand of equation (34) captures the implications of the capital-skill complementarity for optimal marginal tax rates, namely, substitution effect. As capital accumulation can raise the all individual wages through substitution effect, i.e., $\epsilon^{SE}_{w_n,K} > 0, \forall n$, this term is negative. The intuition behind is that lower marginal tax rate can stimulate investment, then increase individual wages. This general equilibrium effect contributes to the government revenue raised from individual labor incomes. However, higher wages lead to the increasing in labor supply, hence rental rate due to general equilibrium effect (Note that $\epsilon_{R,L_n} > 0, \forall n$), which may increase capital income inequality, i.e., higher $\frac{1-F_x(x_{q^*})}{x_{q^*} f_x(x_{q^*})}$. Thus, taking into account the general equilibrium effect may also raise the optimal marginal tax rate, especially at the high level of capital income, for the motivation of redistribution. Thus, we predict that the optimal capital income tax may become more progressive relative the exogenous prices setting.

The third term shows how optimal marginal tax rate responses to automated technical change. As capital accumulation compresses individual wages through automation effect, i.e., $\epsilon^{AE}_{w_n,K} < 0$, this term is positive. In the context of automation technology, raising marginal capital tax rate may have general equilibrium contribution to government revenue since it prevents the decrease of individual wages or labor income. However, one should not neglect the general equilibrium effect on capital income distribution, which composes an important ingredient for the design of optimal taxation. As the foregoing analysis suggests that substitution effect is always contrary to automation effect, it is intuitive that the automated technical change may reduce the progressivity of optimal capital tax to some extent.

5.2.2 Linear Capital Income Taxation

We now turn to optimal linear capital income taxation, which is more implementable in practice. Following the same techniques, we first specify an elementary tax reform τ^1_x, then find the correspondingly counteracting tax perturbation τ^2_x and τ_z, the only difference it that we convert the elementary tax reform from Dirac perturbation to Uniform perturbation.

Lemma 6 *Given an Uniform capital income tax perturbation τ^1_x, i.e., $\tau^1_x(x) = x$ and $\tau^{1'}_x(x) = 1$, the* *counteracting tax perturbation τ_x^2 and τ_z can be given by*

$$\tau_x^2(x_q) = -\epsilon_{R,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T_x'}}{1-T_x'(x_q)} dq (1-T_x'(x_q)) x_q$$

$$\tau_z(z_n) = -\int \omega_q \frac{\epsilon_{a_q,1-T_x'}}{1-T_x'(x_q)} dq \int_0^{z_n} \epsilon_{w,K}^{total} (1-T_z'(z) - T_z''(z)z) dz.$$

Moreover, the adjustments of wage and rental rate can be given by

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T_x'}}{1-T_x'(x_q)} dq, \quad \frac{dR}{R} = -\epsilon_{R,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T_x'}}{1-T_x'(x_q)} dq.$$

Proof. See Appendix E.5 ■

Proposition 7 *Consider the tax reform given in Lemma 6, with $dW = 0$ satisfied, the optimal linear capital income tax formula is given as follow*

$$\begin{aligned} \frac{t_x}{1-t_x} = & \frac{\int (1-g_q(q)) x_q f_q(q) dq}{\int \epsilon_{a_q,1-T_x'} x_q f_q(q) dq} \\ & \underbrace{-\int \frac{1}{RK(1-t_x)} \left[(1-\bar{g}_{z_n})(1-T_z'(z_n))(1-F_z(z_n)) z_n \right]' \epsilon_{w_n,K}^{SE} dz_n}_{\text{Substitution Effect}} \\ & \underbrace{-\int \frac{1}{RK(1-t_x)} \left[(1-\bar{g}_{z_n})(1-T_z'(z_n)(1-F_z(z_n)) z_n \right]' \epsilon_{w_n,K}^{AE} dz_n}_{\text{Automation Effect}} \end{aligned} \tag{35}$$

Proof. See Appendix E.6 ■

Proposition 7 provides an expression that decomposes the optimal linear capital income tax formula into three terms. When wage is set to be exogenous, only the first term exists, which is in line with Saez and Stantcheva (2018). It actually depicts the classic trade-off between equity, captured by the numerator, and efficiency, captured by the denominator. $\int g_q(q) f_q(q) dq = 1$ means there is no redistributive concern along the capital income dimension, so that $t_x = 0$. As $\int g_q(q) f_q(q) dq = 0$, it collapses to a revenue maximizing tax rate, i.e., $t_x = 1/(1+e^K)$, where we assume $\epsilon_{a_q,1-T_x'}$ equals constant e^K.

Our formula is distinct with Saez and Stantcheva (2018) in two ways: First, we introduce general equilibrium substitution effect, captured by the second term on the right hand of equation (35). Due to the "trickle down" forces on wage and labor income, it is desirable to lower capital income tax rate. However, this general equilibrium effect may promote the redistributive motivation of the government, i.e., higher $\int (1 - g_q(q)) x_q f_q(q) dq$, which can in turn raise the optimal tax rate. Second, we take into account the automated technical change, which is depicted by the third term. As automation effect is on the contrary of substitution effect, the optimal capital income tax may be lower. One caveat is that, no matter nonlinear or linear capital income taxation, the general equilibrium effects for capital income taxation, i.e., substitution effect or automation effect, involve the schedule of nonlinear labor income taxation $T_z(\cdot)$. Thus, the design of the optimal income tax is not isolated, but needs to consider the relevance of tax instruments. We apply these optimal tax formulas to quantitative analysis in the following section.

6 Quantitative Analysis

To quantitatively assess these general equilibrium effects on optimal tax system, we first calibrate a model of the U.S. economy in 2019, then we apply our formulas in Section 5 to the designs of optimal tax system: separable nonlinear capital income tax system (NLIT-NCIT), and separable linear capital income tax system (NLIT-LCIT), respectively.

6.1 Calibration

Our model with multiple factor incomes requires us to calibrate both of the distribution of labor income and capital income. We use the Distributional National Accounts (DINAs) micro-files of Piketty et al. (2018), which provides both of pretax labor income (*plinc*) and pretax capital income (*pkinc*) at the individual level. Discretizing the observed factor incomes, we are able to calibrate the probability density functions (PDF) of labor and capital income directly, i.e., $f_z(z_n)$ and $f_x(x_q)$. Figure 2 displays the distributional features of observed factor incomes.[14]

[14]The details for calibrating the income distributions can be found in Appendix G.1.

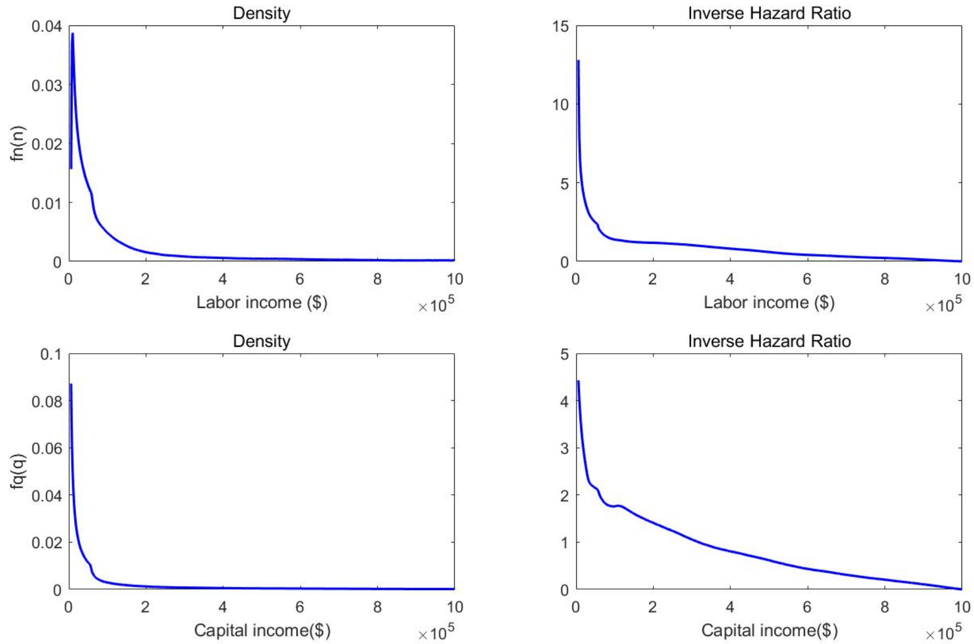

Figure 2: Labor and capital income distribution

The left two panels depict the density across labor income and capital income respectively. One takeaway is that capital income is more concentrated at low income levels, implying capital income is more unequally distributed than labor income (Piketty, 2013)[15]. The right panels reflect the corresponding inverse Hazard Ratios of labor and capital income, which are decreasing with the level of factor incomes. As appeared in the first term of equation (33) and (34), they make an important component of the optimal tax system. Intuitively, the unequal distribution of capital income needs for capital income tax in addition to an optimal nonlinear labor income tax. In the aggregate, the labor share in the 2019 economy is about 52.7%. It indicates the automated technical change to some extent, while Hubmer and Restrepo (2021) has shown that the US labor share declined from a peak of 62% in the 1980s to 55% in 2012, and one of the driver is automating more tasks. Technically, we have two sets of income distribution data in addition to the labor share, thus we are able to calibrate two sets of parameter, B_q and δ_n, and one comparative advantage parameter η. To save space, we leave the calibration details in Appendix G.1. Table 3 summarizes the calibrated parameters.

[15]Saez and Stantcheva (2018) use IRS tax data for 2007 on labor and capital income distributions, they find that the bottom 80% earn essentially zero capital income.

Table 3: List of calibrated parameter values

Description		Value	Target/Source
Preference			
B_q	One source of heterogeneity	vector	Target capital income distribution
ϵ_k	Elasticity of capital supply	0.65	Acemoglu et al. (2020)
ϵ_l	Elasticity of labor supply	0.33	Chetty (2012)
Technology			
δ_n	Comparative advantage across skills	vector	Target labor income distribution
η	Comparative advantage parameter	5.54	Target labor share 52.7%
$\sigma = \frac{1}{1-\rho}$	Substitution elasticity between skills	3	Sachs et al. (2020)
α_n	Automation technology across skills	vector	Target wage distribution
R	Capital rental rate	0.15	Target $K/Y = 3$
Government			
κ	Parameter for redistribution motivation	1	Sachs et al. (2020),Saez (2001)
τ	Parameter for tax function	-3	Heathcote et al. (2017)
ϕ	Parameter of progressivity	0.181	Heathcote et al. (2017)
τ_k	Initinal capital income tax rate	0.1	Acemoglu et al. (2020)

For the individual preference, we borrow the following utility function from Acemoglu et al. (2020) :

$$u(y_q - a_q) = -B_q \frac{a_q^{1+1/\epsilon_k}}{1+1/\epsilon_k} - a_q, \quad v(l_n) = \frac{l_n^{1+1/\epsilon_l}}{1+1/\epsilon_l}.$$

The parameter B_q is calibrated to match the distribution of pretax capital income. ϵ_k and ϵ_l denote the Hicksian capital and labor supply elasticities respectively. We set our baseline capital supply elasticity to 0.65 (Acemoglu et al., 2020), and labor supply elasticity to 0.33 (Chetty, 2012). As for the comparative advantage in Definition 1, $\mu_n(i) = \delta_n \cdot i^\eta$, the parameter δ_n is calibrated to match the distribution of pretax labor income, and η is calibrated to match labor share in the 2019 U.S. economy, that is 52.7%. We take a capital-output ratio of $K/Y = 3$, which implies a rental rate of capital $R = 15\%$, which is greater than $R = 11.5\%$ in the 1980 economy(see Moll et al. (2021)). Turn to the government, we adopt a concave social welfare function $G(V) = \frac{V^{1-\kappa}}{1-\kappa}$, where κ governs the desire for redistribution, as $\kappa = 0$, the government is utilitarian. We set $\kappa = 1$ in our baseline calibration, we also provide sensitivity analyses for alternative $\kappa = 3$. The initial nonlinear labor income tax function is given by $T_0(z) = z - \frac{1-\tau}{1-\phi} z^{1-\phi}$. $\phi = 0.181$ is the rate of progressivity, and τ governs the marginal tax rate in the lower bound (Heathcote et al., 2017). The initial capital income tax rate is set to 10% following Acemoglu et al. (2020). Finally, the elasticity of substitution between skills in production is set to 3 (Sachs et al., 2020), which indicates $\rho = 2/3$. We also consider an alternative value $\rho = 9/10$. Thus, the substitution elasticity in production is 10, for sensitivity analyses.

To gain insights into the intuition behind our model, Figure 3 displays the calibration characterizes along the level of labor income, which can also be indexed by skill or occupation type. The left upper panel shows that the level of automation is almost monotonically decreasing in the skill of occupation, with 0.87 at the lowest skill and 0.24 at the highest skill. The average level of automation is 0.47, which is consistent with the capital income share in the economy. The intuition behind is that the more complex the occupation, the fewer tasks can be replaced by capital, thus the less level of automation. For example, in highly skilled occupations like teachers, doctors and lawyers, more productive tasks are actually thinking, reading, writing and speaking, which can hardly be replaced by robots. In other words, almost all produced tasks in these occupations require labor input. On the contrary, for ordinary workers, many of their tasks such as driving, carrying, building, and assembling can be replaced by robots, therefore, in these occupations, the degree of automation is high. Moll et al. (2021) has adopted the shift-share specification to calibrate the share of automated tasks at each wage percentile, which also finds greater exposure to the automation of routine jobs among workers at the middle and bottom of the wage distribution.

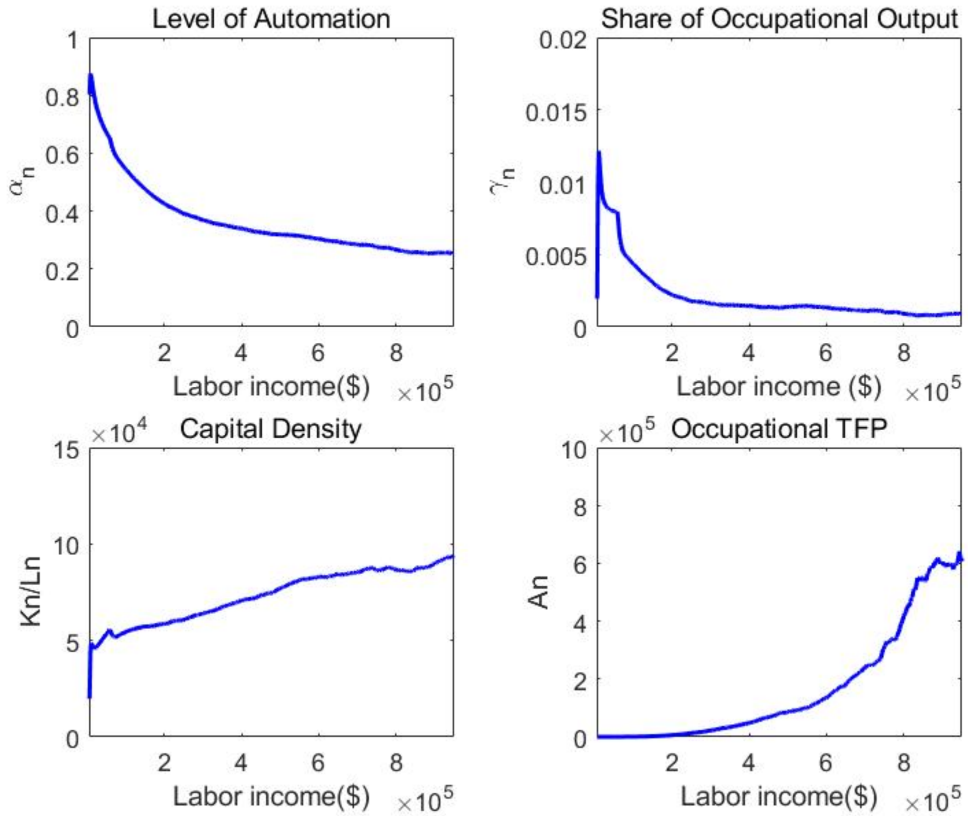

Figure 3: Calibration features across occupation

We also display the output share of each occupation type, γ_n, see the right upper panel of Figure 3. The share of output in occupation with higher incomes is lower, mainly due to there is few of them in the whole populations. However, the middle income group contributes to the most of the total output. In addition, the trend of capital density along occupation, K_n/L_n, is on the contrary of automation, as showed in the lower-left panel. Individuals with high skill leverage more capitals in their occupation, but these capitals are concentrated in a small set of tasks, thus lowering degree of automation.[16] The lower-right panel of Figure 3 shows the trend of occupational total factor productivity (TFP). It is intuitive that occupation with high-skilled individual is more productive than low-skilled individual.

[16]An alternative explanation is that, the level of automation in occupation with skill type n is $\alpha_n = \frac{RK_n}{w_n L_n + RK_n}$, which is increasing with capital density K_n/L_n, but decreasing with w_n/R. When the price driver dominates, it is intuitive that α_n decreases with n as well as K_n/L_n increases with n.

6.2 Simulation

Thanks to the optimal tax formulas, (33)-(35), being expressed in sufficient statistics, we are allowed to simulate optimal tax system following the fixed-point algorithm provided by Mankiw et al. (2009). The simulation details are provided in Appendix G.2. We first consider two alternative tax systems, as depicted, NLIT-NCIT tax system and NLIT-LCIT tax system, which are counterpart to separable nonlinear (SN) and separable linear (SL) tax systems in Ferey et al. (2021), where they study tax on saving. Finally, we investigate optimal capital income tax when suppose the nonlinear labor income is held at the status quo, that is restricting the NLIT to HSV tax functional form(see Heathcote et al. (2017)).

6.2.1 NLIT-NCIT Tax System

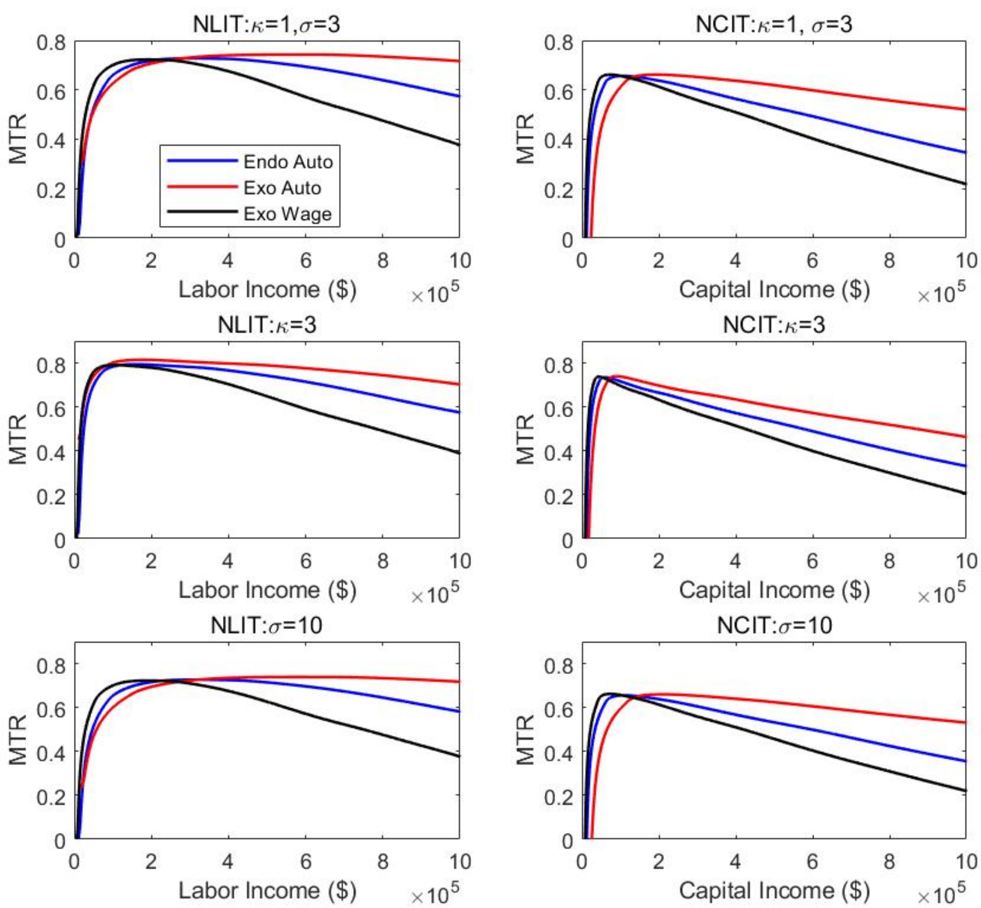

Figure 4: Optimal NLIT-NCIT Tax System

Figure 4 displays the optimal NLIT-NCIT tax system. The upper two panels are our baseline results, i.e., the case of $\kappa = 1$ and $\sigma = 3$. For the optimal nonlinear labor income tax, we find that substitution effect, the case of exogenous automation, makes optimal NLIT more progressive relative to the exogenous-wage setting. That is the marginal labor income tax rate becomes higher on the high income individuals, but the marginal tax rate is lower for the low income individuals. The intuition follows the general equilibrium "trickle-down" forces proposed by Sachs et al. (2020). However, in the case of endogenous automation, the general equilibrium "trickle-down" forces will be moderated, since the automation effect is always on the contrary of substitution effect in our previously analyses (see Section 3). The intuition behind is that, lower tax rate raises the incentive of high individual's labor supply, which can increase their wages through automation effect. Thus, automated technical change makes the optimal NLIT more regressive relative the exogenous-automation setting, but sill more progressive than the exogenous-wage setting, implying that the substitution effect dominates the automation effect.

Turn to the optimal nonlinear capital income tax (NCIT), we find that it shares the similar shape as NLIT. The intuition behind is that the distribution of capital income is unequal, NCIT are desirable for redistribution. The general equilibrium effects also have implications for NCIT. Relative to the exogenous-wage setting, we find that substitution effect makes optimal NCIT more progressive. To some extent, it may stimulate the investment of middle capital income individuals, then increasing the level of wage due to capital-skill complementarity, which can ultimately benefit the government revenue raising from labor income. In addition, automation effect regulates the substitution effect again, it makes the optimal NCIT becomes less progressive relative to the exogenous-automation setting.

We display some robust results in the rest of panels. In the middle two panels, we consider a more redistributive government and set $\kappa = 3$. In the bottom two panels, we consider an alternative case where $\sigma = 10$. Compared to the baseline results, the above intuitions are robust. The caveat is that, the automated technical change is not negligible in the design of optimal tax system. When moved from the initial tax system to optimal NLIT-NCIT, labor share raises to 57.8%.

6.2.2 NLIT-LCIT Tax System

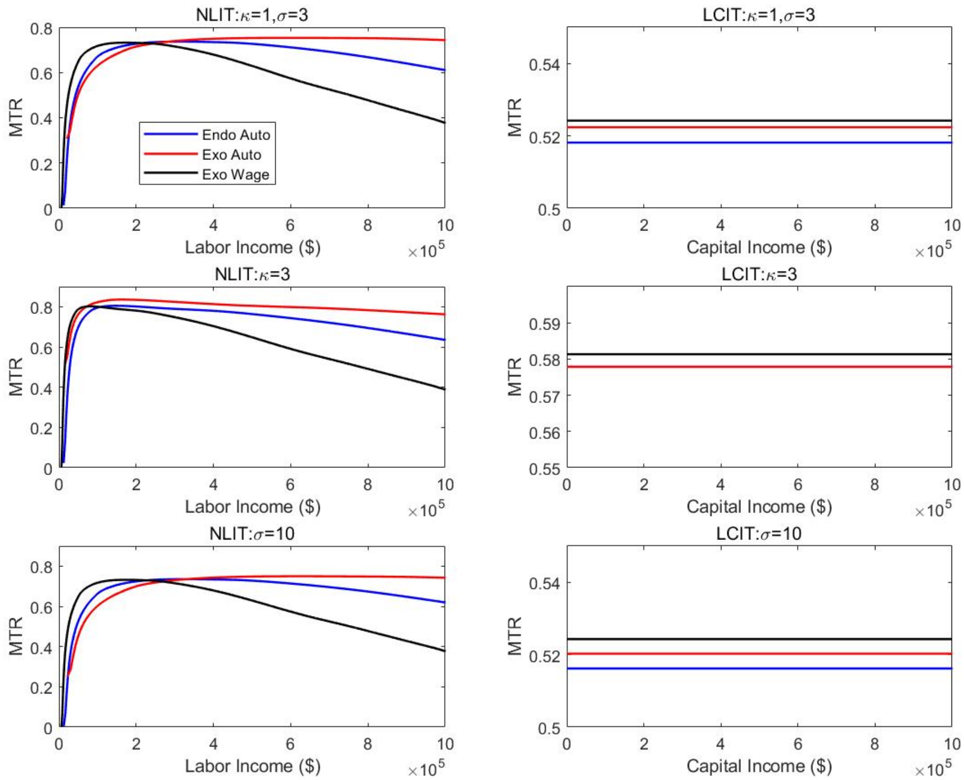

Figure 5: Optimal NLIT-LCIT Tax System

In this subsection, we restrict capital income tax to be linear, which are more implementable in practice, then consider an alternative NLIT-LCIT tax system. The discussion of NLIT follows the same intuition as the NLIT-NCIT tax system. Figure 5 shows that the general equilibrium effect reduces the optimal linear capital income tax rate. The intuition behind is direct, since lower tax rate can stimulate investment, which can increase wages through general equilibrium effect, then benefit the government revenue raising from labor income. The adjustment of optimal LCIT due to automated technical change is light. Moreover, the more redistributive the government (higher κ), the higher the optimal LCIT. Moving to the optimal NLIT-LCIT tax system, labor share in the economy will raise to 58.7%.

6.2.3 NCIT or LCIT Tax System

We now consider a more practical case, by assuming the nonlinear labor income tax follows the functional form of Heathcote et al. (2017), that is $T_z(z) = z - \frac{1-\tau}{1-\phi} z^{1-\phi}$, where $\tau = -3$, and $\phi = 0.181$ governs the progressivity of NLIT. We resimulate the optimal NCIT and LCIT respectively.

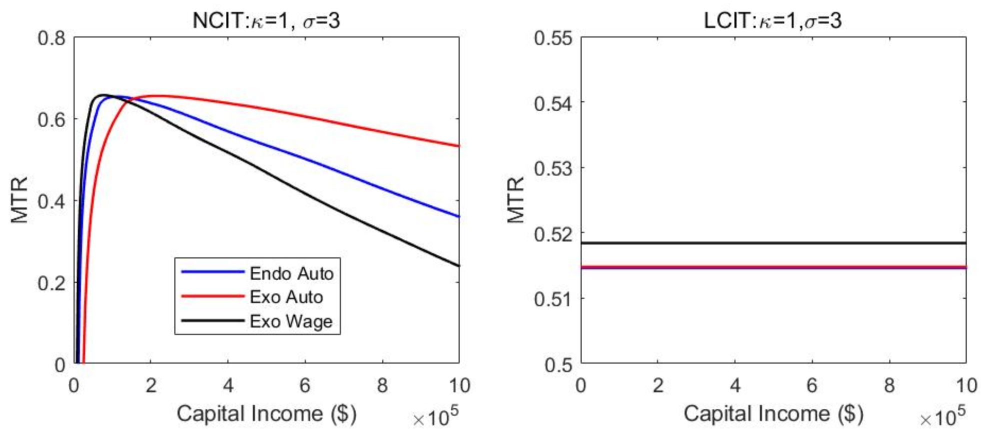

Figure 6: Optimal NLIT or LCIT Tax System

The left panel of Figure 6 displays the optimal NCIT in difference case. Relative to the exogenous-wage setting, the general equilibrium effect makes the optimal NCIT more progressive. In addition, automated technical change plays a central role in governing the progressivity of taxation. Moving from the initial capital tax to the optimal NCIT, labor share raises to 57%. The right panel of Figure 6 displays the optimal LCIT. As anticipated above, optimal LCIT is lower in the general equilibrium framework. Moving to the optimal LCIT, labor share raises to 57.8%.

7 Conclusion

This paper develops a tractable model to investigate tax incidence and optimal income taxation in general equilibrium with multiple factor incomes and automation technology.

Our model indicates that there are two contrary channels through which factor inputs affect factor prices, i.e., substitution effect and automation effect. When taken into account the automated technical change, we find that the real wage stagnation of bottom- and middle-income individuals can go hand-in-hand with capital accumulation, and the responses of wages with respect to labor input are less sensitive than that of capital input. Moreover, capital accumulation induces the adjustment of automation technology, which leads to the increase of wage inequality.

The implications of factor inputs and automation technology for both level and distribution of wages, play a central role in our tax incidence analysis. We first give an integral equation system of labor and capital supply. Then, we find that both labor income tax reform and capital income tax reform have general equilibrium "trickle down" forces on the government revenue. In addition, these tax reforms are equipped with redistributive forces on both labor income and capital income, thus having implications for social welfare through substitution effect and automation effect.

In the context of multiple factor incomes, we derive optimal tax system, NLIT-NCIT and NLIT-LCIT, using variational approach (namely, dual approach). We show that both labor income and capital income tax formulas can be expressed in sufficient statistics. Both substitution effect and automation effect play a central role in the design of optimal tax system. In addition, we prove the equivalence between dual approach and primal approach in the context of dual-tax system.

After calibrating the U.S. economy in 2019, we find that the degree of automation is decreasing with occupational earnings. We then simulate three alternative tax systems. For the NLIT-NCIT tax system, while substitution effect makes both labor income taxation and capital income taxation more progressive relative to the exogenous-wage setting, automation effect regulates the substitution effect and makes the nonlinear tax system less progressive relative to the exogenous-automation setting. When we restrict capital income taxation to be linear, which is the NLIT-LCIT tax system, we find that the general equilibrium "trickle down" forces reduce the optimal capital income tax rate relative to the exogenous-wage setting. However, due to the redistributive forces on capital income, the optimal capital income tax rate is higher than the current capital income tax rate. Finally, we restrict nonlinear labor income tax with constant progressivity, and simulate optimal NCIT and LCIT respectively. Capital income tax reform is still desirable relative the current U.S. tax system. The goal of this paper is not to provide a precisely calculated value of the "correct" optimal marginal tax rates, As an extensive literature points out, there are many drivers that determine the optimal tax. Our major conclusion is that one should not ignore automated technical change when designing the optimal tax system.

This paper can be extended in several respects. First, our methods can be applied to higher multi-dimensional heterogeneity, e.g., more kinds of factor incomes. Second, by introducing heterogeneous returns of capital, or income-shifting between labor and capital income, the practical implications of this theoretical work can become more abundant. Finally, one can study alternative optimal tax system based on our model, such as a linear earning-dependent capital tax. We leave these issues for further studies.

References

Acemoglu, D. and D. Autor (2011): "Skills, tasks and technologies: Implications for employment and earnings," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1043–1171.

Acemoglu, D. and J. Loebbing (2022): "Automation and Polarization," Tech. rep., National Bureau of Economic Research.

Acemoglu, D., A. Manera, and P. Restrepo (2020): "Does the US Tax Code Favor Automation?" Tech. rep., National Bureau of Economic Research.

Acemoglu, D. and P. Restrepo (2018a): "Modeling automation," in *AEA Papers and Proceedings*, vol. 108, 48–53.

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

——— (2020): "Robots and jobs: Evidence from US labor markets," *Journal of Political Economy*, 128, 2188–2244.

——— (2022): "Tasks, automation, and the rise in us wage inequality," *Econometrica*, 90, 1973–2016.

Ales, L., M. Kurnaz, and C. Sleet (2015): "Technical change, wage inequality, and taxes," *American Economic Review*, 105, 3061–3101.

Bergholt, D., F. Furlanetto, and N. Maffei-Faccioli (2022): "The decline of the labor share: new empirical evidence," *American Economic Journal: Macroeconomics*, 14, 163–98.

Blundell, R., D. A. Green, and W. Jin (2022): "The UK as a Technological Follower: Higher Education Expansion and the College Wage Premium," *The Review of Economic Studies*, 89, 142–180.

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

Chetty, R. (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80, 969–1018.

CLEMENS, M. A., E. G. LEWIS, AND H. M. POSTEL (2018): "Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion," *American Economic Review*, 108, 1468–87.

COSTINOT, A. AND J. VOGEL (2010): "Matching and inequality in the world economy," *Journal of Political Economy*, 118, 747–786.

COSTINOT, A. AND I. WERNING (2018): "Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation," *NBER Working Paper*.

CUI, X., L. GONG, AND W. LI (2021): "Supply-side optimal capital taxation with endogenous wage inequality," *Journal of Public Economics*, 198, 104421.

DIAMOND, P. AND E. SAEZ (2011): "The case for a progressive tax: from basic research to policy recommendations," *Journal of Economic Perspectives*, 25, 165–90.

DIAMOND, P. A. (1998): "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates," *American Economic Review*, 83–95.

FEREY, A., B. LOCKWOOD, AND D. TAUBINSKY (2021): "Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity," Tech. rep., National Bureau of Economic Research.

GERRITSEN, A. (2016): "Optimal Nonlinear Taxation: The Dual Approach," *Working Papers*.

GERRITSEN, A., B. JACOBS, A. V. RUSU, AND K. SPIRITUS (2020): "Optimal taxation of capital income with heterogeneous rates of return," .

GOLOSOV, M. AND I. KRASIKOV (2023): "The Optimal Taxation of Couples," Tech. rep., National Bureau of Economic Research.

GUERREIRO, J., S. REBELO, AND P. TELES (2022): "Should robots be taxed?" *The Review of Economic Studies*, 89, 279–311.

HARBERGER, A. C. (1962): "The incidence of the corporation income tax," *Journal of Political economy*, 70, 215–240.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2017): "Optimal Tax Progressivity: An Analytical Framework*," *The Quarterly Journal of Economics*, 132, 1693–1754.

HÉMOUS, D. AND M. OLSEN (2022): "The rise of the machines: Automation, horizontal innovation, and income inequality," *American Economic Journal: Macroeconomics*, 14, 179–223.

HUBMER, J. AND P. RESTREPO (2021): "Not a typical firm: The joint dynamics of firms, labor shares, and capital–labor substitution," Tech. rep., National Bureau of Economic Research.

JACQUET, L. AND E. LEHMANN (2021): "How to tax different Incomes?" *CEPR Discussion Paper No. DP16571*.

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

KIM, H. Y. (2000): "The Antonelli versus Hicks elasticity of complementarity and inverse input demand systems," *Australian Economic Papers*, 39, 245–261.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1053.

LOEBBING, J. (2020): "Redistributive Income Taxation with Directed Technical Change," *CESifo Working Paper*.

MANKIW, N. G., M. WEINZIERL, AND D. YAGAN (2009): "Optimal taxation in theory and practice," *Journal of Economic Perspectives*, 23, 147–74.

MIRRLEES, J. A. (1971): "An exploration in the theory of optimum income taxation," *The Review of Economic Studies*, 38, 175–208.

MOLL, B., L. RACHEL, AND P. RESTREPO (2021): "Uneven growth: automation's impact on income and wealth inequality," Tech. rep., National Bureau of Economic Research.

——— (2022): "Uneven growth: automation's impact on income and wealth inequality," *Econometrica*, 90, 2645–2683.

PIKETTY, T. (2013): "Capital in the 21st Century," *Cambridge, MA: President and Fellows, Harvard College*.

PIKETTY, T., E. SAEZ, AND G. ZUCMAN (2018): "Distributional national accounts: methods and estimates for the United States," *The Quarterly Journal of Economics*, 133, 553–609.

ROTHSCHILD, C. AND F. SCHEUER (2013): "Redistributive taxation in the roy model," *The Quarterly Journal of Economics*, 128, 623–668.

SACHS, D., A. TSYVINSKI, AND N. WERQUIN (2020): "Nonlinear tax incidence and optimal taxation in general equilibrium," *Econometrica*, 88, 469–493.

SAEZ, E. (2001): "Using elasticities to derive optimal income tax rates," *The Review of Economic Studies*, 68, 205–229.

SAEZ, E. AND S. STANTCHEVA (2018): "A simpler theory of optimal capital taxation," *Journal of Public Economics*, 162, 120–142.

SEADE, J. K. (1977): "On the shape of optimal tax schedules," *Journal of public Economics*, 7, 203–235.

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

SPIRITUS, K., E. LEHMANN, S. RENES, AND F. ZOUTMAN (2022): "Optimal Taxation with Multiple Incomes and Types," *NHH Dept. of Business and Management Science Discussion Paper*.

STIGLITZ, J. E. (1982): "Self-selection and Pareto efficient taxation," *Journal of Public Economics*, 17, 213–240.

THUEMMEL, U. (2023): "Optimal taxation of robots," *Journal of the European Economic Association*, 21, 1154–1190.

Appendices

A Macro Aggregation

A.1 Proof of Lemma 1

Proof. Let p_n and $p_n(i)$ denote the price of output Y_n in skill-type n and the price of output $Y_n(i)$ produced within the task i respectively. From the skill output production function,

$$\ln Y_n = \int_0^1 \ln Y_n(i) di,$$

the relationship between skill output and task-level output is $Y_n(i) = p_n Y_n / p_n(i)$. From the task-level production function,

$$Y_n(i) = \begin{cases} \psi_n^k(i) K_n(i) & \text{if } i \in [0, \alpha_n] \\ \psi_n^l(i) L_n(i) & \text{if } i \in (\alpha_n, 1] \end{cases}$$

we obtain that $K_n(i) = Y_n(i)/\psi_n^k(i)$, $R = p_n(i)\psi_n^k(i)$ for $0 \leq i \leq \alpha_n$ and $L_n(i) = Y_n(i)/\psi_n^l(i)$, $w_n = p_n(i)\psi_n^l(i)$ for $\alpha_n \leq i \leq 1$. Combining with the expression of $Y_n(i)$, it is easy to get

$$K_n(i) = \frac{p_n Y_n}{R}, \forall i \in [0, \alpha_n]; \quad L_n(i) = \frac{p_n Y_n}{w_n}, \forall i \in (\alpha_n, 1].$$

Remember that $K_n = \int_0^{\alpha_n} K_n(i) di$ and $L_n = \int_{\alpha_n}^1 L_n(i) di$, thus capital and labor assigned to skill-type n and corresponding task i satisfy the following relationships

$$K_n(i) = \frac{K_n}{\alpha_n}, \quad L_n(i) = \frac{L_n}{1 - \alpha_n} \quad \forall n \in N, i \in [0, 1].$$

Substituting them back into the task-level production function and using the Cobb-Douglas aggregator, we derive the skill output production function,

$$Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}, \quad \text{where} \quad A_n(\alpha_n) = \frac{e^{\int_0^{\alpha_n} \ln \psi_n^k(i) di + \int_{\alpha_n}^1 \ln \psi_n^l(i) di}}{\alpha_n^{\alpha_n}(1-\alpha_n)^{1-\alpha_n}}. \tag{A.1}$$

Without loss of generality, we normalize the price of aggregate output as one. Using the CES production function of aggregate output $Y = \left\{ \int_{\underline{n}}^{\bar{n}} \beta_n Y_n^{\rho} dn \right\}^{1/\rho}$, one obtains the relationship between skill output and aggregate output as $Y_n = \beta_n^{\frac{1}{1-\rho}} p_n^{\frac{1}{\rho-1}} Y$. Since $K_n = \int_0^{\alpha_n} K_n(i)di = \frac{\alpha_n p_n Y_n}{R}$ and $K = \int_{\underline{n}}^{\bar{n}} K_n dn$, the share of capital allocated to skill-type n over aggregate capital can be given as follow,

$$\phi_n(\alpha_n) = \frac{K_n}{K} = \frac{\alpha_n \beta_n^{\frac{1}{1-\rho}} p_n^{\frac{\rho}{\rho-1}}}{\int_{\underline{n}}^{\bar{n}} \alpha_n \beta_n^{\frac{1}{1-\rho}} p_n^{\frac{\rho}{\rho-1}} dn}.$$

Substituting $\phi_n(\alpha_n)$ and $Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}$ into $Y = \left\{ \int_{\underline{n}}^{\bar{n}} \beta_n Y_n^{\rho} dn \right\}^{1/\rho}$, aggregate output could be expressed in the following reduced form:

$$Y \equiv F(K, \mathcal{L}; \boldsymbol{\alpha}) = \left\{ \int_{\underline{n}}^{\bar{n}} \beta_n \left[\tilde{A}_n(\alpha_n) K^{\alpha_n} L_n^{1-\alpha_n} \right]^{\rho} dn \right\}^{1/\rho}, \quad \text{where} \quad \tilde{A}_n(\alpha_n) = A_n(\alpha_n) \phi_n^{\alpha_n}(\alpha_n). \tag{A.2}$$

■

A.2 Proof of Lemma 2

Proof. We know that the skill output function is given as $Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}$, profit maximizing means that factor prices equal their marginal productivity, thus we have the following first-order conditions:

$$w_n = \frac{(1-\alpha_n) p_n Y_n}{L_n}, \quad R = \frac{\alpha_n p_n Y_n}{K_n}.$$

Denote $\gamma_n = p_n Y_n / Y$ as the share of output value produced by skill-type n in the total output value, $\int_{\underline{n}}^{\bar{n}} \gamma_n dn = 1$ since aggregate production function is CES. Moreover, we denote $\alpha = \int_{\underline{n}}^{\bar{n}} \alpha_n \gamma_n dn$ as the average degree of automation in economy. Using $K = \int_{\underline{n}}^{\bar{n}} K_n dn$ and rewrite above two first-order conditions, equilibrium factor prices can be given as follows,

$$w_n(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{\partial F(K, \mathcal{L}; \boldsymbol{\alpha})}{\partial L_n} = \frac{(1-\alpha_n)\gamma_n Y}{L_n}, \quad R(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{\partial F(K, \mathcal{L}; \boldsymbol{\alpha})}{\partial K} = \frac{\alpha Y}{K}. \tag{A.3}$$

As for the equilibrium automation technology, there is an indifference condition $\mu_n(\alpha_n) \equiv \frac{\psi_n^l(\alpha_n)}{\psi_n^k(\alpha_n)} = \frac{w_n}{R}$, such that producing with capital or labor is indifference in task α_n. Combining with (A.3), one obtains

$$\alpha_n = 1 - \frac{\alpha}{\gamma_n} \frac{\mu_n(\alpha_n) L_n}{K}.$$

Remember that $\alpha = \int_{\underline{n}}^{\bar{n}} \alpha_n \gamma_n dn$, so the equilibrium automation can be solved by the following equations:

$$\alpha_n \equiv \alpha_n(K, \mathcal{L}) = 1 - \frac{1}{\gamma_n} \frac{\mu_n(\alpha_n) L_n}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n) L_n dn}, \forall n \in N \quad \alpha \equiv \alpha(K, \mathcal{L}) = \frac{K}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n) L_n dn}. \tag{A.4}$$

We do not give the analytical expression of automation technology, but turn to numerical solution in Section 6. ■

B Proof of Elasticities

B.1 Proof of Supply-side Elasticities.

Proof. First, we consider linear labor and capital income tax system with marginal tax rate t_z and t_x. From equation (1), individuals maximize their utilities according to the following two first-order conditions

$$v'(l_n) = (1 - t_z)\, w_n, \quad u'(y_q - a_q) = 1 + (1 - t_x)R. \tag{B.1}$$

Differentiating above equations with respect to marginal retention rate $1 - t_z$, $1 - t_x$, we obtain

$$v''(l_n) \frac{dl_n}{d(1 - t_z)} = w_n, \quad -u''(y_q - a_q) \frac{da_q}{d(1 - t_x)} = R.$$

Following the definition of elasticity, the behavior elasticities in the context of linear tax system can be derived as follows:

$$e_{l_n, 1-t_z} \equiv -\frac{1 - t_z}{l_n} \frac{dl_n}{dt_z} = \frac{v'(l_n)}{l_n v''(l_n)}, \quad e_{a_q, 1-t_x} \equiv -\frac{a_q}{1 - t_x} \frac{da_q}{dt_x} = -\frac{u'(y_q - a_q) - 1}{a_q u''(y_q - a_q)}.$$

Rearrange and one get

$$v''(l_n) = \frac{v'(l_n)}{l_n e_{l_n, 1-t_z}}, \quad -u''(y_q - a_q) = \frac{u'(y_q - a_q) - 1}{a_q e_{a_q, 1-t_x}}. \tag{B.2}$$

Next we turn to the behavior elasticities under nonlinear tax system. Bear in mind that we have given the form of tax perturbation in equation (7), i.e., $\tilde{T}_i = T_i + \kappa_i \tau_i$ with $i \in [z, x]$, thus as

$\kappa_i \to 0$, the two first-order conditions can be rewritten as follows,

$$\begin{aligned} v'(l_n) &= (1 - T_z'(z_n) - \kappa_z \tau_z'(z_n))w_n, \quad \forall n \in N \\ u'(y_q - a_q) &= 1 + (1 - T_x'(x_q) - \kappa_x \tau_x'(x_q))R, \quad \forall q \in Q. \end{aligned} \tag{B.3}$$

Taking derivative of above equations with respect to κ_z and κ_x respectively, one obtains

$$v''(l_n)\frac{dl_n}{d\kappa_z}|_{\kappa_z=0} = -w_n\left[T_z''(z_n)w_n\frac{dl_n}{d\kappa_z}|_{\kappa_z=0} + \tau_z'(z_n)\right],$$

$$-u''(y_q - a_q)\frac{da_q}{d\kappa_x}|_{\kappa_x=0} = -R\left[T_x''(x_q)R\frac{da_q}{d\kappa_x}|_{\kappa_x=0} + \tau_x'(x_q)\right].$$

Plug (B.2) into above equations,we have

$$\frac{dl_n}{d\kappa_z}|_{\kappa_z=0} = -\frac{\tau_z'(z_n)}{T_z''(z_n)w_n + \frac{1-T_z'(z_n)}{l_n e_{l_n,1-t_z}}}, \quad \frac{da_q}{d\kappa_x}|_{\kappa_x=0} = -\frac{\tau_x'(x_q)}{T_x''(x_q)R + \frac{1-T_x'(x_q)}{a_q e_{a_q,1-t_x}}},$$

where we use $v'(l_n) = (1 - T_z'(z_n))w_n$ and $u'(y_q - a_q) = 1 + (1 - T_x'(x_q))R$. The elasticities of factor supply with respect to marginal retention tax rate can be expressed as follows,

$$\epsilon_{l_n,1-T_z'} \equiv -\frac{1 - T_z'(z_n)}{l_n}\frac{dl_n}{\tau_z'(z_n)d\kappa_z}|_{\kappa_z=0} = \frac{[1 - T_z'(z_n)]\, e_{l_n,1-t_z}}{1 - T_z'(z_n) + e_{l_n,1-t_z}T_z''(z_n)z_n},$$

$$\epsilon_{a_q,1-T_x'} \equiv -\frac{1 - T_x'(x_q)}{a_q}\frac{da_q}{\tau_x'(x_q)d\kappa_x}|_{\kappa_x=0} = \frac{\left[1 - T_x'(x_q)\right] e_{a_q,1-t_x}}{1 - T_x'(x_q) + e_{a_q,1-t_x}T_x''(x_q)x_q}.$$

As for the elasticities of factor supply with respect to price, we take derivative of equation (B.3) about w_n and R at $\kappa_i \to 0$, respectively, then get

$$v''(l_n)\frac{dl_n}{dw_n} = 1 - T_z'(z_n) - T_z''(z_n)\left[l_n + w_n\frac{dl_n}{dw_n}\right]w_n,$$

$$-u''(y_q - a_q)\frac{da_q}{dR} = 1 - T_x'(x_q) - T_x''(x_q)\left[a_q + R\frac{da_q}{dR}\right]R.$$

Following the same techniques, one obtains

$$\frac{dl_n}{dw_n} = \frac{1 - T_z'(z_n) - T_z''(z_n)z_n}{\frac{(1-T_z'(z_n))w_n}{l_n e_{l_n,1-t_z}} + w_n T_z''(z_n)w_n}, \quad \frac{da_q}{dR} = \frac{1 - T_x'(x_q) - T_x''(x_q)x_q}{\frac{(1-T_x'(x_q))R}{a_q e_{a_q,1-t_x}} + RT_x''(x_q)R}.$$

Thus,

$$\epsilon_{l_n,w_n} \equiv \frac{w_n}{l_n}\frac{dl_n}{dw_n} = \frac{\left[1 - T_z'(z_n) - T_z''(z_n)z_n\right] e_{l_n,1-t_z}}{1 - T_z'(z_n) + e_{l_n,1-t_z}T_z''(z_n)z_n},$$

$$\epsilon_{a_q,R} \equiv \frac{R}{a_q}\frac{da_q}{dR} = \frac{\left[1 - T_x'(x_q) - T_x''(x_q)x_q\right] e_{a_q,1-t_x}}{1 - T_x'(x_q) + e_{a_q,1-t_x}T_x''(x_q)x_q}.$$

■

B.2 Proof of Demand-side Elasticities

Proof. Bear in mind that the equilibrium factor prices are given as

$$w_n \equiv w_n(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{(1-\alpha_n)\gamma_n Y}{L_n}, \quad R \equiv R(K, \mathcal{L}; \boldsymbol{\alpha}) = \frac{\alpha Y}{K}.$$

where $\gamma_n = \frac{p_n Y_n}{Y}$, $p_n = \beta_n \left(\frac{Y_n}{Y}\right)^{\rho-1}$, and $Y_n = \tilde{A}_n(\alpha_n)K^{\alpha_n}L_n^{\alpha_n}$. Taking the logarithm on both side of the above equations, one obtains

$$\ln w_n = \ln \beta_n + \ln(1-\alpha_n) + (1-\rho)\ln Y + \rho \ln Y_n - \ln L_n, \quad \ln R = \ln \alpha + \ln Y - \ln K, \quad \text{(B.4)}$$

where $\ln Y_n = \ln \tilde{A}_n(\alpha_n) + \alpha_n \ln K + (1-\alpha_n)\ln L_n$. Next we differentiate equation (B.4) with respect to labor input $\mathcal{L}$ and following the definition of elasticity, we have

$$\epsilon^D_{w_n,L_n} = \frac{d\ln w_n}{d\ln L_n} = \rho(1-\alpha_n) - 1, \quad \epsilon_{w_n,L_{n'}} = \frac{d\ln w_n}{d\ln Y}\frac{d\ln Y}{d\ln L_{n'}} = (1-\rho)(1-\alpha_{n'})\gamma_{n'}, \quad \forall n, n' \in N.$$

where we use $w_{n'}L_{n'} = (1-\alpha_{n'})\gamma_{n'}Y$ and $w_{n'} = dY/dL_{n'}$. Since the price of aggregate out is normalized to one, thus, $\epsilon_{R,L_n} = \frac{d\ln R}{d\ln Y}\frac{d\ln Y}{d\ln L_n} = (1-\alpha_n)\gamma_n$.

As for the elasticities with respect to capital input K, we take derivative of equation (B.4) about K, then get

$$\epsilon_{w_n,K} = \frac{d\ln w_n}{d\ln K} = \frac{d\ln w_n}{d\ln Y}\frac{d\ln Y}{d\ln K} + \frac{d\ln w_n}{d\ln Y_n}\frac{d\ln Y_n}{d\ln K} = (1-\rho)\alpha + \rho\alpha_n,$$

where we use $\frac{RK}{Y} = \alpha$, $R = \frac{dY}{dK}$, and $\alpha = \frac{d\ln Y}{d\ln K}$. Moreover, $\epsilon_{R,K} = \alpha - 1$.

Turn to the elasticities with respect to automation technology $\boldsymbol{\alpha}$, following the same techniques, one could obtain

$$\epsilon_{w_n,\alpha_n} = \frac{d\ln w_n}{d\ln \alpha_n} = -\frac{\alpha_n}{1-\alpha_n}, \quad \epsilon_{R,\alpha} = \frac{d\ln R}{d\ln \alpha} = 1.$$

Here we use envelope conditions, i.e. $\frac{d\ln Y}{d\ln \alpha} = \frac{d\ln Y_n}{d\ln \alpha_n} = 0$ in equilibrium. To elucidate the intuition, one should know that there is no improvement in production at the optimal level of automation technology, i.e., $\frac{dF(K,\mathcal{L},\boldsymbol{\alpha})}{d\boldsymbol{\alpha}} = 0$.

The rest of demand-side elasticities are about automation technology. From lemma 2, we know that

$$\alpha_n \equiv \alpha_n(K,\mathcal{L}) = 1 - \frac{1}{\gamma_n}\frac{\mu_n(\alpha_n)L_n}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn}, \quad \alpha \equiv \alpha(K,\mathcal{L}) = \frac{K}{K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn}, \quad \forall n \in N.$$

Rearrange and take the logarithm, one obtains

$$\ln(1-\alpha_n) = -\ln\beta_n + \ln\mu_n(\alpha_n) + \ln L_n + \rho\ln Y - \rho\ln Y_n - \ln\left(K + \int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn\right), \quad \text{(B.5)}$$

$$\ln\alpha = \ln K - \ln\left(K + \int_{\underline{n}}^{\bar{n}} L_n\mu_n(\alpha_n)dn\right), \quad \text{(B.6)}$$

where we use $\gamma_n = p_nY_n/Y$ and $p_n = \beta_n\left(\frac{Y_n}{Y}\right)^{\rho-1}$. Following the same process, it is easy to show that

$$\epsilon^D_{\alpha_n,L_n} = \frac{d\ln\alpha_n}{d\ln(1-\alpha_n)}\frac{d\ln(1-\alpha_n)}{d\ln L_n} = -\frac{1-\alpha_n}{\alpha_n}(1-\rho(1-\alpha_n)),$$

$$\epsilon_{\alpha_n,L_{n'}} = \frac{d\ln\alpha_n}{d\ln(1-\alpha_n)}\frac{d\ln(1-\alpha_n)}{d\ln L_{n'}} = -\frac{1-\alpha_n}{\alpha_n}\left[\frac{d\ln\mu_n(\alpha_n)}{d\ln L_{n'}} + \rho\frac{d\ln Y}{d\ln L_{n'}} - \frac{d\ln(K+\int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn)}{d\ln L_{n'}}\right].$$

Denote $\epsilon_{\mu_n(\alpha_n),\alpha_n} = \frac{dln\mu_n(\alpha_n)}{dln\alpha_n}$ as the elasticity of comparative advantage with respect to automation. With

$$\begin{aligned}\frac{d\ln(K+\int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn)}{d\ln L_{n'}} &= \frac{L_{n'}}{K+\int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn}\left(\mu_{n'}(\alpha_{n'}) + \int_{\underline{n}}^{\bar{n}} L_n\mu_n'(\alpha_n)\frac{d\alpha_n}{dL_{n'}}dn\right) \\ &= (1-\alpha_{n'})\gamma_{n'} + \frac{\int_{\underline{n}}^{\bar{n}} L_n\mu_n(\alpha_n)\frac{\mu_n'(\alpha_n)\alpha_n}{\mu_n(\alpha_n)}\frac{L_{n'}}{\alpha_n}\frac{d\alpha_n}{dL_{n'}}dn}{K+\int_{\underline{n}}^{\bar{n}} \mu_n(\alpha_n)L_n dn} \\ &= (1-\alpha_{n'})\gamma_{n'} + \int_{\underline{n}}^{\bar{n}} (1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,Ln'}dn,\end{aligned}$$

thus

$$\epsilon_{\alpha_n,L_{n'}} = -\frac{1-\alpha_n}{\alpha_n}\left[\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,L_{n'}} + \rho(1-\alpha_{n'})\gamma_{n'} - (1-\alpha_{n'})\gamma_{n'} - \int_{\underline{n}}^{\bar{n}}(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,Ln'}dn\right].$$

Rearrange then we prove that

$$\epsilon_{\alpha_n,L_{n'}} = \frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'} + \int_{\underline{n}}^{\bar{n}}(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,L_{n'}}dn}{\epsilon_{\mu_n(\alpha_n),\alpha_n} + \alpha_n/(1-\alpha_n)}, \quad \forall n \in N.$$

As for the automation elasticity of capital, we differentiate (B.5) with respect to K, thus

$$\epsilon_{\alpha_n,K} = -\frac{1-\alpha_n}{\alpha_n}\left[\frac{dln\mu_n(\alpha_n)}{dlnK} + \rho\frac{dlnY}{dlnK} - \rho\frac{dlnY_n}{dlnK} - \frac{dln(K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn)}{dlnK}\right].$$

Symmetrically, we have

$$\begin{aligned}\frac{d\ln(K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn)}{d\ln K} &= \frac{K}{K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn}\left(1+\int_{\underline{n}}^{\bar{n}}L_n\mu_n'(\alpha_n)\frac{d\alpha_n}{dK}dn\right)\\ &= \alpha + \frac{\int_{\underline{n}}^{\bar{n}}L_n\mu_n(\alpha_n)\frac{\mu_n'(\alpha_n)\alpha_n}{\mu_n(\alpha_n)}\frac{K}{\alpha_n}\frac{d\alpha_n}{dK}dn}{K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn}\\ &= \alpha + \int_{\underline{n}}^{\bar{n}}(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn,\end{aligned}$$

plug into above equation, one obtains

$$\epsilon_{\alpha_n,K} = -\frac{1-\alpha_n}{\alpha_n}\left[\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K} + \rho\alpha - \rho\alpha_n - \alpha - \int_{\underline{n}}^{\bar{n}}(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn\right].$$

Thus we can prove that

$$\epsilon_{\alpha_n,K} = \frac{(1-\rho)\alpha + \rho\alpha_n + \int(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn}{\epsilon_{\mu_n(\alpha_n),\alpha_n} + \alpha_n/(1-\alpha_n)}, \quad \forall n \in N.$$

Finally, differentiating equation (B.6) with respect to factor inputs, we have

$$\begin{aligned}\epsilon_{\alpha,L_n} = \frac{d\ln\alpha}{d\ln L_n} &= -\frac{d\ln(K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn)}{d\ln L_n}\\ &= -(1-\alpha_n)\gamma_n - \int_{\underline{n}}^{\bar{n}}(1-\alpha_{n'})\gamma_{n'}\epsilon_{\mu_{n'}(\alpha_{n'}),\alpha_{n'}}\epsilon_{\alpha_{n'},L_n}dn', \quad \forall n \in N,\end{aligned}$$

and

$$\epsilon_{\alpha,K} = \frac{d\ln\alpha}{d\ln K} = 1 - \frac{d\ln(K+\int_{\underline{n}}^{\bar{n}}\mu_n(\alpha_n)L_n dn)}{d\ln K} = 1 - \alpha - \int_{\underline{n}}^{\bar{n}}(1-\alpha_n)\gamma_n\epsilon_{\mu_n(\alpha_n),\alpha_n}\epsilon_{\alpha_n,K}dn.$$

■

C Proof of Automated Technical Change

C.1 Proof of Section 3.1

Proof. Under the assumption of $0 < \rho < 1$ and $0 < \alpha_n < 1, \forall n \in N$, the demand-side elasticities in Table 2 except $\epsilon_{\alpha_n,L_{n'}}$,$\epsilon_{\alpha_n,K}$, ϵ_{α,L_n} and $\epsilon_{\alpha,K}$ can be signed directly, we do not go into details. Bear in mind that Definition 1 means the elasticity of comparative advantage with respect to automation is constant in equilibrium, i.e., $\epsilon_{\mu_n(\alpha_n),\alpha_n} = \eta > 0$, thus $\epsilon_{\alpha_n,L_{n'}}$ and $\epsilon_{\alpha_n,K}$ can be rewritten as follows,

$$\epsilon_{\alpha_n,L_{n'}} = \frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'} + \int (1-\alpha_n)\gamma_n \eta \epsilon_{\alpha_n,L_{n'}} dn}{\eta + \alpha_n/(1-\alpha_n)},$$

$$\epsilon_{\alpha_n,K} = \frac{(1-\rho)\alpha + \rho\alpha_n + \int (1-\alpha_n)\gamma_n \eta \epsilon_{\alpha_n,K} dn}{\eta + \alpha_n/(1-\alpha_n)}.$$

Multiply $(1-\alpha_n)\gamma_n$ on both sides, one obtains

$$(1-\alpha_n)\gamma_n\epsilon_{\alpha_n,L_{n'}} = \frac{(1-\alpha_n)\gamma_n}{\eta + \alpha_n/(1-\alpha_n)}(1-\rho)(1-\alpha_{n'})\gamma_{n'} + \frac{\eta(1-\alpha_n)\gamma_n}{\eta + \alpha_n/(1-\alpha_n)} \int (1-\alpha_n)\gamma_n\epsilon_{\alpha_n,L_{n'}} dn,$$

$$(1-\alpha_n)\gamma_n\epsilon_{\alpha_n,K} = \frac{(1-\alpha_n)\gamma_n}{\eta + \alpha_n/(1-\alpha_n)}((1-\rho)\alpha + \rho\alpha_n) + \frac{\eta(1-\alpha_n)\gamma_n}{\eta + \alpha_n/(1-\alpha_n)} \int (1-\alpha_n)\gamma_n\epsilon_{\alpha_n,K} dn.$$

Integrate both sides and rearrange, we have

$$\int (1-\alpha_n)\gamma_n\epsilon_{\alpha_n,L_{n'}} dn = \frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'} \int \frac{(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}{1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn},$$

$$\int (1-\alpha_n)\gamma_n\epsilon_{\alpha_n,K} dn = \frac{(1-\rho)\alpha \int \frac{(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn + \rho \int \frac{\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}{1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}.$$

Substitute them back into above equations, $\epsilon_{\alpha_n,L_{n'}}$ and $\epsilon_{\alpha_n,K}$ can be given in the following forms,

$$\epsilon_{\alpha_n,L_{n'}} = \frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'}}{[\eta + \alpha_n/(1-\alpha_n)]\left[1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn\right]} > 0, \tag{C.1}$$

$$\epsilon_{\alpha_n,K} = \frac{(1-\rho)\alpha + \rho\alpha_n \left(1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn\right) + \rho \int \frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}{[\eta + \alpha_n/(1-\alpha_n)] \left[1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn\right]} > 0. \tag{C.2}$$

In the last inequality, we use $1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn > 1 - \int (1-\alpha_n)\gamma_n dn = \alpha > 0$.

Next we turn to the signs of ϵ_{α,L_n} and $\epsilon_{\alpha,K}$. In appendix B.2, we have proven that

$$\epsilon_{\alpha,L_n} = -(1-\alpha_n)\gamma_n - \int_{\underline{n}}^{\bar{n}} (1-\alpha_{n'})\gamma_{n'}\eta\epsilon_{\alpha_{n'},L_n} dn', \quad \forall n \in N.$$

it is obviously that $\epsilon_{\alpha,L_n} < 0$ since $\epsilon_{\alpha_n,L_{n'}} > 0$ for all $n, n' \in N$.

With $\epsilon_{\alpha_n,K} = \frac{(1-\rho)\alpha+\rho\alpha_n+\int(1-\alpha_n)\gamma_n\eta\epsilon_{\alpha_n,K}dn}{\eta+\alpha_n/(1-\alpha_n)}$, eliminating the integral term, then $\epsilon_{\alpha,K}$ can be rewritten as follow,

$$\epsilon_{\alpha,K} = 1 - \alpha + (1-\rho)\alpha + \rho\alpha_n - \left(\eta + \frac{\alpha_n}{1-\alpha_n}\right)\epsilon_{\alpha_n,K}.$$

Substituting equation (C.2), one would get

$$\begin{aligned}
\epsilon_{\alpha,K} &= 1 - \alpha + (1-\rho)\alpha - \frac{(1-\rho)\alpha + \rho \int \frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}{1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn} \\
&= 1 - \alpha - \frac{(1-\rho)\alpha \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn + \rho \int \frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn}{1 - \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn} \\
&> 1 - \alpha - \frac{(1-\rho)\alpha(1-\alpha) + \rho\alpha(1-\alpha)}{\alpha} = 0.
\end{aligned}$$

In the last inequality, we use $\int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn < \int (1-\alpha_n)\gamma_n dn = 1 - \alpha$, and $\int \frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn < \int \alpha_n(1-\alpha_n)\gamma_n dn < \alpha(1-\alpha)$.[17] ■

C.2 Proof of Proposition 2

Proof. From equation (9), we know that,

$$\frac{dw_n}{w_n}|_{AE} = \epsilon_{w_n,\alpha_n} \left[\epsilon^D_{\alpha_n,L_n} \frac{dL_n}{L_n} + \int_{\underline{n}}^{\bar{n}} \epsilon_{\alpha_n,L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{\alpha_n,K} \frac{dK}{K}\right], \quad \forall n.$$

[17] Note that $E[\alpha_n(1-\alpha_n)] = E[\alpha_n]E[1-\alpha_n] + cov[\alpha_n, 1-\alpha_n]$, obviously, the covariance between α_n and $1-\alpha_n$ is negative.

For the sake of brevity, we denote $\tilde{\eta} = \int \frac{\eta(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn$ and $\varphi = (1-\rho)\alpha + \rho \int \frac{\eta\alpha_n(1-\alpha_n)\gamma_n}{\eta+\alpha_n/(1-\alpha_n)} dn$. Appendix C.1 shows that $0 < \tilde{\eta} < 1$ and $\varphi > 0$, thus $\epsilon_{\alpha_n, L_{n'}}$ and $\epsilon_{\alpha_n, K}$ are reduced to

$$\epsilon_{\alpha_n, L_{n'}} = \frac{(1-\rho)(1-\alpha_{n'})\gamma_{n'}}{(\eta + \alpha_n/(1-\alpha_n))(1-\tilde{\eta})}, \quad \epsilon_{\alpha_n, K} = \frac{\rho\alpha_n(1-\tilde{\eta}) + \varphi}{(\eta + \alpha_n/(1-\alpha_n))(1-\tilde{\eta})}.$$

For any $n, \tilde{n} \in N$ and $n > \tilde{n}$, we substitute these expressions of elasticities into above equation and subtract, one obtains,

$$\begin{aligned}
\frac{dw_n}{w_n}|_{AE} - \frac{dw_{\tilde{n}}}{w_{\tilde{n}}}|_{AE} = {} & [1-\rho(1-\alpha_n)] \frac{dL_n}{L_n} - [1-\rho(1-\alpha_{\tilde{n}})] \frac{dL_{\tilde{n}}}{L_{\tilde{n}}} \\
& + \left[\frac{1-\rho}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}} + 1} - \frac{1-\rho}{\eta(1-\alpha_n)/\alpha_n + 1} \right] \frac{1}{1-\tilde{\eta}} \int (1-\alpha_{n'})\gamma_{n'} \frac{dL_{n'}}{L_{n'}} dn' \\
& + \left[\frac{\rho(1-\tilde{\eta})\alpha_{\tilde{n}} + \varphi}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}} + 1} - \frac{\rho(1-\tilde{\eta})\alpha_n + \varphi}{\eta(1-\alpha_n)/\alpha_n + 1} \right] \frac{1}{1-\tilde{\eta}} \frac{dK}{K}, \quad \forall n > \tilde{n}.
\end{aligned}$$

Under the conditions of $dL_n/L_n = dl_{n'}/L_{n'} = dL/L > 0$, $dK/K > 0$, $0 < \rho < 1$ and $\dot{\alpha}_n < 0$, it is easy to find that

$$[1-\rho(1-\alpha_n)] \frac{dL_n}{L_n} < [1-\rho(1-\alpha_{\tilde{n}})] \frac{dL_{\tilde{n}}}{L_{\tilde{n}}},$$

$$\frac{1-\rho}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}} + 1} > \frac{1-\rho}{\eta(1-\alpha_n)/\alpha_n + 1}, \quad \frac{\rho(1-\tilde{\eta})\alpha_{\tilde{n}} + \varphi}{\eta(1-\alpha_{\tilde{n}})/\alpha_{\tilde{n}} + 1} > \frac{\rho(1-\tilde{\eta})\alpha_n + \varphi}{\eta(1-\alpha_n)/\alpha_n + 1}.$$

Thus for automation effect, labor inputs decrease wage premium directly but increase wage premium indirectly, while capital input always increase wage premium. As for substitution effect, equation (11) shows that,

$$\frac{dw_n}{w_n}|_{SE} = \epsilon^D_{w_n, L_n} \frac{dL_n}{L_n} + \int_{\underline{n}}^{\tilde{n}} \epsilon_{w_n, L_{n'}} \frac{dL_{n'}}{L_{n'}} dn' + \epsilon_{w_n, K} \frac{dK}{K}, \quad \forall n.$$

Subtract between n and $\tilde{n}$ and leverage the demand-side elasticities in Table 2, the following equation can be deduced,

$$\frac{dw_n}{w_n}|_{SE} - \frac{dw_{\tilde{n}}}{w_{\tilde{n}}}|_{SE} = [\rho(1-\alpha_n) - 1] \frac{dL_n}{L_n} - [\rho(1-\alpha_{\tilde{n}}) - 1] \frac{dL_{\tilde{n}}}{L_{\tilde{n}}} + \rho(\alpha_n - \alpha_{\tilde{n}}) \frac{dK}{K} \quad \forall n > \tilde{n}.$$

Under the same conditions, we have

$$[\rho(1-\alpha_n) - 1] \frac{dL_n}{L_n} > [\rho(1-\alpha_{\tilde{n}}) - 1] \frac{dL_{\tilde{n}}}{L_{\tilde{n}}}, \quad \rho(\alpha_n - \alpha_{\tilde{n}}) \frac{dK}{K} < 0.$$

Thus, labor inputs increase wage premium directly through substitution effect, while capital input decrease it. ■

D Tax Incidence Analysis

D.1 Proof of Proposition 3

Proof. In equilibrium, as labor market clear, i.e., $l_n = L_n$, equation (16) can be rewritten as follow directly,

$$\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T_z'}\frac{\tau_z'(z_n)}{1-T_z'(z_n)} + \epsilon_{l_n,w_n}\int \epsilon_{w_n,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn' + \epsilon_{l_n,w_n}\epsilon_{w_n,K}^{total}\frac{dK}{K}, \tag{D.1}$$

where we denote $\epsilon_{w_n,L_{n'}}^{total} = \epsilon_{w_n,L_{n'}} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,L_{n'}}$ and $\epsilon_{w_n,K}^{total} = \epsilon_{w_n,K} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,K}$. The rest of work is to eliminate dK/K. Note that capital market clearing means $K = \int_N \int_Q a_q f(n,q)dqdn = \int_Q a_q f_q(q)dq$, where $f_q(q) = \int_N f(n,q)dn$. It is easily to derive that

$$\frac{dK}{K} = \int \omega_q \frac{da_q}{a_q}dq, \quad \text{with} \quad \omega_q = \frac{a_q f_q(q)}{K}.$$

Substitute the expression of da_q/a_q in equation (16) into above equation, one obtains

$$\frac{dK}{K} = -\int \omega_q \epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)} + \int \omega_q \epsilon_{a_q,R}dq \int \epsilon_{R,L_{n'}}^{total}\frac{dL_{n'}}{L_{n'}}dn' + \int \omega_q \epsilon_{a_q,R}dq\,\epsilon_{R,K}^{total}\frac{dK}{K}.$$

Denote $\bar{\epsilon}_{K,R} = \int \omega_q \epsilon_{a_q,R}dq$ and $\chi = \frac{1}{1-\bar{\epsilon}_{K,R}\epsilon_{R,K}^{total}}$. Rearrange above equation, the change in aggregate capital can be reduced as

$$\frac{dK}{K} = -\chi \int \omega_q \epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)}dq + \chi\bar{\epsilon}_{K,R}\int \epsilon_{R,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn'.$$

That is equation (19). Substitute it back into equation (D.1), the integral equation of labor supply can be reduced as follow

$$\frac{dl_n}{l_n} = \underbrace{-\epsilon_{l_n,1-T_z'}\frac{\tau_z'(z_n)}{1-T_z'(z_n)}}_{DE}\underbrace{-\epsilon_{l_n,w_n}\epsilon_{w_n,K}^{total}\chi\int \omega_q \epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)}dq}_{CE}$$
$$\underbrace{+\epsilon_{l_n,w_n}\int\left[\epsilon_{w_n,L_{n'}}^{total} + \epsilon_{w_n,K}^{total}\chi\bar{\epsilon}_{K,R}\epsilon_{R,L_{n'}}^{total}\right]\frac{dl_{n'}}{l_{n'}}dn'}_{GE}.$$

■

D.2 Proof of Proposition 4

Proof. Solve the individual utility maximization problem (1), one obtains the following indirect utility function,

$$V(n,q) = u(y_q - a_q) + w_n l_n + (1+R)a_q - T_z(w_n l_n) - T_x(Ra_q) - v(l_n). \tag{D.2}$$

Moreover, government revenue is given by

$$B = \int\int \left[T_z(w_n l_n) + T_x(Ra_q)\right] f(n,q) dndq. \tag{D.3}$$

With the definition of social welfare function $W = \frac{1}{\lambda} G\left(\{V(n,q)\}_{n\times q \in N\times Q}\right) + B$ and social welfare weight $g(n,q) = \frac{1}{\lambda}\frac{dG}{dV(n,q)}$, the change in social welfare can be given by

$$dW = \int\int g(n,q) dV(n,q) f(n,q) dndq + dB. \tag{D.4}$$

In the following process, we show that how individual utility $V(n,q)$ and government revenue B response to a given tax reform (τ_z, τ_x), through mechanical effect (ME), behavior effect (BE), substitution effect (SE) and automation effect (AE), respectively. Bear in mind that tax reform leads to the change in factor supplies and factor prices, i.e., $\frac{dl_n}{l_n}$, $\frac{da_q}{q_q}$, $\frac{dw_n}{w_n}$ and $\frac{dR}{R}$, armed with equation (D.2), it is easily to deduce that

$$\begin{aligned} dV(n,q) = &\underbrace{-\tau_z(z_n) - \tau_x(x_q)}_{ME} \\ &+ \underbrace{\left[w_n - T_z'(z_n)w_n - v'(l_n)\right] l_n(-\epsilon_{l_n,1-T_z'})\frac{\tau_z'(z_n)}{1-T_z'(z_n)} + \left[1 + R - T_x'(x_q)R - u'(y_q - a_q)\right] a_q(-\epsilon_{a_q,1-T_x'})\frac{\tau_x'(x_q)}{1-T_x'(x_q)}}_{BE} \\ &+ \underbrace{(1-T_z'(z_n))z_n\frac{dw_n}{w_n}|_{SE} + (1-T_x'(x_q))x_q\frac{dR}{R}|_{SE}}_{SE} + \underbrace{(1-T_z'(z_n))z_n\frac{dw_n}{w_n}|_{AE} + (1-T_x'(x_q))x_q\frac{dR}{R}|_{AE}}_{AE}. \end{aligned}$$

Note that individual first-order conditions imply $w_n - T_z'(z_n)w_n - v'(l_n) = 0$ and $1 + R - T_x'(x_q)R - u'(y_q - a_q) = 0$, thus $dV(n,q)$ can be reduced to

$$
\begin{aligned}
dV(n,q) = & \underbrace{-\tau_z(z_n) - \tau_x(x_q)}_{ME} + \underbrace{(1 - T_z'(z_n))z_n \frac{dw_n}{w_n}|_{SE} + (1 - T_x'(x_q))x_q \frac{dR}{R}|_{SE}}_{SE} \\
& + \underbrace{(1 - T_z'(z_n))z_n \frac{dw_n}{w_n}|_{AE} + (1 - T_x'(x_q))x_q \frac{dR}{R}|_{AE}}_{AE}.
\end{aligned}
\tag{D.5}
$$

Follow the same logic, armed with equation (D.3), the change in government revenue can be given by

$$
\begin{aligned}
dB = & \underbrace{\int\int \left[\tau_z(z_n) + \tau_x(x_q)\right] f(n,q)dndq}_{ME} \\
& + \underbrace{\int\int T_z'(z_n)z_n(-\epsilon_{l_n,1-T_z'})\frac{\tau_z'(z_n)}{1 - T_z'(z_n)}f(n,q)dndq + \int\int T_x'(x_q)x_q(-\epsilon_{a_q,1-T_x'})\frac{\tau_x'(x_q)}{1 - T_x'(x_q)}f(n,q)dndq}_{BE} \\
& + \underbrace{\int\int T_z'(z_n)z_n\left[1 + \epsilon_{l_n,w_n}\right]\frac{dw_n}{w_n}|_{SE}f(n,q)dndq + \int\int T_x'(x_q)x_q\left[1 + \epsilon_{a_q,R}\right]\frac{dR}{R}|_{SE}f(n,q)dndq}_{SE} \\
& + \underbrace{\int\int T_z'(z_n)z_n\left[1 + \epsilon_{l_n,w_n}\right]\frac{dw_n}{w_n}|_{AE}f(n,q)dndq + \int\int T_x'(x_q)x_q\left[1 + \epsilon_{a_q,R}\right]\frac{dR}{R}|_{AE}f(n,q)dndq}_{AE}.
\end{aligned}
\tag{D.6}
$$

To simply our exposition, we denote $g_n(n) = \frac{\int_Q g(n,q)f(n,q)dq}{f_n(n)}$ and $g_q(q) = \frac{\int_N g(n,q)f(n,q)dn}{f_q(q)}$, note that we have defined $f_n(n) = \int_Q f(n,q)dq$ and $f_q(q) = \int_N f(n,q)dn$. Substitute equation (D.5) and (D.6) into (D.4), then the incidence of tax reform on social welfare can be given by equation (25). ■

E Variational Approach for Optimal Taxation

E.1 Proof of Lemma 4

Proof. Shutting down the general equilibrium feedback requires that $da_q/a_q = 0$. As $K = \int a_q f_q(q) dq$, it is easily to verify that $dK/K = 0$. Remember that

$$\frac{da_q}{a_q} = -\epsilon_{a_q,1-T'_x} \frac{\tau'_x(x_q)}{1-T'_x(x_q)} + \epsilon_{a_q,R} \int \epsilon^{total}_{R,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn' + \epsilon_{a_q,R} \epsilon^{total}_{R,K} \frac{dK}{K}, \tag{E.1}$$

thus we have

$$\tau'_x(x_q) = \frac{(1-T'_x(x_q))\epsilon_{a_q,R}}{\epsilon_{a_q,1-T'_x}} \int \epsilon^{total}_{R,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn'.$$

Substitute the expressions of $\epsilon_{a_n,R}$ and $\epsilon_{a_n,1-T'_x}$, which have been summarized in table 1, into above equation, one obtains

$$\tau'_x(x_q) = (1 - T'_x(x_q) - T''_x(x_q)x_q) \int \epsilon^{total}_{R,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn'.$$

Integrating across x_q, the counteracting tax perturbation on capital income is given by

$$\tau_x(x_q) = (1 - T'_x(x_q))x_q \int \epsilon^{total}_{R,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn'.$$

Turn to the counteracting perturbation τ^2_z, as the integral equation of labor supply is given by

$$\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T'_z} \frac{\tau'_z(z_n)}{1-T'_z(z_n)} - \epsilon_{l_n,w_n} \epsilon^{total}_{w_n,K} \cdot \chi \bar{\epsilon}_{K,R} \int \epsilon^{total}_{R,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn' + \epsilon_{l_n,w_n} \int \zeta_{w_n,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn'.$$

Replacing $\frac{dl_n}{d_n} = -\epsilon_{l_n,1-T'_z} \frac{\tau^{1'}_z(z_n)}{1-T'_z(z_n)}$ and $\tau'_z(z_n) = \tau^{1'}_z(z_n) + \tau^{2'}_z(z_n)$, one can obtain

$$-\epsilon_{l_n,1-T'_z} \frac{\tau^{1'}_z(z_n)}{1-T'_z(z_n)} = -\epsilon_{l_n,1-T'_z} \frac{\tau^{1'}_z(z_n) + \tau^{2'}_z(z_n)}{1-T'_z(z_n)} + \epsilon_{l_n,w_n} \epsilon^{total}_{w_n,K} \cdot \chi \bar{\epsilon}_{K,R} \int \epsilon^{total}_{R,L_{n'}} \epsilon_{l_{n'},1-T'_z} \frac{\tau^{1'}_z(z_{n'})}{1-T'_z(z_{n'})} dn'$$
$$- \epsilon_{l_n,w_n} \int \zeta_{w_n,L_{n'}} \epsilon_{l_{n'},1-T'_z} \frac{\tau^{1'}_z(z_{n'})}{1-T'_z(z_{n'})} dn'.$$

Rearrange and use the definition $\tau_z^{1'}(z_n) = \delta_{z_{n^*}}(z_n)$, one obtains

$$\begin{aligned}\tau_z^{2'}(z_n) &= \frac{(1-T_z'(z_n))\epsilon_{l_n,w_n}}{\epsilon_{l_n,1-T_z'}}\left(\epsilon_{w_n,K}^{total}\chi\bar{\epsilon}_{K,R}\epsilon_{R,L_{n^*}}^{total} - \zeta_{w_n,L_{n^*}}\right)\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1} \\ &= -\epsilon_{w_n,L_{n^*}}^{total}\frac{(1-T_z'(z_n))\epsilon_{l_n,w_n}}{\epsilon_{l_n,1-T_z'}}\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1},\end{aligned}$$

where we use $\zeta_{w_n,L_{n'}} = \epsilon_{w_n,L_{n'}}^{total} + \epsilon_{w_n,K}^{total}\chi\bar{\epsilon}_{K,R}\epsilon_{R,L_{n'}}^{total}$ for any $n \in N$. Remember that, in Table 1, we have

$$\epsilon_{l_n,w_n} = \frac{[1-T_z'(z_n) - T_z''(z_n)z_n]\, e_{l_n,1-t_z}}{1-T_z'(z_n) + e_{l_n,1-t_z}T_z''(z_n)z_n}, \quad \epsilon_{l_n,1-T_z'} = \frac{[1-T_z'(z_n)]\, e_{l_n,1-t_z}}{1-T_z'(z_n) + e_{l_n,1-t_z}T_z''(z_n)z_n}.$$

Substitute them into above equation and integrate across z_n, $\tau_z^2(z_n)$ can be given by

$$\tau_z^2(z_n) = -\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z^*)}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int_0^{z_n}\tilde{\epsilon}_{w,L_{n^*}}^{total}[1-T_z'(z) - T_z''(z)z]dz.$$

Finally, we are going to show how these perturbations affect wages and rental rate. In section 4.2, we investigate the channels through which tax perturbation affects wages and rental rate, with elasticities defined previously, the implications of tax perturbation for them have be reduced as follows (i.e., equation (21) and (22))

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total}\chi\int\omega_q\epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)}dq + \int\left[\epsilon_{w_n,L_{n'}}^{total} + \epsilon_{w_n,K}^{total}\chi\bar{\epsilon}_{K,R}\epsilon_{R,L_{n'}}^{total}\right]\frac{dl_{n'}}{l_{n'}}dn',$$

$$\frac{dR}{R} = -\epsilon_{R,K}^{total}\chi\int\omega_q\epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)}dq + \int\left[\epsilon_{R,L_{n'}}^{total} + \epsilon_{R,K}^{total}\chi\bar{\epsilon}_{K,R}\epsilon_{R,L_{n'}}^{total}\right]\frac{dl_{n'}}{l_{n'}}dn'.$$

As $da_q/a_q = dK/K = 0$, integrate equation (E.1), we know that

$$\int\omega_q\epsilon_{a_q,1-T_x'}\frac{\tau_x'(x_q)}{1-T_x'(x_q)}dq = \bar{\epsilon}_{K,R}\int\epsilon_{R,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn'.$$

Substitute it into above equations, we have

$$\frac{dw_n}{w_n} = \int\epsilon_{w_n,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn', \quad \frac{dR}{R} = \int\epsilon_{R,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn'.$$

Bear in mind that $\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T'_z} \frac{\tau_z^{1'}(z_n)}{1-T'_z(z_n)}$, thus, with the tax perturbations τ_z^1, τ_z^2 and τ_x defined as above, the incidence of tax reform on factor prices can be given as follows

$$\frac{dw_n}{w_n} = -\epsilon^{total}_{w_n,L_{n^*}} \frac{\epsilon_{l_{n^*},1-T'_{z^*}}}{1-T'_z(z_{n^*})} \left(\frac{dz_{n^*}}{dn^*}\right)^{-1}, \quad \frac{dR}{R} = -\epsilon^{total}_{R,L_{n^*}} \frac{\epsilon_{l_{n^*},1-T'_{z^*}}}{1-T'_z(z_{n^*})} \left(\frac{dz_{n^*}}{dn^*}\right)^{-1}.$$

■

E.2 Proof of Proposition 5

Proof. The key idea is that when the tax system is at optimal, there should be no marginal improvement on social welfare given any tax perturbation. Normally,

$$dW = \int\int g(n,q)f(n,q)dV(n,q)dndq + dB = 0.$$

We know that the adjustment of individual utility $V(n,q)$ and government revenue B actually come from the adjustments of factor supplies and factor prices. In Appendix D.2, we have shown that

$$dV(n,q) = z_n(1-T'_z(z_n))\frac{dw_n}{w_n} + x_q(1-T'_x(x_q))\frac{dR}{R} - \tau_z(z_n) - \tau_x(x_q),$$

$$\begin{aligned} dB &= \int\int d[T_z(z_n) + T_x(x_q)]f(n,q)dndq \\ &= \int\int \left[T'_z(z_n)z_n\left(\frac{dw_n}{w_n} + \frac{dl_n}{l_n}\right) + \tau_z(z_n) + T'_x(x_q)x_q\left(\frac{dR}{R} + \frac{da_q}{a_q}\right) + \tau_x(x_q)\right] f(n,q)dndq. \end{aligned}$$

In Appendix E.1, we have derived the expressions of these adjustments, dw_n/w_n, dR/R, dl_n/l_n and da_q/a_q, under certain tax perturbation $\tau_z = \tau_z^1 + \tau_z^2$ and τ_x. The rest of work is to substitute them into above equations, we display the calculation in turn. For the incidence of tax perturbation on individual utility, we have

$$
\begin{aligned}
dV_n =& z_n(1-T_z^{'}(z_n))\frac{dw_n}{w_n} + x_q(1-T_x^{'}(x_q))\frac{dR}{R} - \tau_z^1(z_n) - \tau_z^2(z_n) - \tau_x(x_q) \\
=& z_n(1-T_z^{'}(z_n))\int \epsilon_{w_n,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn' + x_q(1-T_x^{'}(x_q))\int \epsilon_{R,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn' \\
& - \tau_z^1(z_n) - \tau_z^2(z_n) - (1-T_x^{'}(x_q))x_q\int \epsilon_{R,L_{n'}}^{total}\frac{dl_{n'}}{l_{n'}}dn' \\
=& -\frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\epsilon_{w_n,L_{n^*}}^{total}z_n(1-T_z^{'}(z_n)) \\
& - \mathbf{I}_{z_n\geq z_{n^*}} + \frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int_0^{z_n}\epsilon_{w,L_{n^*}}^{total}[1-T_z^{\prime}(z)-T_z^{\prime\prime}(z)z]dz.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\int\int g(n,q)f(n,q)dV(n,q)dndq \\
&= -\frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\int \epsilon_{w_n,L_{n^*}}^{total}z_n(1-T_z^{'}(z_n))g(n,q)f(n,q)dndq - \int\int_{n^*}g(n,q)f(n,q)dndq \\
&\quad + \frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\int\int_0^{z_n}\epsilon_{w,L_{n^*}}^{total}[1-T_z^{\prime}(z)-T_z^{\prime\prime}(z)z]dzg(n,q)f(n,q)dndq \\
&= -\frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon_{w_n,L_{n^*}}^{total}z_n(1-T_z^{'}(z_n))g_n(n)f_n(n)dn - \int_{n^*}g_n(n)f_n(n)dn \\
&\quad + \frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon_{w,L_{n^*}}^{total}[1-T_z^{\prime}(z)-T_z^{\prime\prime}(z)z]\left(\int_z g_z(z_n)f_z(z_n)dz_n\right)dz.
\end{aligned}
$$

In the last equation, we use $g_z(z_n) \equiv g_n(n) = \frac{\int g(n,q)f(n,q)dq}{f_n(n)}$, and exchange the order of integration. Bear in mind that we have defined the average welfare weight above a certain labor income z_n as $\bar{g}_{z_n} \equiv \bar{g}_n = \frac{\int_{n'>n}g_{n'}(n')f_{n'}(n')dn'}{1-F_n(n)}$, thus we have

$$
\begin{aligned}
&\int\int g(n,q)f(n,q)dV(n,q)dndq \\
&= -\bar{g}_{z_{n^*}}(1-F_z(z_{n^*})) \\
&\quad + \frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon_{w,L_{n^*}}^{total}\left[\left(1-T_z^{'}(z)-T_z^{''}(z)z\right)\bar{g}_z(1-F_z(z)) - z(1-T_z^{'}(z))g_z(z)f_z(z)\right]dz \\
&= -\bar{g}_{z_{n^*}}(1-F_z(z_{n^*})) + \frac{\epsilon_{l_{n^*},1-T_z^{'}}}{1-T_z^{'}(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon_{w_n,L_{n^*}}^{total}\left[(1-T_z^{'}(z_n))z_n\bar{g}_{z_n}(1-F_z(z_n))\right]^{'}dz_n.
\end{aligned}
$$

We now turn to the incidence of the given tax perturbation on government revenue,

$$
\begin{aligned}
dB =& \int\int \left[T_z'(z_n)z_n\left(\frac{dw_n}{w_n}+\frac{dl_n}{l_n}\right)+\tau_z(z_n)+T_x'(x_q)x_q\left(\frac{dR}{R}+\frac{da_q}{a_q}\right)+\tau_x(x_q)\right]f(n,q)dndq\\
=& \int\left[T_z'(z_n)z_n\left(\int \epsilon^{total}_{w_n,L_{n'}}\frac{dl_{n'}}{l_{n'}}dn'+\frac{dl_n}{l_n}\right)+\tau_z^1(z_n)+\tau_z^2(z_n)\right]f_n(n)dn\\
&+\int\left[T_x'(x_q)x_q\left(\int \epsilon^{total}_{R,L_{n'}}\frac{dl_{n'}}{l_{n'}}dn'+0\right)+(1-T_x'(x_q))x_q\int \epsilon^{total}_{R,L_{n'}}\frac{dl_{n'}}{l_{n'}}dn'\right]f_q(q)dq\\
=& -\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}T_z'(z_n)z_nf_n(n)dn-\int\frac{\epsilon_{l_n,1-T_z'}}{1-T_z'(z_n)}\delta_{z_{n^*}}(z_n)T_z'(z_n)z_nf_n(n)dn\\
&+\int \mathbf{I}_{z_n\geq z_{n^*}}f_n(n)dn-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\int_0^{z_n}\tilde{\epsilon}^{total}_{w',L_{n^*}}\left[1-T_z'(z')-T_z''(z')z'\right]dz'f_z(z_n)dz_n\\
&-\frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int x_q\epsilon^{total}_{R,L_{n^*}}f_q(q)dq.
\end{aligned}
$$

By adjusting the order of integration, we convert the double integral into a single integral, thus

$$
\begin{aligned}
dB =& -\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}T_z'(z_n)z_nf_z(z_n)dz_n-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}T_z'(z_{n^*})z_{n^*}f_z(z_{n^*})\\
&+(1-F_z(z_{n^*}))-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w',L_{n^*}}\left[1-T_z'(z')-T_z''(z')z'\right](1-F_z(z'))dz'\\
&-\frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int x_q\epsilon^{total}_{R,L_{n^*}}f_q(q)dq\\
=&1-F_z(z_{n^*})-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}T_z'(z_{n^*})z_{n^*}f_z(z_{n^*})-\frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int x_q\epsilon^{total}_{R,L_{n^*}}f_q(q)dq\\
&-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[T_z'(z_n)z_nf_z(z_n)+\left(1-T_z'(z_n)-T_z''(z_n)z_n\right)(1-F_z(z_n))\right]dz_n\\
=&1-F(z_{n^*})-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}T_z'(z_{n^*})z_{n^*}f_z(z_{n^*})\\
&-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[(1-T_z'(z_n))z_n(1-F(z_n))\right]'dz_n\\
&-\frac{\epsilon_{l_{n^*},1-T_{z^*}'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\left[\int x_q\epsilon^{total}_{R,L_{n^*}}f_q(q)dq+\int z_n\epsilon^{total}_{w_n,L_{n^*}}f_n(n)dn\right]\\
=&1-F(z_{n^*})-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}T_z'(z_{n^*})z_{n^*}f_z(z_{n^*})\\
&-\frac{\epsilon_{l_{n^*},1-T_z'}}{1-T_z'(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[\left(1-T_z'(z_n)\right)z_n(1-F_z(z_n))\right]'dz_n.
\end{aligned}
$$

In the last equation, we use the following Euler's homogeneous function theorem,

Theorem 1 *Given an aggregate CES production function described in Lemma 1, Euler's homogeneous function theorem implies that*[18]

$$\int x_q \epsilon^{total}_{R,L_{n^*}} f_q(q)dq + \int z_n \epsilon^{total}_{w_n,L_{n^*}} f_n(n)dn = 0. \tag{E.2}$$

Armed with these expressions, the marginal effects of the tax perturbation on social welfare can be given by

$$\begin{aligned}
dW =& \int\int g(n,q)f(n,q)dV(n,q)dndq + dB \\
=& -\bar{g}_{z_{n^*}}(1-F(z_{n^*})) + \frac{\epsilon_{l_{n^*},1-T'_z}}{1-T'_z(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[(1-T'_z(z_n))z_n\bar{g}_{z_n}(1-F_z(z_n))\right]' dz_n \\
&+ 1 - F_z(z_{n^*}) - \frac{\epsilon_{l_{n^*},1-T'_z}}{1-T'_z(z_{n^*})}T'_z(z_{n^*})z_{n^*}f_z(z_{n^*}) \\
&- \frac{\epsilon_{l_{n^*},1-T'_z}}{1-T'_z(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[\left(1-T'_z(z_n)\right)z_n(1-F_z(z_n))\right]' dz_n \\
=&(1-\bar{g}_{z_{n^*}})(1-F_z(z_{n^*})) - \frac{\epsilon_{l_{n^*},1-T'_z}}{1-T'_z(z_{n^*})}T'_z(z_{n^*})z_{n^*}f_z(z_{n^*}) \\
&- \frac{\epsilon_{l_{n^*},1-T'_z}}{1-T'_z(z_{n^*})}\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int \epsilon^{total}_{w_n,L_{n^*}}\left[(1-\bar{g}_{z_n})\left(1-T'_z(z_n)\right)z_n(1-F_z(z_n))\right]' dz_n.
\end{aligned}$$

Equate dW to zero and rearrange, the optimal nonlinear labor income tax formula can be elicited as follow,

$$\begin{aligned}
\frac{T'_z(z_{n^*})}{1-T'_z(z_{n^*})} =& \frac{1}{\epsilon_{l_{n^*},1-T'_z}}(1-\bar{g}_{z_{n^*}})\frac{1-F_z(z_{n^*})}{z_{n^*}f_z(z_{n^*})} \\
&- \left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\left[(1-\bar{g}_{z_n})\left(\frac{1-T'_z(z_n)}{1-T'_z(z_{n^*})}\right)\left(\frac{1-F_z(z_n)}{z_{n^*}f_z(z_{n^*})}\right)z_n\right]' \epsilon^{total}_{w_n,L_{n^*}} dz_n.
\end{aligned}$$

Denote $\epsilon^{SE}_{w_n,L_{n^*}} = \epsilon^{D}_{w_{n^*},L_{n^*}}\Delta_{n^*}(n) + \epsilon_{w_n,L_{n^*}}$ and $\epsilon^{AE}_{w_n,L_{n^*}} = \epsilon_{w_{n^*},\alpha_{n^*}}\epsilon^{D}_{\alpha_{n^*},L_{n^*}}\Delta_{n^*}(n) + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,L_{n^*}}$, we have $\epsilon^{total}_{w_n,L_{n^*}} = \epsilon^{SE}_{w_n,L_{n^*}} + \epsilon^{AE}_{w_n,L_{n^*}}$. Thus, the optimal tax formula can be further decomposed into (33). ■

[18] Using $F(K,\mathcal{L},\boldsymbol{\alpha}) = R\int a_q f_q(q)dq + \int w_n l_n f_n(n)dn$, take the derivative of L_{n^*} and rearrange, one can obtain theorem 1 directly.

E.3 Proof of Lemma 5

Proof. Our goal is to find the counteracting tax perturbation τ_x^2 and τ_z corresponding to the elementary τ_x^1, such that

$$\frac{da_q}{a_q} = -\epsilon_{a_q,1-T_x'} \frac{\tau_x^{1'}(x_q)}{1-T_x'(x_q)}, \quad \frac{dl_n}{l_n} = 0.$$

Bear the integral equation of capital supply in mind, that is

$$\frac{da_q}{a_q} = -\epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} + \epsilon_{a_q,R} \int \epsilon_{R,L_{n'}}^{total} \frac{dl_{n'}}{l_{n'}} dn' + \epsilon_{a_q,R} \epsilon_{R,K}^{total} \frac{dK}{K}.$$

As $\tau_x'(x_q) = \tau_x^{1'}(x_q) + \tau_x^{2'}(x_q)$, it is easily to obtain that

$$\tau_x^{2'}(x_q) = \frac{\epsilon_{a_q,R}}{\epsilon_{a_q,1-T_x'}} \left(1 - T_x'(x_q)\right) \epsilon_{R,K}^{total} \frac{dK}{K} = \left(1 - T_x'(x_q) - T_x''(x_q) x_q\right) \epsilon_{R,K}^{total} \frac{dK}{K}.$$

In the last equation, we use the expressions of $\epsilon_{a_q,R}$ and $\epsilon_{a_q,1-T_x'}$ displayed in Table 1. Integrate across x_q, the counteracting perturbation $\tau_x^2(x_q)$ can be given by

$$\tau_x^2(x_q) = (1 - T_x'(x_q)) x_q \epsilon_{R,K}^{total} \frac{dK}{K}. \tag{E.3}$$

As shown in Proposition 3, the integral equation system can be expressed as follows,

$$\frac{dl_n}{l_n} = -\epsilon_{l_n,1-T_z'} \frac{\tau_z'(z_n)}{1-T_z'(z_n)} - \epsilon_{l_n,w_n} \epsilon_{w_n,K}^{total} \chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} dq + \epsilon_{l_n,w_n} \int \zeta_{w_n,L_{n'}} \frac{dl_{n'}}{l_{n'}} dn',$$

$$\frac{dK}{K} = -\chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} dq + \chi \bar{\epsilon}_{K,R} \int \epsilon_{R,L_{n'}}^{total} \frac{dl_{n'}}{l_{n'}} dn'.$$

Using $dl_n/l_n = 0, \forall n$, we have $\frac{dK}{K} = -\chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} dq$ and

$$\tau_z'(z_n) = \frac{\epsilon_{l_n,w_n}}{\epsilon_{l_n,1-T_z'}} (1 - T_z'(z_n)) \epsilon_{w_n,K}^{total} \frac{dK}{K} = (1 - T_z'(z_n) - T_z''(z_n) z_n) \epsilon_{w_n,K}^{total} \frac{dK}{K}.$$

Again, the supply-side elasticities ϵ_{l_n,w_n} and $\epsilon_{l_n,1-T_z'}$ defined in Table 1 are used in the last equation. Integrate across z_n, the counteracting perturbation $\tau_z(z_n)$ is given as follow

$$\tau_z(z_n) = \int_0^{z_n} (1 - T_z'(z) - T_z''(z)z)\epsilon_{w,K}^{total} dz \frac{dK}{K}. \tag{E.4}$$

Substitute $\frac{dK}{K} = -\chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x'(x_q)}{1-T_x'(x_q)} dq$ and $\frac{dl_n}{l_n} = 0, \forall n$ into the expressions of factor price adjustments, that is Corollary 1,

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total} \chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1 - T_x'(x_q)} dq + \int \left[\epsilon_{w_n,L_{n'}}^{total} + \epsilon_{w_n,K}^{total} \chi \bar{\epsilon}_{K,R} \epsilon_{R,L_{n'}}^{total} \right] \frac{dl_{n'}}{l_{n'}} dn',$$

$$\frac{dR}{R} = -\epsilon_{R,K}^{total} \chi \int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{'}(x_q)}{1 - T_x'(x_q)} dq + \int \left[\epsilon_{R,L_{n'}}^{total} + \epsilon_{R,K}^{total} \chi \bar{\epsilon}_{K,R} \epsilon_{R,L_{n'}}^{total} \right] \frac{dl_{n'}}{l_{n'}} dn'.$$

the incidence of above tax reform on wage and rental rate can be reduced to

$$\frac{dw_n}{w_n} = \epsilon_{w_n,K}^{total} \frac{dK}{K}, \quad \frac{dR}{R} = \epsilon_{R,K}^{total} \frac{dK}{K}. \tag{E.5}$$

As the aggregate adjustment of capital supply can be expressed as follow

$$\frac{dK}{K} = \int \omega_q \frac{da_q}{a_q} dq = -\int \omega_q \epsilon_{a_q,1-T_x'} \frac{\tau_x^{1'}(x_q)}{1 - T_x'(x_q)} dq. \tag{E.6}$$

The remaining work is to specify the elementary tax perturbation $\tau_x^1(x_q)$. Consider a Dirac tax perturbation $\tau_x^1 = \mathbf{I}_{x \geq x^*}$ and $\tau_x^{1'}(x) = \delta_{x^*}(x)$, we have $\frac{dK}{K} = -\frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1}$, plug it back into equation (E.3), (E.4) and (E.5), we prove that

$$\tau_x^2(x_q) = -\epsilon_{R,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1 - T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1} (1 - T_x^{'}(x_q))x_q,$$

$$\tau_z(z_n) = -\frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1 - T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1} \int_0^{z_n} \epsilon_{w,K}^{total} [1 - T_z'(z) - T_z''(z)z] dz,$$

and

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1 - T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1}, \quad \frac{dR}{R} = -\epsilon_{R,K}^{total} \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1 - T_x'(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1}.$$

■

E.4 Proof of Proposition 6

The proof of optimal nonlinear capital income tax formula is indeed symmetric with the one of optimal nonlinear labor income tax. The incidence analysis of tax reform on social welfare imply that

$$dV(n,q) = z_n(1-T_z'(z_n))\frac{dw_n}{w_n} + x_q(1-T_x'(x_q))\frac{dR}{R} - \tau_x^1(x_q) - \tau_x^2(x_q) - \tau_z(z_n).$$

Bear Lemma 5 in mind, it is easily to obtain

$$\begin{aligned} dV(n,q) =& z_n(1-T_z^{'}(z_n))\epsilon_{w_n,K}^{total}\frac{dK}{K} + x_q(1-T_x^{'}(x_q))\epsilon_{R,K}^{total}\frac{dK}{K} - \mathbf{I}_{x_q \geq x_{q^*}} \\ & - (1-T_x^{'}(x_q))x_q\epsilon_{R,K}^{total}\frac{dK}{K} - \int_0^{z_n}(1-T_z^{'}(z) - T_z^{''}(z)z)\epsilon_{w,K}^{total}dz\frac{dK}{K} \\ =& - \mathbf{I}_{x_q \geq x_{q^*}} - z_n(1-T_z^{'}(z_n))\epsilon_{w_n,K}^{total}\frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1} \\ & + \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int_0^{z_n}(1-T_z^{'}(z) - T_z^{''}(z)z)\epsilon_{w,K}^{total}dz. \end{aligned}$$

Aggregate across individuals, one obtains

$$\begin{aligned} & \int\int g(n,q)f(n,q)dV(n,q)dndq \\ =& - \int\int_{q^*} g(n,q)f(n,q)dndq - \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}z_n(1-T_z^{'}(z_n))g_n(n)f_n(n)dn \\ & + \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int\int\int_0^{z_n}(1-T_z^{'}(z) - T_z^{''}(z)z)\epsilon_{w,K}^{total}dzg(n,q)f(n,q)dndq \\ =& - \bar{g}_{x_{q^*}}(1-F_x(x_{q^*})) - \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}z_n(1-T_z^{'}(z_n))g_z(z_n)f_z(z_n)dz_n \\ & + \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int(1-T_z^{'}(z) - T_z^{''}(z)z)\epsilon_{w,K}^{total}\int_z g_z(z_n)f_z(z_n)dz_ndz \\ =& - \bar{g}_{x_{q^*}}(1-F_x(x_{q^*})) - \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}z_n(1-T_z^{'}(z_n))g_z(z_n)f_z(z_n)dz_n \\ & + \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w,K}^{total}(1-T_z^{'}(z) - T_z^{''}(z)z)\bar{g}_z(1-F_z(z))dz \\ =& - \bar{g}_{x_{q^*}}(1-F_x(x_{q^*})) \\ & + \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}\left[(1-T_z^{'}(z_n))z_n\bar{g}_{z_n}(1-F_z(z_n))\right]^{'}dz_n. \end{aligned}$$

In the last equation, we use

$$\left[(1-T_z'(z_n))z_n\bar{g}_{z_n}(1-F_z(z_n))\right]' = \left[(1-T_z'(z_n)-T_z''(z_n)z_n)\bar{g}_{z_n}(1-F_z(z_n)) - z_n(1-T_z'(z_n)g_z(z_n)f_z(z_n)\right].$$

Next, we turn to the incidence of tax reform on government revenue, bear in mind that

$$\begin{aligned} dB =& \int\int d\left[T_z(z_n)+T_x(x_q)\right] f(n,q)dndq \\ =& \int\int \left[T_z'(z_n)z_n\left(\frac{dw_n}{w_n}+\frac{dl_n}{l_n}\right)+\tau_z(z_n)+T_x'(x_q)x_q\left(\frac{dR}{R}+\frac{da_q}{a_q}\right)+\tau_x^1(x_q)+\tau_x^2(x_q)\right] f(n,q)dndq. \end{aligned}$$

Lemma 5 implies that

$$\begin{aligned} dB =& \int T_z'(z_n)z_n\epsilon_{w_n,K}^{total}f_z(z_n)dz_n\frac{dK}{K} + \int\int_0^{z_n}(1-T_z'(z)-T_z''(z)z)\epsilon_{w,K}^{total}dzf_n(n)dn\frac{dK}{K} \\ &+ \int T_x'(x_q)x_qf_q(q)dq\epsilon_{R,K}^{total}\frac{dK}{K} + \int T_x'(x_q)x_q\frac{da_q}{a_q}f_q(q)dq \\ &+ (1-F_x(x_{q^*})) + \int(1-T_x'(x_q))x_qf_q(q)dq\epsilon_{R,K}^{total}\frac{dK}{K} \\ =&1-F_x(x_{q^*}) - \int T_x'(x_q)x_q\epsilon_{a_q,1-T_x'}\frac{\tau_x^{1'}(x_q)}{1-T_x'(x_q)}f_q(q)dq + \int \epsilon_{R,K}^{total}x_qf_q(q)dq\frac{dK}{K} \\ &+ \int T_z'(z_n)z_n\epsilon_{w_n,K}^{total}f_z(z_n)dz_n\frac{dK}{K} + \int(1-T_z'(z)-T_z''(z)z)\epsilon_{w,K}^{total}\int_z f_z(z_n)dz_ndz\frac{dK}{K} \\ =&1-F_x(x_{q^*}) - T_x'(x_{q^*})x_{q^*}f_x(x_{q^*})\frac{\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} + \int \epsilon_{R,K}^{total}x_qf_q(q)dq\frac{dK}{K} \\ &+ \int \epsilon_{w_n,K}^{total}\left[T_z'(z_n)z_nf_z(z_n)+(1-T_z'(z_n)-T_z'(z_n)z_n)(1-F_z(z_n))\right]dz_n\frac{dK}{K} \\ =&1-F_x(x_{q^*}) - T_x'(x_{q^*})x_{q^*}f_x(x_{q^*})\frac{\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \\ &- \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}\left[(1-T_z'(z_n))z_n(1-F_z(z_n))\right]'dz_n \\ &- \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\left(\int \epsilon_{R,K}^{total}x_qf_q(q)dq + \int \epsilon_{w_n,K}^{total}z_nf_n(n)dn\right) \\ =&1-F_x(x_{q^*}) - T_x'(x_{q^*})x_{q^*}f_x(x_{q^*})\frac{\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})} \\ &- \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x'}}{1-T_x'(x_{q^*})}\left(\frac{dx_{q^*}}{dq^*}\right)^{-1}\int \epsilon_{w_n,K}^{total}\left[(1-T_z'(z_n))z_n(1-F_z(z_n))\right]'dz_n. \end{aligned}$$

In the last equation, we use the Euler's homogeneous function theorem about K, which is symmetric with Theorem 1, that is

$$\int \epsilon_{R,K}^{total} x_q f_q(q) dq + \int \epsilon_{w_n,K}^{total} z_n f_n(n) dn = 0.$$

Equalizing $dW = \int \int g(n,q) f(n,q) dV(n,q) dn dq + dB$ to zero, one obtains

$$\begin{aligned} 0 =& (1-\bar{g}_{x_{q^*}})(1-F_x(x_{q^*})) - T_x^{'}(x_{q^*}) x_{q^*} f_x(x_{q^*}) \frac{\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})} \\ & - \frac{\omega_{q^*}\epsilon_{a_{q^*},1-T_x^{'}}}{1-T_x^{'}(x_{q^*})} \left(\frac{dx_{q^*}}{dq^*}\right)^{-1} \int \epsilon_{w_n,K}^{total} \left[(1-\bar{g}_{z_n})(1-T_z^{'}(z_n)) z_n (1-F_z(z_n))\right]^{'} dz_n. \end{aligned}$$

Rearrange above equation, the optimal nonlinear capital income tax formula can be derived as follow

$$\begin{aligned} \frac{T_x^{'}(x_{q^*})}{1-T_x^{'}(x_{q^*})} =& \frac{1}{\epsilon_{a_{q^*},1-T_x^{'}}}(1-\bar{g}_{x_{q^*}}) \frac{1-F_x(x_{q^*})}{x_{q^*} f_x(x_{q^*})} \\ & - \frac{1}{RK(1-T_x^{'}(x_{q^*}))} \int \left[(1-\bar{g}_{z_n}) \left(1-T_z^{'}(z_n)\right) (1-F_z(z_n))\, z_n\right]^{'} \epsilon_{w_n,K}^{total} dz_n. \end{aligned}$$

Note that

$$\frac{\omega_{q^*}(dx_{q^*}/dq^*)^{-1}}{x_{q^*} f_x(x_{q^*})} = \frac{(a_{q^*} f_q(q^*)/K)(dx_{q^*}/dq^*)^{-1}}{R a_{q^*} f_x(x_{q^*})} = \frac{a_{q^*} f_x(x_{q^*})/K}{R a_{q^*} f_x(x_{q^*})} = \frac{1}{RK}.$$

Denote $\epsilon_{w_n,K}^{total} = \epsilon_{w_n,K} + \epsilon_{w_n,\alpha_n}\epsilon_{\alpha_n,K} = \epsilon_{w_n,K}^{SE} + \epsilon_{w_n,K}^{AE}$, the optimal tax formula can be further decomposed into (34).

E.5 Proof of Lemma 6

In the case of Uniform tax perturbation $\tau_x^1(x) = x$ and $\tau_x^{1'}(x) = 1$, equation (E.6) converts to

$$\frac{dK}{K} = -\int \omega_q \frac{\epsilon_{a_q,1-T_x^{\prime}}}{1-T_x^{\prime}(x_q)} dq,$$

Plug back into equation (E.3) - (E.5), the counteracting tax perturbations can be given by

$$\tau_x^2(x_q) = -\epsilon_{R,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T_x^{\prime}}}{1-T_x^{\prime}(x_q)} dq (1-T_x^{\prime}(x_q)) x_q,$$

$$\tau_z(z_n) = -\int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int_0^{z_n} \epsilon_{w,K}^{total}(1-T'_z(z)-T''_z(z)z)dz,$$

and the corresponding adjustment of factor prices can be given by

$$\frac{dw_n}{w_n} = -\epsilon_{w_n,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq, \quad \frac{dR}{R} = -\epsilon_{R,K}^{total} \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq.$$

E.6 Proof of Proposition 7

Proof. Armed with Lemma 6, we now have $\frac{dK}{K} = -\int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq$. Follow the same processes in Appendix E.4, the incidence of tax reform on individual utilities can be given by

$$\begin{aligned}
dV(n,q) =& z_n(1-T'_z(z_n))\frac{dw_n}{w_n} + x_q(1-T'_x(x_q))\frac{dR}{R} - \tau_x^1(x_q) - \tau_x^2(x_q) - \tau_z(z_n) \\
=& z_n(1-T'_z(z_n))\epsilon_{w_n,K}^{total}\frac{dK}{K} + x_q(1-T'_x(x_q))\epsilon_{R,K}^{total}\frac{dK}{K} - x_q - (1-T'_x(x_q))x_q\epsilon_{R,K}^{total}\frac{dK}{K} \\
& - \int_0^{z_n} \epsilon_{w,K}^{total}(1-T'_z(z)-T''_z(z)z)dz\frac{dK}{K} \\
=& - x_q - z_n(1-T'_z(z_n))\epsilon_{w_n,K}^{total}\int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \\
& + \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int_0^{z_n} \epsilon_{w,K}^{total}(1-T'_z(z)-T''_z(z)z)dz.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\int\int g(n,q)f(n,q)dV(n,q)dndq \\
&= -\int x_q g_q(q)f_q(q)dq - \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int z_n(1-T'_z(z_n))\epsilon_{w_n,K}^{total} g_n(n)f_n(n)dn \\
&\quad + \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int\int\int_0^{z_n} \epsilon_{w,K}^{total}(1-T'_z(z)-T''_z(z)z)dz g(n,q)f(n,q)dndq \\
&= -\int x_q g_q(q)f_q(q)dq - \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int z_n(1-T'_z(z_n))\epsilon_{w_n,K}^{total} g_n(n)f_n(n)dn \\
&\quad + \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int \epsilon_{w,K}^{total}(1-T'_z(z)-T''_z(z)z)\bar{g}_z(1-F_z(z))dz \\
&= -\int x_q g_q(q)f_q(q)dq + \int \omega_q \frac{\epsilon_{a_q,1-T'_x}}{1-T'_x(x_q)} dq \int \epsilon_{w_n,K}^{total}\left[(1-T'_z(z_n))z_n\bar{g}_{z_n}(1-F_z(z_n))\right]' dz_n.
\end{aligned}$$

As for the government revenue, we have

$$
\begin{aligned}
dB &= \int\int d\left[T_z(z_n) + T_x(x_q)\right] f(n,q)dndq \\
&= \int\int \left[T_z^{'}(z_n)z_n\left(\frac{dw_n}{w_n} + \frac{dl_n}{l_n}\right) + \tau_z(z_n) + \tau_x^1(x_q) + \tau_x^2(x_q) + T_x^{'}(x_q)x_q\left(\frac{dR}{R} + \frac{da_q}{a_q}\right)\right] f(n,q)dndq \\
&= \int T_z^{'}(z_n)z_n\epsilon_{w_n,K}^{total}f_n(n)dn\frac{dK}{K} + \int\int\int_0^{z_n}(1 - T_z^{'}(z) - T_z^{''}(z)z)\epsilon_{w,K}^{total}dzf(n,q)dndq\frac{dK}{K} + \int x_qf_q(q)dq \\
&\quad + \int(1 - T_x^{'}(x_q))x_qf_q(q)dq\epsilon_{R,K}^{total}\frac{dK}{K} + \int T_x^{'}(x_q)x_qf_q(q)dq\epsilon_{R,K}^{total}\frac{dK}{K} + \int T_x^{'}(x_q)x_q\frac{da_q}{a_q}f_q(q)dq \\
&= \int x_qf_q(q)dq - \int \epsilon_{a_q,1-T_x^{'}}\frac{T_x^{'}(x_q)}{1 - T_x^{'}(x_q)}x_qf_q(q)dq + \int x_qf_q(q)dq\epsilon_{R,K}^{total}\frac{dK}{K} \\
&\quad + \int T_z^{'}(z_n)z_n\epsilon_{w_n,K}^{total}f_n(n)dn\frac{dK}{K} + \int \epsilon_{w,K}^{total}(1 - T_z^{'}(z) - T_z^{''}(z)z)(1 - F_z(z))dz\frac{dK}{K}.
\end{aligned}
$$

Restrict the tax function to be linear, i.e., $T_x(x_q) = t_x$ for all $q \in Q$, thus we have

$$
\begin{aligned}
dB &= \int x_qf_q(q)dq - \frac{t_x}{1 - t_x}\int \epsilon_{a_q,1-t_x}x_qf_q(q)dq - \int \omega_q\frac{\epsilon_{a_q,1-t_x}}{1 - t_x}dq\int \epsilon_{w_n,K}^{total}\left[(1 - T_z^{'}(z_n))z_n(1 - F_z(z_n))\right]^{'}dz_n \\
&\quad - \int \omega_q\frac{\epsilon_{a_q,1-t_x}}{1 - t_x}dq\left(\int \epsilon_{R,K}^{total}x_qf_q(q)dq + \int \epsilon_{w_n,K}^{total}z_nf_n(n)dn\right) \\
&= \int x_qf_q(q)dq - \frac{t_x}{1 - t_x}\int \epsilon_{a_q,1-t_x}x_qf_q(q)dq - \int \omega_q\frac{\epsilon_{a_q,1-t_x}}{1 - t_x}dq\int \epsilon_{w_n,K}^{total}\left[(1 - T_z^{'}(z_n))z_n(1 - F_z(z_n))\right]^{'}dz_n.
\end{aligned}
$$

In the last equation, we use the Euler's homogeneous function theorem again. Equalizing $dW = \int\int g(n,q)f(n,q)dV(n,q)dndq + dB$, one obtains

$$
\begin{aligned}
0 &= \int\int g(n,q)f(n,q)dV(n,q)dndq + dB \\
&= \int(1 - g_q(q))x_qf_q(q)dq - \frac{t_x}{1 - t_x}\int \epsilon_{a_q,1-t_x}x_qf_q(q)dq \\
&\quad - \int \omega_q\frac{\epsilon_{a_q,1-t_x}}{1 - t_x}dq\int \epsilon_{w_n,K}^{total}\left[(1 - \bar{g}_{z_n})(1 - T_z^{'}(z_n))z_n(1 - F_z(z_n))\right]^{'}dz_n.
\end{aligned}
$$

Rearranging above equation leads to the optimal linear capital income tax formula,

$$
\begin{aligned}
\frac{t_x}{1 - t_x} &= \frac{\int(1 - g_q(q))x_qf_q(q)dq}{\int \epsilon_{a_q,1-t_x}x_qf_q(q)dq} \\
&\quad - \frac{1}{RK(1 - t_x)}\int \left[(1 - \bar{g}_{z_n})(1 - T_z^{'}(z_n))z_n(1 - F_z(z_n))\right]^{'}\epsilon_{w_n,K}^{total}dz_n.
\end{aligned}
$$

Note that

$$\frac{\int \omega_q \epsilon_{a_q,1-t_x} dq}{\int \epsilon_{a_q,1-t_x} x_q f_q(q) dq} = \frac{\int \frac{a_q f_q(q)}{K} \epsilon_{a_q,1-t_x} dq}{R \int \epsilon_{a_q,1-t_x} a_q f_q(q) dq} = \frac{1}{RK}.$$

Using $\epsilon^{total}_{w_n,K} = \epsilon^{SE}_{w_n,K} + \epsilon^{AE}_{w_n,K}$, the optimal tax formula can be further decomposed into (35). ■

F Equivalence to Primal Approach

In the main text, we derive the optimal tax formulas with dual approach, or variational approach. We now turn to primal approach, or mechanism design method, to solve the optimal tax formulas in general general equilibrium, then we prove these two approaches are indeed equivalent.

F.1 Tax formulas of Primal Approach

Proof. Bear in mind that there is multi-dimensional heterogeneity in our model, wealth endowment indexed by $q \in Q$, and individual skill indexed by $n \in N$, both of them are unobserved for the government. While individual derives utility in the form of $V(n,q) = u(y_q - a_q) + c(n,q) - v(l_n)$, then incentive compatibility requires that

$$u(y_q - a_q) + c(n,q) - v(l_n) \geq u(y_q - a_{q'}) + c(n',q') - v(\frac{w_{n'} l_{n'}}{w_n}), \quad \forall n', q'.$$

the incentive compatibility constraint (IC) can be given by

$$\nabla c(n,q) = [\nabla^n c(n,q), \nabla^q c(n,q)] = \left[v'(l_n)(\dot{w}_n l_n + w_n \dot{l}_n) \frac{1}{w_n}, u'(y_q - a_q)\dot{a}_q \right], \quad \forall n, q. \tag{F.1}$$

and the resource constraint (RC) is given by

$$\int \int c(n,q) f(n,q) dn dq = F(K, \mathcal{L}, \boldsymbol{\alpha}) + K - B. \tag{F.2}$$

In the same spirit of Loebbing (2020), we go through the following steps to derive the optimal tax formulas.

Step 1: Combine IC (F.1) and RC (F.2) to get an integrated constraint condition, that $c(n,q)$

is represented as a function of l_n and a_q,

$$
\begin{aligned}
c(n,q) &= c(\underline{n},\underline{q}) + \int_{\underline{n}}^{n} \nabla^n c(n',q)dn' + \int_{\underline{q}}^{q} \nabla^q c(n,q')dq \\
&= c(\underline{n},\underline{q}) + \int_{\underline{n}}^{n} v'(l_{n'})(\dot{w}_{n'}l_{n'} + w_{n'}\dot{l}_{n'})\frac{1}{w_{n'}}dn' - \int_{\underline{q}}^{q} u'(y_{q'}-a_{q'})(\dot{y}_{q'}-\dot{a}_{q'})dq' + \int_{\underline{q}}^{q} u'(y_{q'}-a_{q'})\dot{y}_{q'}dq' \\
&= c(\underline{n},\underline{q}) + \int_{\underline{n}}^{n} v'(l_{n'})\hat{w}_{n'}l_{n'}dn' + v(l_n) - v(l_{\underline{n}}) - u(y_q - a_q) + u(y_{\underline{q}} - a_{\underline{q}}) + \int_{\underline{q}}^{q} u'(y_{q'}-a_{q'})\dot{y}_{q'}dq'.
\end{aligned}
$$

Integrating both sides and using RC, we obtain

$$
\begin{aligned}
F(K,\mathcal{L};\boldsymbol{\alpha}) + K - B =& c(\underline{n},\underline{q}) + \int\int_{\underline{n}}^{n} v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n}f_n(n)dn + \int v(l_n)f_n(n)dn - v(l_{\underline{n}}) \\
&- \int u(y_q - a_q)f_q(q)dq + u(y_{\underline{q}} - a_{\underline{q}}) + \int\int_{\underline{q}}^{q} u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q}f_q(q)dq \\
=& c(\underline{n},\underline{q}) + \int\int_{\tilde{n}}^{\bar{n}} f_n(n)dnv'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} + \int v(l_n)f_n(n)dn - v(l_{\underline{n}}) \\
&- \int u(y_q - a_q)f_q(q)dq + u(y_{\underline{q}} - a_{\underline{q}}) + \int\int_{\tilde{q}}^{\bar{q}} f_q(q)dqu'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q} \\
=& c(\underline{n},\underline{q}) + \int(1 - F_n(\tilde{n}))v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} + \int v(l_n)f_n(n)dn - v(l_{\underline{n}}) \\
&- \int u(y_q - a_q)f_q(q)dq + u(y_{\underline{q}} - a_{\underline{q}}) + \int(1 - F_q(\tilde{q}))u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q}.
\end{aligned}
$$

Eliminating $c(\underline{n},\underline{q})$, the integrated constraint condition can be given as follow,

$$
\begin{aligned}
c(n,q) =& F(K,\mathcal{L};\boldsymbol{\alpha}) + K - B - \int(1 - F_n(\tilde{n}))\, v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} - \int v(l_{\tilde{n}})f_n(\tilde{n})d\tilde{n} + \int_{\underline{n}}^{n} v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} + v(l_n) \\
&+ \int u(y_{\tilde{q}} - a_{\tilde{q}})f_q(\tilde{q})d\tilde{q} - u(y_q - a_q) + \int_{\underline{q}}^{q} u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q} - \int(1 - F_q(\tilde{q}))u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q}.
\end{aligned}
\tag{F.3}
$$

Step 2: Substituting $c(n,q)$ back into $V(n,q)$, then differentiating social welfare W with respect to l_n and a_q to get the first-order conditions for social planner,

$$
\begin{aligned}
V(n,q) =& u(y_q - a_q) + c(n,q) - v(l_n) \\
=& F(K,\mathcal{L};\boldsymbol{\alpha}) + K - B - \int(1 - F_n(\tilde{n}))\, v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} - \int v(l_{\tilde{n}})f_n(\tilde{n})d\tilde{n} + \int_{\underline{n}}^{n} v'(l_{\tilde{n}})\hat{w}_{\tilde{n}}l_{\tilde{n}}d\tilde{n} \\
&+ \int u(y_{\tilde{q}} - a_{\tilde{q}})f_q(\tilde{q})d\tilde{q} + \int_{\underline{q}}^{q} u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q} - \int(1 - F_q(\tilde{q}))u'(y_{\tilde{q}} - a_{\tilde{q}})\dot{y}_{\tilde{q}}d\tilde{q}.
\end{aligned}
$$

Given the social welfare function as $W = \frac{1}{\lambda}G\left(\{V(n,q)\}_{n\times q\in N\times Q}\right) + B$, the spirit of primal approach is to choose optimal allocations $\{l_n\}_N$ and $\{a_q\}_Q$ such that social welfare is maximized.

Differentiate W with respect to l_n and a_q, and denote $\tilde{V} = V + B$, we have

$$\frac{\partial W}{\partial l_n} = \int\int g(n',q')f(n',q')\frac{\partial V(n',q')}{\partial l_n}dn'dq' + \frac{\partial B}{\partial l_n} = \int\int g(n',q')f(n',q')\frac{\partial \tilde{V}(n',q')}{\partial l_n}dn'dq' \tag{F.4}$$

$$\frac{\partial W}{\partial a_q} = \int\int g(n',q')f(n',q')\frac{\partial V(n',q')}{\partial a_q}dn'dq' + \frac{\partial B}{\partial a_q} = \int\int g(n',q')f(n',q')\frac{\partial \tilde{V}(n',q')}{\partial a_q}dn'dq'. \tag{F.5}$$

The remaining work is to find $\frac{\partial \tilde{V}(n',q')}{\partial l_n}$ and $\frac{\partial \tilde{V}(n',q')}{\partial a_q}$. Since

$$\begin{aligned}\frac{\partial \tilde{V}(n',q')}{\partial l_n} =& w_n f_n(n) - v'(l_n)f_n(n)\\ &- (1-F_n(n))(v''(l_n)l_n + v'(l_n))\hat{w}_n - \int (1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}\\ &+ (v''(l_n)l_n + v'(l_n))\hat{w}_n \cdot \mathbf{1}_{n'\geq n} + \int_{\underline{n}}^{n'} v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n},\end{aligned}$$

$$\begin{aligned}\frac{\partial \tilde{V}(n',q')}{\partial a_q} =& (1+R)f_q(q) - \int (1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial a_q}d\tilde{n} + \int_{\underline{n}}^{n'} v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial a_q}d\tilde{n} - u'(y_q - a_q)f_q(q)\\ &- u''(y_q-a_q)\dot{y}_q \cdot \mathbf{1}_{q'\geq q} + (1-F_q(q))u''(y_q-a_q)\dot{y}_q,\end{aligned}$$

plug above equations into (F.4) and (F.5), the derivations of social welfare W with respect to factor supplies can be given by

$$\begin{aligned}\frac{\partial W}{\partial l_n} =& w_n f_n(n) - v'(l_n)f_n(n)\\ &- (1-F_n(n))(v''(l_n)l_n + v'(l_n))\hat{w}_n - \int (1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}\\ &+ \int\int_n^{\bar{n}} g(n',q')f(n',q')(v''(l_n)l_n + v'(l_n))\hat{w}_n dn'dq' + \int\int g(n',q')f(n',q')\int_{\underline{n}}^{n'} v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}dn'dq'\\ =& w_n f_n(n) - v'(l_n)f_n(n)\\ &- (1-F_n(n))(v''(l_n)l_n + v'(l_n))\hat{w}_n - \int (1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}\\ &+ \bar{g}_n(n)(1-F_n(n))(v''(l_n)l_n + v'(l_n))\hat{w}_n + \int\int\int_{\tilde{n}}^{\bar{n}} g(n',q')f(n',q')dn'dq'v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}\\ =& w_n f_n(n) - v'(l_n)f_n(n) - (1-\bar{g}_n(n))(1-F_n(n))(v''(l_n)l_n + v'(l_n))\hat{w}_n\\ &- \int (1-\bar{g}_n(\tilde{n}))\,(1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}.\end{aligned}$$

Note that $\int\int g(n,q)f(n,q)dndq = 1$ and $\bar{g}_n = \frac{\int\int_n g(n',q')f(n',q')dn'dq'}{1-F_n(n)}$. Similarly, one obtains

$$\begin{aligned}
\frac{\partial W}{\partial a_q} =&(1+R)f_q(q) - u'(y_q - a_q)f_q(q) \\
&+ (1-F_q(q))u''(y_q-a_q)\dot{y}_q - \int (1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial a_q}d\tilde{n} \\
&- \int\int_q^{\bar{q}} g(n',q')f(n',q')u''(y_q-a_q)\dot{y}_q dn'dq' + \int\int g(n',q')f(n',q')\int_{\underline{n}}^{n'} v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n}dn'dq' \\
=&(1+R)f_q(q) - u'(y_q-a_q)f_q(q) + (1-\bar{g}_q(q))(1-F_q(q))u''(y_q-a_q)\dot{y}_q \\
&- \int (1-\bar{g}_n(\tilde{n}))\,(1-F_n(\tilde{n}))\, v'(l_{\tilde{n}})l_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial a_q}d\tilde{n}.
\end{aligned}$$

Step 3: Using individual's first-order conditions to reintroduce the optimal marginal income tax rates.

Individual first-order conditions are given by

$$v'(l_n) = (1-T_z'(z_n))w_n, \quad u'(y_q-a_q) = 1 + (1-T_x'(x_q))R.$$

Bear in mind that $e_{l_n,1-t_z} = \frac{v'(l_n)}{v''(l_n)l_n}$ and $e_{a_q,1-t_x} = -\frac{u'(y_q-a_q)-1}{a_q u''(y_q-a_q)}$. Thus we have

$$v''(l_n)l_n = \frac{(1-T_z'(z_n))w_n}{e_{l_n,1-t_z}}, \quad u''(y_q-a_q)a_q = -\frac{(1-T_x'(x_q))R}{e_{a_q,1-t_x}}.$$

Equalize $\frac{\partial W}{\partial l_n} = \frac{\partial W}{\partial a_q} = 0$, and plug into above equations, one obtains

$$\begin{aligned}
0 =&T_z'(z_n)w_n f_n(n) - \left(1 + \frac{1}{\epsilon_{l_n,1-t_z}}\right)(1-\bar{g}_n(n))(1-F_n(n))(1-T_z'(z_n))\dot{w}_n \\
&- \int (1-\bar{g}_n(\tilde{n}))\,(1-F_n(\tilde{n}))\,(1-T_z'(z_{\tilde{n}}))z_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial l_n}d\tilde{n},
\end{aligned}$$

$$\begin{aligned}
0 =&T_x'(x_q)Rf_q(q) - \frac{1}{\epsilon_{a_q,1-t_x}}(1-\bar{g}_q(q))(1-F_q(q))(1-T_x'(x_q))R\frac{\dot{y}_q}{a_q} \\
&- f_q(q)\int (1-\bar{g}_n(\tilde{n}))\,(1-F_n(\tilde{n}))\,(1-T_z'(z_{\tilde{n}}))z_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial K}d\tilde{n}.
\end{aligned}$$

Rearrange, then lead to the following optimal tax formulas

$$
\begin{aligned}
\frac{T_z'(z_n)}{1-T_z'(z_n)} &= \left(1+\frac{1}{e_{l_n,1-t_z}}\right)(1-\bar{g}_n(n))\frac{1-F_n(n)}{f_n(n)}\hat{w}_n \\
&\quad + \int (1-\bar{g}_n(\tilde{n}))\left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right)\left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial L_n/L_n}d\tilde{n},
\end{aligned} \tag{F.6}
$$

$$
\begin{aligned}
\frac{T_x'(x_q)}{1-T_x'(x_q)} &= \frac{1}{e_{a_q,1-t_x}}(1-\bar{g}_q(q))\frac{1-F_q(q)}{f_q(q)a_q}\dot{y}_q \\
&\quad + \int (1-\bar{g}_n(\tilde{n}))\left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_x'(x_q)}\right)\left(\frac{1-F_n(\tilde{n})}{RK}\right) z_{\tilde{n}}\frac{\partial \hat{w}_{\tilde{n}}}{\partial K/K}d\tilde{n}.
\end{aligned} \tag{F.7}
$$

■

F.2 Equivalence

We now show that equation (F.6) and (F.7) are indeed equivalent to equation (33) and (34), that is, primal approach is equivalence to dual approach in general equilibrium with multi-heterogeneity.

Equivalence of labor income taxation. Start from the first term of (F.6), with $F_z(z_n) = F_w(w_n) = F_n(n)$, we know that $f_w(w_n)\dot{w}_n = f_n(n)$ and $f_z(z_n)\frac{dz_n}{dw_n} = f_w(w_n)$. Moreover, $z_n = w_n l_n$ implies that $dz_n/dw_n = l_n(1+\epsilon_{l_n,w_n})$. Since $\hat{w}_n = \dot{w}_n/w_n$, it easily to show that

$$
\frac{1-F_n(n)}{f_n(n)}\hat{w}_n = \frac{1-F_z(z_n)}{w_n f_w(w_n)} = \frac{1-F_z(z_n)}{(1+\epsilon_{l_n,w_n})z_n f_z(z_n)}.
$$

Given the expressions of ϵ_{l_n,w_n} and $\epsilon_{l_n,1-T_z'}$ in Table 1, that is

$$
\epsilon_{l_n,1-T_z'} = \frac{[1-T_z'(z_n)]\, e_{l_n,1-t_z}}{1-T_z'(z_n)+e_{l_n,1-t_z}T_z''(z_n)z_n}, \quad \epsilon_{l_n,w_n} = \frac{[1-T_z'(z_n)-T_z''(z_n)z_n]\, e_{l_n,1-t_z}}{1-T_z'(z_n)+e_{l_n,1-t_z}T_z''(z_n)z_n}.
$$

One can verify that $\frac{1}{\epsilon_{l_n,1-T_z'}} = \frac{1}{1+\epsilon_{l_n,w_n}}(1+\frac{1}{e_{l_n,1-t_z}})$. Therefore the first term of (F.6) can be rewritten as

$$
\left(1+\frac{1}{e_{l_n,1-t_z}}\right)(1-\bar{g}_n(n))\frac{1-F_n(n)}{f_n(n)}\hat{w}_n = \frac{1}{\epsilon_{l_n,1-T_z'}}(1-\bar{g}_{z_n})\frac{1-F_z(z_n)}{z_n f_z(z_n)},
$$

which is equivalent to the first term of equation (33).

We now turn to the general equilibrium term in the optimal nonlinear labor income tax formula. With $\frac{d\hat{w}_{\tilde{n}}}{dL_n/L_n} = \frac{d(d\ln w_{\tilde{n}}/d\ln L_n)}{d\tilde{n}} = \frac{d\epsilon^{total}_{w_{\tilde{n}},L_n}}{d\tilde{n}}$, the second term in equation (F.6) derived by primal approach can be rewritten as

$$\int (1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}} d\tilde{\epsilon}^{total}_{w_{\tilde{n}},L_n}.$$

Next, using integration by parts, we have

$$\begin{aligned}&\int (1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}} d\tilde{\epsilon}^{total}_{w_{\tilde{n}},L_n} \\ &= (1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}} \epsilon^{total}_{w_{\tilde{n}},L_n}|^{\bar{n}}_{\underline{n}} - \int \epsilon^{total}_{w_{\tilde{n}},L_n} d\left[(1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}}\right].\end{aligned}$$

Bear in mind that $\bar{g}_n = \frac{\int \int_n^{\bar{n}} g(n',q) f(n',q) dn' dq}{1-F_n(n)}$, and $\int \int_{\underline{n}}^{\bar{n}} g(n',q) f(n',q) dn' dq$ has been normalized to one, thus we have $1-\bar{g}_{\underline{n}} = 0$. In addition, with $1-F(\bar{n}) = 0$, the first part of above equation equals to zero, which leads to

$$\begin{aligned}&\int (1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}} d\epsilon^{total}_{w_{\tilde{n}},L_n} \\ &= -\int \epsilon^{total}_{w_{\tilde{n}},L_n} d\left[(1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}}\right] \\ &= -\int \left[(1-\bar{g}_{\tilde{n}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_n(\tilde{n})}{z_n f_n(n)}\right) z_{\tilde{n}}\right]' \epsilon^{total}_{w_{\tilde{n}},L_n} dz_{\tilde{n}} \\ &= -\left(\frac{dz_n}{dn}\right)^{-1} \int \left[(1-\bar{g}_{z_{\tilde{n}}}) \left(\frac{1-T_z'(z_{\tilde{n}})}{1-T_z'(z_n)}\right) \left(\frac{1-F_z(z_{\tilde{n}})}{z_n f_z(z_n)}\right) z_{\tilde{n}}\right]' \epsilon^{total}_{w_{\tilde{n}},L_n} dz_{\tilde{n}},\end{aligned}$$

which is consistent with the second term derived by dual approach. Using the definition of $\epsilon^{total}_{w_{\tilde{n}},L_n} = \epsilon^{SE}_{w_{\tilde{n}},L_n} + \epsilon^{AE}_{w_{\tilde{n}},L_n}$, the proving of equivalence between (F.6) and (33) is completed.

Equivalence of capital income taxation. Start from the first term of (F.7), our goal is to replace $\dot{y}_q$ by $\dot{a}_q$. Bear in mind that $u'(y_q - a_q) = 1 + (1 - T_x'(x_q))R$, this first-order condition implies

$$u''(y_q - a_q)(\dot{y}_q - \dot{a}_q) = -R^2 T_x''(x_q)\dot{a}_q.$$

thus we have

$$\dot{y}_q = \left(1 - \frac{R^2 T_x''(x_q)}{u''(y_q - a_q)}\right)\dot{a}_q = \frac{1 - T_x'(x_q) + e_{a_q,1-t_x} T_x''(x_q) x_q}{1 - T_x'(x_q)}\dot{a}_q.$$

in the last equation, we use $u''(y_q - a_q)a_q = -\frac{(1-T_x'(x_q))R}{e_{a_q,1-t_x}}$. Now the first term of (F.7) can be rewritten as

$$\frac{1}{e_{a_q,1-t_x}}(1-\bar{g}_q(q))\frac{1-F_q(q)}{f_q(q)a_q}\dot{y}_q = \frac{1}{e_{a_q,1-t_x}}\frac{1-T_x'(x_q)+e_{a_q,1-t_x}T_x''(x_q)x_q}{1-T_x'(x_q)}(1-\bar{g}_{x_q})\frac{1-F_x(x_q)}{x_q f_x(x_q)}$$
$$= \frac{1}{\epsilon_{a_q,1-T_x'}}(1-\bar{g}_{x_q})\frac{1-F_x(x_q)}{x_q f_x(x_q)}.$$

In the second equation, we use $F_x(x_q) = F_a(a_q) = F_q(q)$, and in the last equation, we use the expression of $\epsilon_{a_q,1-T_x'}$ given in Table 1. Thus, the first terms of (F.7) and (34) are equivalence. The equivalence of the general equilibrium term between (F.7) and (34) can be proved following the same techniques as the one of nonlinear labor income tax formula.

G Details for Quantitative Analysis

G.1 Calibration

Data description-We use the Distributional National Accounts micro-files of Piketty et al. (2018), this dataset combines tax, survey, and national accounts data to estimate the distribution of national income in the United States since 1913. The advantage of this dataset is that it captures 100% of national income relative to IRS tax return data, which is always used to calibrate the earning distribution(Saez and Stantcheva, 2018). Instead of specifying a certain distributional function of incomes, we calibrate the labor and capital income distribution using the observed income distribution directly. The DINAs micro-files contain the pretax labor income (*plinc*) and pretax capital income (*pkinc*) at the individual level. To bring the data to our model, we first restrict the value of *plinc* and *pkinc* between 5,000 and 10,000,000. Where *plinc*$= 5,000$ correspond to $\underline{n}$ in our model. The number of observations is 31387, and the median value of labor income is about $33,000$\$. Next, we discretize the income using bins, and the length of bin is $2,000$. By counting the number of observations in each bin as a proportion of the total number of observations, we approximate the probability density of the mean income of each bin, that is $f_n(n)$ and $f_q(q)$. After the processes of normalizing and smoothing, we can calculate the inverse Hazard Ratio at each bins, that is $(1-F_z(z_n))/f_z(z_n)z_n$. Note that $F_z(z_n) = F_n(n)$, so $f_z(z_n)dz_n = f_n(n)dn$. The same is for capital income distribution. After calibrating the distribution of labor income and capital income, we turn to the calibration of the 2019 U.S. economy.

On the preference side, we have $u(y_q - a_q) = -B_q \frac{a_q^{1+1/\epsilon_k}}{1+1/\epsilon_k} - a_q$ and $v(l_n) = \frac{l_n^{1+1/\epsilon_l}}{1+1/\epsilon_l}$, given the initial tax system $T_z(z_n) = z_n - \frac{1-\zeta}{1-\phi} z_n^{1-\phi}$ and $t_x = 10\%$, we can use the following first-order conditions

$$v'(l_n) = (1 - T_z'(z_n))w_n, \quad u'(y_q - a_q) = 1 + (1 - t_x)R,$$

to calibrate parameter B_q and the distribution of labor supply l_n and wage w_n. Note that we have $x_q = Ra_q$ and $z_n = w_n l_n$, and x_q, z_n can be observed from the data.

On the technology side, bear in mind that we have given the aggregate production function in Appendix A.1,

$$Y = \left\{ \int \beta_n Y_n^{\rho} dn \right\}^{\frac{1}{\rho}}, \quad \text{with} \quad Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}.$$

We specify the distributional parameter $\beta_n = f_n^{\rho}(n)$. Profit maximizing means that $w_n = \partial Y / \partial L_n$, one can obtain the following equation,

$$w_n^{1-\rho} = Y^{1-\rho}(1-\alpha_n)^{1-\rho} L_n^{\rho-1} p_n^{-\rho} f_n^{\rho}(n) \tag{I.1}$$

Where we use $Y_n = \frac{w_n L_n}{1-\alpha_n} \frac{1}{p_n}$, the rest of work is to find p_n. Substitute $K_n = \alpha_n p_n Y_n / R$ and $L_n = (1-\alpha_n) p_n Y_n / w_n$ into $Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}$, one obtains

$$A_n(\alpha_n) \alpha_n^{\alpha_n} (1-\alpha_n)^{1-\alpha_n} \left(\frac{R}{w_n} \right)^{1-\alpha_n} \frac{p_n}{R} = 1.$$

Rearrange above equation, p_n can be given by

$$p_n = \frac{\left(\frac{w_n}{R}\right)^{1-\alpha_n} R}{e^{\int_0^{\alpha_n} \ln \psi_n^k(i) di + \int_{\alpha_n}^1 \ln \psi_n^l(i) di}}. \tag{I.2}$$

Note that we have $A_n(\alpha_n) = \frac{e^{\int_0^{\alpha_n} \ln \psi_n^k(i) di + \int_{\alpha_n}^1 \ln \psi_n^l(i) di}}{\alpha_n^{\alpha_n} (1-\alpha_n)^{1-\alpha_n}}$. Without loss of generality, we assume $\psi_n^k(i) = 1$ and $\psi_n^l(i) = \delta_n \cdot i^{\eta}$, thus the price of occupational output can be reduced to $p_n = \frac{w_n^{1-\alpha_n} R_n^{\alpha}}{\delta_n^{1-\alpha_n} e^{-\eta(1-\alpha_n)} (\alpha_n)^{-\eta\alpha_n}}$. In the context of endogenous automated technical change, we have $\frac{w_n}{R} = \frac{\psi_n^l(\alpha_n)}{\psi_n^k(\alpha_n)} = \delta_n \alpha_n^{\eta}$ in equilibrium, thus the price can be reduced to $p_n = e^{\eta(1-\alpha_n)} R \alpha_n^{\eta}$.

Substitute $p_n = e^{\eta(1-\alpha_n)} R \alpha_n^{\eta}$ into (I.1), we get,

$$w_n^{1-\rho} = Y^{1-\rho}(1-\alpha_n)^{1-\rho} L_n^{\rho-1} (R\alpha_n^{\eta})^{-\rho} f_n^{\rho}(n) e^{-\eta\rho(1-\alpha_n)}, \quad \forall n. \tag{I.3}$$

In addition, set the labor share to 52.7%, the following equation must be satisfied,

$$\frac{\int w_n L_n f_n(n)dn}{\int \frac{\alpha_n w_n L_n}{1-\alpha_n} f_n(n)dn} = \frac{52.7\%}{1-52.7\%}. \tag{I.4}$$

Equation (I.1) and (I.4) are used to calibrate the set of automation $\{\alpha_n\}_N$ and the comparative parameter η. Then $\{\delta_n\}_N$ can be calibrated using $\frac{w_n}{R} = \frac{\psi_n^l(\alpha_n)}{\psi_n^k(\alpha_n)} = \delta_n \alpha_n^\eta$. After that, it is easily to calculate $A_n(\alpha_n)$, γ_n and K_n/L_n.

G.2 Simulation

The simulation of optimal tax system is just following the fixed-point algorithm introduced by Mankiw et al. (2009).

Step 1: Given the tax systems, we solve the equilibrium. Namely, We find a set of prices $\{w_n\}_{n\in N}$ and R, allocations $\{l_n, a_q\}_{n\times q\in N\times Q}$, and automation levels $\{\alpha_n\}_{n\in N}$ such that the first-order conditions for workers and firms are satisfied. Moreover factors and final goods markets must be clearing.

Starting from a given $\{w_n\}_{n\in N}$ and R, we use the two first-order conditions of individuals to indicate optimal factor supplies $\{l_n, a_q\}_{n\times q\in N\times Q}$,

$$l_n^{1/\epsilon_l} = w_n(1-T_z'), \quad B_n a_n^{1/\epsilon_k} = (1-T_x')R.$$

then we are allowed to update the level of automation using equation (I.3), and calculate alternative p_n, K_n and Y_n using the following equations,

$$p_n = e^{\eta(1-\alpha_n)} R \alpha_n^\eta, \qquad K_n = \frac{\alpha_n w_n l_n}{1-\alpha_n}\frac{1}{R}, \quad Y_n = A_n(\alpha_n) K_n^{\alpha_n} L_n^{1-\alpha_n}.$$

We calculate iterative aggregate output Y and capital K using

$$Y_{iter} = \int p_n Y_n f_n(n)dn, \quad K_{iter} \int K_n f_n(n)dn.$$

and iterative government revenue B and social welfare W using

$$B_{iter} = \int z_n T_z'(z_n) f_n(n)dn + \int x_q T_x'(x_q) f_q(q)dq, \quad W_{iter} = \frac{1}{\lambda}\int\int \frac{V^{1-\kappa}(n,q)}{1-\kappa} f(n,q)dndq + B.$$

Finally, we update factor prices using

$$R = \frac{\alpha Y}{K}, \quad w_n = \frac{(1-\alpha_n)RK_n}{\alpha_n L_n}.$$

where $\alpha = \int \alpha_n \gamma_n dn$. In equilibrium, Y_{iter}, K_{iter}, B_{iter} and W_{iter} must converge, which is our condition to terminate the first loop.

Step 2: After calculating the equilibrium results, we will further calculate the sufficient statistics required for the optimal tax expressions. The supply-side elasticities $\epsilon_{l_n,1-T_z'}$ and $\epsilon_{a_n,1-T_z'}$ can be found in Table 1, and the demand-side elasticities have been given in Table 2. The remaining work is to update the social welfare weights,

$$g(n,q) = \frac{V^{-\kappa}(n,q)}{\lambda}, \quad \bar{g}_{z_n} = \frac{\int\int_{z_{n'}>z_n} g(n',q)f(n',q)dn'dq}{1-F_z(z_n)}, \quad \bar{g}_{x_q} = \frac{\int\int_{x_{q'}>x_q} g(n,q')f(n,q')dndq'}{1-F_x(x_q)}.$$

and the inverse Hazard Ratio $\frac{1-F_z(z_n)}{f_z(z_n)z_n}$ and $\frac{1-F_x(x_q)}{f_x(x_q)x_q}$.

Step 3: Given above sufficient statistics, compute the alternative tax schedules using the planner's first-order condition, i.e., the optimal tax formulas. For NLIT-NCIT tax system, we use

$$\begin{aligned}\frac{T_z'(z_{n^*})}{1-T_z'(z_{n^*})} &= \frac{1}{\epsilon_{l_{n^*},1-T_z'}}(1-\bar{g}_{z_{n^*}})\frac{1-F_z(z_{n^*})}{z_{n^*}f_z(z_{n^*})} \\ &\underbrace{-\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\left[(1-\bar{g}_{z_n})\left(\frac{1-T_z'(z_n)}{1-T_z'(z_{n^*})}\right)\left(\frac{1-F_z(z_n)}{z_{n^*}f_z(z_{n^*})}\right)z_n\right]'\epsilon^{SE}_{w_n,L_{n^*}}dz_n}_{\text{Substitution Effect}} \\ &\underbrace{-\left(\frac{dz_{n^*}}{dn^*}\right)^{-1}\int\left[(1-\bar{g}_{z_n})\left(\frac{1-T_z'(z_n)}{1-T_z'(z_{n^*})}\right)\left(\frac{1-F_z(z_n)}{z_{n^*}f_z(z_{n^*})}\right)z_n\right]'\epsilon^{AE}_{w_n,L_{n^*}}dz_n}_{\text{Automation Effect}}\end{aligned}$$

$$\begin{aligned}\frac{T_x'(x_{q^*})}{1-T_x'(x_{q^*})} &= \frac{1}{\epsilon_{a_{q^*},1-T_x'}}(1-\bar{g}_{x_{q^*}})\frac{1-F_x(x_{q^*})}{x_{q^*}f_x(x_{q^*})} \\ &\underbrace{-\frac{1}{RK(1-T_x'(x_{q^*}))}\int\left[(1-\bar{g}_{z_n})\left(1-T_z'(z_n)\right)(1-F_z(z_n))z_n\right]'\epsilon^{SE}_{w_n,K}dz_n}_{\text{Substitution Effect}} \\ &\underbrace{-\frac{1}{RK(1-T_x'(x_{q^*}))}\int\left[(1-\bar{g}_{z_n})\left(1-T_z'(z_n)\right)(1-F_z(z_n))z_n\right]'\epsilon^{AE}_{w_n,K}dz_n}_{\text{Automation Effect}}\end{aligned}$$

Turn to NLIT-LCIT tax system, we consider an alternative optimal capital income tax formula

$$\frac{t_x}{1-t_x} = \frac{\int (1-g_q(q))x_q f_q(q)dq}{\int \epsilon_{a_q,1-T_x'} x_q f_q(q)dq}$$
$$\underbrace{-\int \frac{1}{RK(1-t_x)} \left[(1-\bar{g}_{z_n})(1-T_z'(z_n))(1-F_z(z_n))z_n\right]' \epsilon^{SE}_{w_n,K} dz_n}_{\text{Substitution Effect}}$$
$$\underbrace{-\int \frac{1}{RK(1-t_x)} \left[(1-\bar{g}_{z_n})(1-T_z'(z_n)(1-F_z(z_n))z_n\right]' \epsilon^{AE}_{w_n,K} dz_n}_{\text{Automation Effect}}$$

When the nonlinear labor income tax is restricted to be the following functional form

$$T_z(z_n) = z_n - \frac{1-\tau}{1-\phi} z_n^{1-\phi},$$

we calculate the alternative capital income tax rate using above NCIT and LCIT formulas.
Step 4: Repeat above steps until the updating is negligible. When the fixed point system is solved, the resource constraint must be satisfied

$$\int\int c(n,q)f(n,q)dndq + B = Y + K.$$