Redistributive Income Taxation with Directed Technical Change
– Preliminary and Incomplete –

Jonas Loebbing*

February 2019

This paper studies the implications of (endogenously) directed technical change for the design of non-linear labor income taxes in a Mirrleesian economy augmented to include endogenous technology development and adoption choices by firms. First, I identify conditions under which any progressive tax reform induces equalizing technical change, that is, technical change that compresses the pre-tax wage distribution. The key intuition is that progressive tax reforms tend to reduce labor supply of more skilled relative to less skilled workers, while the increased relative supply of less skilled workers induces firms to develop and use technologies that are more complementary to less skilled workers. Second, I provide conditions under which the endogenous response of technology raises the welfare gains from progressive tax reforms. Third, I show that the endogenous technical change effects tend to make the optimal tax scheme more progressive, raising marginal tax rates at the right tail of the income distribution and lowering them (potentially below zero) at the left tail. Finally, I aim to assess the new effects quantitatively, which is still in progress.

JEL: H21, H23, H24, J31, O33 Keywords: Optimal Taxation, Directed Technical Change, Endogenous Technical Change, Wage Inequality.

1. Introduction

Technical change is widely considered an important determinant of changes in the wage structure of an economy and hence of first-order importance for the design of redistributive tax schemes. Existing work analyzes how redistributive taxes respond optimally to exogenous changes in production technology that affect the wage distribution (e.g. Ales, Kurnaz and Sleet, 2015). But technologies are developed and adopted by firms pursuing economic objectives (cf. Acemoglu, 1998, 2007), so they should respond to perturbations of the economy such as tax reforms. In particular, previous work on directed technical change has theoretically proposed and empirically substantiated that the supply of skills in an economy is an important determinant of the extent to which technology favors skilled workers and thereby raises

*Center for Macroeconomic Research, University of Cologne. Website: https://cmr.uni-koeln.de/en/team/phd-students/research-teaching-assistant/loebbing/. Email: loebbing[at]wiso.uni-koeln.de. Telephone: +49 221 470 8650.
wage inequality (e.g. Acemoglu, 1998; Morrow and Trefler, 2017; Carneiro, Liu and Salvanes, 2018). At the same time a large literature on redistributive taxation shows that (non-linear) labor income taxes distort the supply of labor at different levels of skill in a quantitatively significant way (cf. Saez, Slemrod and Giertz, 2012). Changes in labor income taxes should thus be expected to induce changes in technology, which in turn affect pre-tax wage inequality. 

Taking into account these technology responses in the analysis of tax policy seems an important task for taxation theory.

This paper therefore analyzes the design of non-linear labor income taxes when technology is determined endogenously through the profit-maximizing decisions of firms. For that purpose, I develop a general but tractable model of the economy that features both endogenous labor supply of a continuum of differentially skilled workers and endogenous technology development and adoption choices of firms. 

In the model, technical change is driven by technology firms’ decisions in which type of technology to invest. Some types are more complementary to high-skilled workers, some are more complementary to low-skilled workers. Technology firms’ investment decisions depend on final good firms’ demand for intermediate goods that embody the different types of technologies. This intermediate good demand in turn crucially depends on the structure of labor supply firms face on the labor market. If there is a relatively large supply of low-skilled workers, firms demand technologies that are relatively complementary to the low-skilled; if the supply of high-skilled workers is relatively large, firms demand more skill-biased technologies.

Income taxes interact with directed technical change via the structure of labor supply. For example, raising marginal tax rates for high incomes and reducing them for low incomes discourages labor supply of high-skilled and encourages labor supply of low-skilled workers. This shifts firms’ demand towards less skill-biased technologies, to which technology firms respond by shifting investment towards such technologies. Intuitively, progressive tax reforms should therefore induce technical change in favor of less skilled workers.

I examine this intuition formally and investigate its implications for the welfare effects of tax reforms and the design of optimal taxes. To this end, I first show that the model’s equilibrium has a parsimonious reduced form, which makes the tax analysis tractable. Importantly, this reduced form is well studied by the theory of directed technical change (e.g. Acemoglu, 2007; Loebbing, 2018). Moreover, Acemoglu (2007) shows that this reduced form applies to a large set of directed technical change models studied in the literature. This makes my tax analysis, which is based on the reduced form, generic within the theory of directed technical change.

Turning to the analysis of income taxes, I first expose my results in a simple version of the model with only two different types of workers. Then I extend the results to a version with a continuum of worker types. Besides generating additional qualitative insights, the continuum version allows me to quantify my results, using empirical estimates of directed technical change effects from the literature.

In the first step I study the effects of tax reforms on the direction of technical change. In line with the intuition developed above, I find that, under certain conditions, progressive tax reforms induce technical change that compresses the wage distribution.
In the next step I analyze how the welfare effects of tax reforms are affected by the presence of directed technical change. Somewhat surprisingly given the preceding result, I find that directed technical change does not unambiguously raise the welfare gains from a given progressive reform. This is because the positive redistributive effect of the induced technical change may be counteracted by an adverse effect on tax revenue if the induced technical change shifts wage income from workers with high to workers with low marginal tax rates. This resembles the finding of Sachs, Tsyvinski and Werquin (2017) that accounting for substitution effects between workers – as first analyzed by Stiglitz (1987) in optimal taxation – does not necessarily lower the welfare gain from progressive reforms because of the associated revenue effects.

I show however that once we consider the scope for welfare improvements by means of progressive reforms instead of the welfare effects of a given reform, the results align with the preceding findings and the intuition given above. In particular, I find that the set of tax schedules that can be improved in terms of welfare by means of progressive reforms increases when taking into account directed technical change. The result is based on a comparison of the perspectives of an endogenous and an exogenous technology planner, where the exogenous technology planner believes, mistakenly within the model, that technology is exogenously fixed. This comparison provides a useful new tool for precisely identifying the role of wage endogeneity in the analysis of income taxes.

Next I characterize optimal tax rates in the presence of directed technical change and compare the results to the optimal tax rates of the exogenous technology planner, that is, the rates perceived as optimal when not taking into account the endogeneity of technology. I show that directed technical change renders optimal taxes more progressive in the sense that it raises the optimal marginal tax rates in the upper tail and lowers them in the lower tail of the income distribution. The optimal marginal tax rates in the lower tail may even become negative.

From a technical point of view, I adapt the approach of Diamond (1998) in the analysis of optimal taxes to environments with endogenous wages. If the aggregate production function takes a CES form, this approach yields a closed-form expression for optimal marginal tax rates, including adjustments for wage endogeneity, which may be useful for the analysis of optimal taxes in general equilibrium environments more generally.

Finally I aim to quantify my results on tax reforms and optimal taxes, using estimates of directed technical change effects from the empirical literature on directed technical change. This part of the paper is still in progress.

The structure of the paper is as follows. Section 2 presents the model and introduces notation. Section 3 states important results from the theory of directed technical change, which provide the basis for the analysis in the present paper. Section 4 contains the analysis of tax reforms. Section 4.1 considers the effects of tax reforms on directed technical change and Section 4.2 considers the impact of directed technical change on the welfare effects of tax reforms. Section 5 analyzes optimal taxes, Section 6 quantifies the results from the preceding sections (in progress, not yet available), and Section 7 concludes.

---

1I use a general Bergson-Samuelson welfare function and impose a mild condition ensuring that the welfare function values equity across workers.
Related Literature  The paper connects the literatures on the optimal design of non-linear labor income taxes and on (endogenously) directed technical change. It is the first to incorporate endogenous technology responses into an analysis of non-linear labor income taxation and the first to rigorously explore normative implications of the theory of directed technical change.

Starting from the literature on optimal taxation, the paper extends Sachs et al. (2017) to incorporate endogenously directed technical change, showing that endogenous technology responses counteract the standard substitution effects between skills analyzed by Sachs et al. (2017). Compared to Ales et al. (2015), who analyze how exogenous technical change affects the optimal tax schedule, I treat technical change as endogenous such that it responds to changes in the tax system.

The paper is also related to recent studies of the taxation of robots (Guerreiro, Rebelo and Teles, 2018; Thuemmel, 2018). In contrast to these studies, I show that technology can be affected indirectly through the income tax system without having to tax specific technologies directly, which might be challenging in practice due to informational and administrative constraints.

Starting from the theory of directed technical change, I build on the seminal ideas of Acemoglu (1998) and Kiley (1999) and explore their normative implications, in particular for the design of labor income taxes. Thereby, I make use of the advances in the theory of directed technical change by Acemoglu (2007) and Loebbing (2018).

2. Setup

The model is a general equilibrium model with endogenous production technology that is embodied in intermediate inputs. The intermediate inputs are supplied under monopolistic competition as, for example, in Romer (1990). The monopolistically competitive suppliers can improve the quality of their intermediate goods by investing R&D resources. Crucially, the model features multiple types of technology-embODYING intermediate inputs, which differ in their complementarity relationships with different types of labor. Hence, the wage distribution is affected differentially by improvements in the quality of different types of intermediates. Changes in the distribution of R&D resources over intermediate good types induced by exogenous shocks (such as tax reforms) constitute endogenous technical change.

The tax analysis will build on a reduced form of the model’s equilibrium. While the model itself imposes several specific assumptions, the reduced form is much more general: the same reduced form can be obtained from a variety of different models of endogenous technical change, as discovered in Acemoglu (2007).

2.1. Model

The model features heterogenous workers, perfectly competitive final good firms, monopolistically competitive technology firms, and a government that levies taxes.

---

2See Acemoglu (2007) and Loebbing (2018) for complementary lists of models that all give rise to the set of reduced form equations derived below.
Workers  There is a continuum of workers with different types $\theta \in \Theta$. I will study two cases:

(i) a two type model where $\Theta = \{\bar{\theta}, \bar{\theta}\}$,
(ii) a continuum of types, $\Theta = [\underline{\theta}, \bar{\theta}]$.

Types are distributed according to the density function $h : \theta \mapsto h_\theta$, with cumulative distribution function $H$.

Workers’ utilities depend on consumption $c_\theta$ and labor supply $l_\theta$ according to

$$u_\theta(c_\theta, l_\theta) = c_\theta - v(l_\theta),$$

that is, there are no income effects. The disutility of labor $v$ is $C^2$, strictly increasing, and strictly convex.

Workers’ pre-tax incomes are $y_\theta = w_\theta l_\theta$ and income taxes are given by the tax function $T : y_\theta \mapsto T(y_\theta)$. The retention function corresponding to tax $T$ is denoted $R_T$. Hence, workers’ budget constraints are

$$c_\theta = R(w_\theta l_\theta) + S,$$  \hspace{1cm} (1)

where $S$ is a lump-sum transfer used to neutralize the government’s budget constraint.

Workers choose their labor supply to maximize utility, taking wages as given. The first-order condition is

$$v'(l_\theta) = R'(w_\theta l_\theta)w_\theta.$$  \hspace{1cm} (2)

Firms  There is a continuum of mass one of identical final good firms indexed by $i$. They produce a final consumption good (the numéraire) according to the production function $G(L_i, Q_i)$.

The first input $L_i = \{L_{i,\theta}\}_{\theta \in \Theta}$ collects the amounts of all different types of labor used by firm $i$. The second input $Q_i = \{Q_{i,j}\}_{j \in \{1,2,...,J\}}$ collects the variables $Q_{i,j}$, each of which is an aggregate of a continuum of technology-embodying intermediate goods:

$$Q_{i,j} = \int_0^1 \phi_{i,j,k} q_{i,j,k}^\alpha dk.$$

The variables $q_{i,j,k}$ denote the amount of intermediate good $(j, k)$ used by firm $i$, while the parameter $\alpha \in (0,1)$ governs the substitutability of intermediates with the same $j$-index. The variables $\phi_{i,j,k}$ give the quality of the corresponding intermediate goods and they represent the endogenous part of production technology in the model. Their determination is described in detail below.

With this structure of final good production, we can write the output of firm $i$ as $\hat{G}(L_i, \phi, q_i)$ where $\phi = \{\phi_{i,j,k}\}_{(i,j,k) \in \{1,2,...,J\} \times [0,1]}$ and $q_i = \{q_{i,j,k}\}_{(i,j,k) \in \{1,2,...,J\} \times [0,1]}$ collect qualities and quantities of all different intermediates. I assume that the function $\hat{G}$ is linear homogeneous and concave in the rival inputs $(L, q)$, satisfying the standard microeconomic replication argument (e.g. Romer, 1994). Since in addition all final good firms are price-takers, the final good sector admits a representative firm, so I drop the index $i$ in what follows.

\footnote{The case with a continuum of different intermediate good types $j, j \in [0,J]$, can be treated analogously.}
Final good firms’ profit maximization leads to the following demand for labor:

\[ w_\theta = D_{L_\theta} \tilde{G}(L, \phi, q). \]

In the case with two types of labor, \( D_{L_\theta} \) means standard partial differentiation with respect to \( L_\theta \). If there is a continuum of labor types, \( D_{L_\theta} \) denotes Gateaux differentiation with respect to \( L \) in direction of the Dirac measure at \( \theta \), which I define rigorously in Section 2.2. Labor market clearing requires that the aggregate labor demand \( L_\theta \) equals the sum of individual workers’ labor supply,

\[ L_\theta = l_\theta f_\theta \quad \text{for all } \theta. \]

Demand for intermediate good \( q_{j,k} \) is given by

\[ p_{j,k} = \alpha \phi_{j,k} q_{j,k}^{\alpha-1} \frac{\partial G(L, Q)}{\partial Q_j}, \quad (3) \]

where \( p_{j,k} \) is the price of the intermediate good.

The technology-embodying intermediate goods are produced under monopolistic competition by technology firms. Each good \((j,k)\) is produced by a single technology firm, which I label by the index \((j,k)\) of its output. Technology firm \((j,k)\) produces its output at constant marginal cost \( \eta_j \) from final good and receives an ad valorem sales subsidy of \( \xi \). It sets the post-subsidy price \( p_{j,k} \) to maximize profits

\[ \left((1+\xi)p_{j,k} - \eta_j\right) q_{j,k} \]

subject to the demand from final good firms (equation (3)). Since the demand from final good firms is iso-elastic, the profit maximizing price is given by a constant markup over marginal cost net of the subsidy:

\[ p_{j,k} = \frac{\eta_j}{(1+\xi)\alpha}. \quad (4) \]

Technology firms can invest R&D resources to improve the quality of their output. In particular, a quality level of \( \phi_{j,k} \) costs \( C_j(\phi_{j,k}) \) units of R&D resources, where the cost function \( C_j \) is convex and strictly increasing for every \( j \). Firm \((j,k)\)’s profits as a function of its quality level \( \phi_{j,k} \) are

\[ \pi_{j,k}(\phi_{j,k}) = \max_q \left\{ \alpha \phi_{j,k} \frac{\partial G(L, Q)}{\partial Q_j} q^\alpha - \eta_j q - p' C_j(\phi_{j,k}) \right\}. \]

where \( p' \) denotes the (competitive) market price of R&D resources. Via an envelope argument, the first-order condition for the choice of quality is given by

\[ \alpha \frac{\partial G(L, Q)}{\partial Q_j} q_{j,k}^\alpha = \frac{dC_j(\phi_{j,k})}{d\phi_{j,k}}, \]

where \( q_{j,k} \) is assumed to take its profit maximizing value implied by equation (4). One can verify that the optimal \( q_{j,k} \) grows at the rate \( 1/(1-\alpha) \) in \( \phi_{j,k} \), such that the left-hand side of equation (2.1) grows at rate \( \alpha/(1-\alpha) \) in \( \phi_{j,k} \). I assume henceforth that \( dC_j/d\phi_{j,k} \) grows at a rate greater than \( \alpha/(1-\alpha) \) in \( \phi_{j,k} \), which ensures that the first-order condition identifies the unique profit maximum. Since profits are symmetric across all firms \((j,k)\) with the same
j-index, uniqueness of the profit maximum implies that the choices of all firms with index \( j \) are the same and we can drop the \( k \)-index henceforth. The supply of R&D resources is exogenous and given by \( C \). Their price adjusts to guarantee market clearing,

\[
\sum_{j=1}^{J} C_j(\phi_j) = \bar{C}.
\]

**Government** The government levies different types of taxes/subsidies. First, it subsidizes the sale of technology-embodying intermediate goods by the ad valorem subsidy \( \xi \) to counteract the inefficiency created by the market power of technology firms. I assume that \( \xi = (1 - \alpha) / \alpha \), such that post-subsidy prices of intermediate goods equal marginal costs, \( p_j = \eta_j \).\(^4\) This implies that absent any other taxes, the equilibrium allocation will be efficient. Hence, income taxes are used for redistributive purposes only and do not contain any Pigouvian elements, which would deviate the focus away from the central points of the paper. Importantly, the government is restricted to impose a uniform subsidy across all technology types \( j \). This precludes policies aimed at changing the relative utilization of different technologies to reduce pre-tax wage inequality, as analyzed, for example, in Thuemmel (2018) and Guerreiro et al. (2018).

Second, the government taxes the profits of technology firms and of the owners of R&D resources. As is standard in the literature on labor income taxation, I assume that these taxes are confiscatory to avoid a role for the distribution of firm ownership without a meaningful theory of wealth formation in the model.\(^5\) Alternatively, I could assume that firm ownership and ownership of R&D resources are uniformly distributed across workers without changing any of the results.

Third, the government taxes income according to the tax function \( T \). Reforms of \( T \) and the characterization of its optimal shape are the central objects of the paper. Taken together, taxes and subsidies generate the following government revenue,

\[
S(y) = \int_{\Theta} T(y_{\theta}) h_{\theta} d\theta + p^r \Phi + \sum_{j=1}^{J} \pi_j - \sum_{j=1}^{J} \xi p_j q_j,
\]

which is redistributed lump-sum across workers.

**Equilibrium** An equilibrium of the model, given a tax function \( T \), is a collection of quantities and prices such that all firms maximize profits, workers maximize utility, and all markets clear. Despite the detailed micro structure of the model, its equilibrium admits a simple reduced form. To derive this reduced form, note first that aggregate production at labor input \( l \) and a given set of quality levels \( \phi \) can be written as (because intermediate goods prices equal

---

\(^4\)This level of subsidies would also be chosen as part of the optimal tax policy if it were included in the optimal tax analysis of Section 5.

\(^5\)Note that confiscatory profit taxes are part of the optimal tax policy whenever ownership shares of firms increase and marginal welfare weights decrease in workers’ income levels at the optimum.
marginal costs):
\[ F(l, \phi) := \max_q \left\{ \tilde{G} \left( \{ f_\theta l \} \theta \in \Theta, \phi, q \right) - \sum_{j=1}^{l} \eta_j q_j \right\}. \]

Note that I used labor market clearing (equation (2.1)) to replace the aggregate labor input \( L \) by the individual labor input \( l \) to save on notation in the following. Via an envelope argument, the labor demand equation (2.1) then implies that in equilibrium wages are given by
\[
 w_\theta(l, \phi) = D_l F(l, \phi) \frac{1}{h_\theta}, \tag{5}
\]
where the adjustment factor \( 1/h_\theta \) is necessitated by the switch from aggregate to individual labor inputs in the aggregate production function.

The condition for profit-maximizing quality choices of technology firms (equation (2.1)) coincides with the first-order condition for a maximum of aggregate production with respect to quality \( \phi \) (simply called technology, henceforth) when \( \phi \) is restricted to the set of feasible technologies \( \Phi = \{ \phi \in \mathbb{R}_+^J | \sum_{j=1}^{J} C_j(\phi_j) \leq \mathcal{C} \} \). Thus,
\[
 \phi^* (l) := \arg\max_{\phi \in \Phi} F(l, \phi) \tag{6}
\]
is an equilibrium technology. In the following I focus on equilibria in which technology satisfies equation (6). Existence of other equilibria can be ruled out by imposing assumptions that guarantee strict quasiconcavity of \( F \) in \( \phi \) under the constraint \( \phi \in \Phi \).

Finally, we can simplify the expression for the government’s budget surplus. To this end, note that marginal cost pricing of intermediate goods implies that technology firms’ profits are equal to the total amount of subsidies minus the cost for R&D resources:
\[
 \sum_{j=1}^{l} \pi_j = \sum_{j=1}^{l} \left( (1 + \xi_j) p_j - \eta_j \right) q_j - p' \mathcal{C} = \sum_{j=1}^{l} \xi_j p_j q_j - p' \mathcal{C}. \]

It follows that the revenue from corporate taxes and the expenses on technology good subsidies offset each other exactly in equation (2.1), such that government revenue reduces to
\[
 S(y) = \int_\Theta T(y_\theta) h_\theta d\theta. \tag{7}
\]

We can now define a reduced form equilibrium for a given income tax \( T \) as a collection of labor inputs \( l \), a technology \( \phi \), government revenue \( S \), consumption levels \( c \), and wages \( w \), such that workers’ first-order conditions (2), their budget constraints (1), the wage equation (5), the technology condition (6), and the equation for government revenue (7) are satisfied. These

\[ F(l, \phi_{-j}) := F(l, \phi_{-j}, \tilde{\phi}_j(\phi_{-j})), \text{ where } \phi_{-j} = \{ \phi_j \}_{j \in \{1,2,\ldots,j-1\}} \text{ and } \tilde{\phi}_j(\phi_{-j}) = C_j^{-1} \left( \mathcal{C} - \sum_{j=1}^{j-1} C_j(\phi_j) \right), \]
is strictly quasiconcave in \( \phi_{-j} \), the first-order conditions for a maximum of \( F \) in \( \phi_{-j} \) are necessary and sufficient and there is a unique value \( \phi^*_{-j}(l) \) that satisfies them. Equivalently, there is a unique value \( \phi^*(l) \) satisfying the first-order conditions of the program (6), which are identical to the equilibrium condition (2.1).
reduced form equations provide the starting point for the tax analysis in the next sections.

2.2. Notation

The tax analysis uses functional derivatives and various elasticities. To simplify the exposition I define a specific notation for several frequently used expressions.

**Derivatives** For derivatives in finite-dimensional spaces I use standard notation. For perturbations of the tax function $T$ and labor input $l$ (if the type space is a continuum) I will frequently use the following functional derivatives.

Let $x : (T, z) \mapsto x(T, z)$ be a function of the tax $T$ and, potentially, further variables $z$. Then,

$$D_\tau x(T, z) := \left. \frac{dx(T + \mu \tau, z)}{d\mu} \right|_{\mu=0}$$

denotes the directional derivative of $x$ with respect to $T$ in direction of the tax reform $\tau$.

Relatedly, let $x : (l, z) \mapsto x(l, z)$ be a function of labor input $l$ and, potentially, further variables $z$. Suppose that the type space $\Theta$ is a continuum, such that $l$ is a non-negative, real-valued function on $[\theta, \theta]$. I formalize the derivative of $x$ with respect to labor supply of a given type $\theta$, $l_{\theta}$, as

$$D_{l_\theta} x(l, z) := \lim_{\Delta \to 0} \frac{1}{\Delta} \left. \frac{dx(l + \mu \tilde{l}_\theta, z)}{d\mu} \right|_{\mu=0},$$

where $\tilde{l}_\theta : \bar{\theta} \mapsto \tilde{l}_{\theta, \bar{\theta}}$ is a real-valued function on the type space given by

\[
\tilde{l}_{\theta, \bar{\theta}} = \begin{cases} 
0 & \text{for } \bar{\theta} < \theta - \Delta \\
\frac{\bar{\theta} - \theta + \Delta}{\Delta} & \text{for } \bar{\theta} \in [\theta - \Delta, \theta] \\
\frac{\theta - \bar{\theta} + \Delta}{\Delta} & \text{for } \bar{\theta} \in [\theta, \theta + \Delta] \\
0 & \text{for } \bar{\theta} > \theta + \Delta 
\end{cases}
\]

for all interior types $\theta \in (\bar{\theta}, \bar{\theta})$, by

\[
\tilde{l}_{\theta, \bar{\theta}} = \begin{cases} 
0 & \text{for } \bar{\theta} < \bar{\theta} - \Delta \\
\frac{2(\bar{\theta} - \theta + \Delta)}{\Delta} & \text{for } \bar{\theta} \in [\bar{\theta} - \Delta, \bar{\theta}] 
\end{cases}
\]

for the highest type $\bar{\theta}$, and by

\[
\tilde{l}_{\theta, \bar{\theta}} = \begin{cases} 
\frac{2(\theta - \bar{\theta} + \Delta)}{\Delta} & \text{for } \bar{\theta} \in [\theta, \theta + \Delta] \\
0 & \text{for } \bar{\theta} > \theta + \Delta 
\end{cases}
\]

for the lowest type $\theta$. Intuitively, the derivative is obtained by perturbing the labor supply function in a continuous way in a neighborhood of type $\theta$ and letting this neighborhood converge to $\theta$. Appendix B.1 demonstrates that the thus defined derivative works in a natural manner.

---

7 The derivative of a function with respect to aggregate labor supply $L_\theta$ is defined analogously.
way by showing in detail that

\[ D_{l_\theta} \int_\theta w_{l_\theta} L_{l_\theta} d\theta = w_\theta \quad \forall \theta. \]

This also proves the labor demand equation (2.1) for the case of a type continuum.

The tax analysis below often distinguishes between the direct effect of changes in \( T \) or \( l \) on an outcome \( x \) and the indirect effect mediated through the response of technology \( \phi^* \). In particular, suppose \( x : (T, \phi) \to x(T, \phi) \) depends (directly) on taxes \( T \) and technology \( \phi \). The direct effect of a tax reform in direction \( \tau \), holding technology fixed, is then given by \( D_{\tau} x(T, \phi) \) as defined above. For the indirect effect of the tax reform via technology (the induced technical change effect, henceforth) I introduce the following notation:

\[ D_{\phi, \tau} x(T, \phi^*(T)) := \left. \frac{dx(T, \phi^*(T + \mu \tau))}{d\mu} \right|_{\mu=0}. \]

Here, \( \phi^*(T) \) denotes the equilibrium technology at tax function \( T \). The total effect of the reform on \( x \) then obtains as the sum of the direct and the induced technical change effect. Writing \( x^*(T) := x(T, \phi^*(T)) \), we get

\[ D_{\tau} x^*(T) = D_{\tau} x(T, \phi^*(T)) + D_{\phi, \tau} x(T, \phi^*(T)). \]

Analogously, if \( x : (l, \phi) \to x(l, \phi) \) is a function of labor input \( l \) and technology \( \phi \), the induced technical change effect of a labor input change in direction \( l_\theta \) is

\[ D_{\phi, l_\theta} = \lim_{\Delta \to 0} \frac{1}{\Delta} \left. \frac{dx(l, \phi^*(l + \mu \tilde{l}_\theta))}{d\mu} \right|_{\mu=0}, \]

where \( \phi^*(l) \) is given by equation (6). I assume throughout the paper that these derivatives exist.

**Wage Elasticities**  The response of wages to labor input changes plays a central role in the tax analysis. Consider wages as given by (5), that is, for each type \( \theta \) the wage \( w_\theta \) is a function of labor inputs \( l \) and technology \( \phi \).

The first set of wage elasticities is concerned with the direct effect of labor inputs on wages, holding technology constant. I call these elasticities substitution elasticities, as they describe the changes in marginal productivities induced by factor substitution within a given technology. The own-wage substitution elasticity, that is, the elasticity of \( w_\theta \) with respect to \( l_\theta \), is defined as\(^8\)

\[ \gamma_{\theta, \theta} := \lim_{\Delta \to 0} \frac{l_\theta}{w_\theta} \left. \frac{dw_\theta(l + \mu \tilde{l}_\theta, \phi)}{d\mu} \right|_{\mu=0}. \]

Alternatively, we could write the wage \( w_\theta \) as a function of \( \phi \), \( l \), and type \( \theta \)'s labor input \( l_\theta \) separately, as typically a type's labor input affects its own wage in a way distinct from the labor input function \( l \) (see for example the CES case in Section 2.3). Then, the own-wage

\(^8\)Here I again consider the case of a type continuum. If there are only two types, the definitions of wage elasticities are straightforward and thus omitted.
substitution elasticity is simply
\[ \gamma_{\theta,\theta} = \frac{l_\theta}{w_\theta} \frac{\partial w_\theta}{\partial l_\theta}(l_\theta, l, \phi). \]

The cross-wage substitution elasticity, that is, the elasticity of \( w_\theta \) with respect to a different type’s labor input \( l_{\bar{\theta}} \) (with \( \bar{\theta} \neq \theta \)), is given by
\[ \gamma_{\theta,\bar{\theta}} := \frac{l_{\bar{\theta}}}{w_\theta} D_{l_{\bar{\theta}}} w_\theta(l, \phi), \]
with the derivative \( D_{l_{\bar{\theta}}} \) as defined above.

The second set of wage elasticities captures the induced technical change effects of changes in labor input on wages. These elasticities are called technical change elasticities in the following. The own-wage technical change elasticity is defined as
\[ \rho_{\theta,\theta} := \frac{l_\theta}{w_\theta} \left. \frac{\partial w_\theta}{\partial l_{\theta}}(l, \phi^*(l + \mu l_{\theta})) \right|_{\mu=0}. \]
Again, the CES case in Section 2.3 clarifies why this is a natural definition of the own-wage technical change elasticity and how it can be expressed in terms of conventional partial derivatives.

The cross-wage technical change elasticity measures how wage \( w_\theta \) is affected by a change in another type’s labor supply \( l_{\bar{\theta}} \) via induced technical change. Formally, it is given by
\[ \rho_{\theta,\bar{\theta}} := \frac{l_{\bar{\theta}}}{w_\theta} D_{\phi, l_{\bar{\theta}}} w_\theta(l, \phi^*(l)), \]
where the derivative \( D_{\phi, l_{\bar{\theta}}} \) is as defined above.

**Rate of Progressivity** The rate of progressivity of a tax schedule \( T \) is defined as minus the elasticity of the marginal retention rate \( R'_T \) with respect to income,
\[ P_T(y) := -\frac{R''_T(y)y}{R'_T(y)}. \]
It measures the progression of marginal tax rates as income increases. If the income tax is linear such that marginal tax rates are constant, \( P_T(y) \) is zero. If the income tax is progressive (regressive) in the sense that marginal tax rates increase (decrease) with income, the rate of progressivity is positive (negative).

**Labor Supply Elasticities** As is usual in the literature, I also define some standard concepts of labor supply elasticities to express the effects of tax reforms compactly. The first is the hypothetical elasticity of labor supply with respect to the marginal retention rate that would obtain if the retention function were linear:
\[ e_\theta(l_\theta) := \frac{v'(l_\theta)}{v''(l_\theta)l_\theta}. \]
Consider now the labor supply of an arbitrary type $\theta$ given by workers’ first-order condition (2) as a function of $T$ and $w_\theta$. The true elasticity of labor supply with respect to the marginal retention rate must take into account potential non-linearities of the retention function, which cause the marginal retention rate to change as labor supply changes. This elasticity is given by

$$\epsilon^R_\theta(T, l, w) := \frac{R'_T(w_\theta l_\theta)}{l_\theta} D_t l_\theta(T, w_\theta),$$

where the auxiliary tax reform $\tilde{\tau}$ is chosen such that, as the scaling factor $\mu$ of the reform goes to zero, it raises the marginal retention rate by one infinitesimal unit:

$$\forall y : \tilde{\tau}(y) = -y,$$

and thus:

$$(y - (T(y) + \mu \tilde{\tau}(y)))' = 1 - T'(y) + \mu.$$

Inserting this into workers’ first-order condition and differentiating with respect to $\mu$ (at $\mu = 0$) then gives exactly the local response of individual labor supply to a one unit increase in the marginal retention rate. This leads to the following expression for the elasticity of labor supply with respect to the marginal retention rate (see Appendix B.2 for details):

$$\epsilon^R_\theta(T, l, w) = \frac{e_\theta(l)}{1 + e_\theta(l) P_T(w_\theta l_\theta)}.$$ (8)

For a locally linear tax function, that is, for $P_T(w_\theta l_\theta) = 0$, the elasticity coincides with the hypothetical elasticity $e_\theta$ defined above.

Note that both elasticities measure the response of labor supply of an individual worker, in the sense that they hold the wage $w_\theta$ constant. If, in contrast, all workers of type $\theta$ change their labor supply jointly, the wage $w_\theta$ will react. The wage response then induces an additional change in labor supply. To construct an elasticity that accounts for this feedback effect, define first the elasticity of (individual) labor supply with respect to the wage by

$$\epsilon^w_\theta(T, l, w) := \frac{w_\theta \partial l_\theta(T, w_\theta)}{l_\theta \partial w_\theta}.$$ (9)

It is a standard result that this elasticity can be written as (see Appendix B.2 for details)

$$\epsilon^w_\theta(T, l, w) = \frac{(1 - P_T(w_\theta l_\theta)) e_\theta(l)}{1 + e_\theta(l) P_T(w_\theta l_\theta)}.$$ (9)

From the elasticities of individual labor supply with respect to the marginal retention rate and to the wage we can construct an elasticity of aggregate labor supply of type $\theta$ that accounts for the endogenous response of type $\theta$’s wage:

$$\bar{\epsilon}^R_\theta(T, l, w) := \frac{\epsilon^R_\theta(T, l, w)}{1 - (\gamma_{\theta, \theta} + \rho_{\theta, \theta}) \epsilon^w_\theta(T, l, w)}.$$ (10)

Relative to the individual elasticity, the aggregate elasticity is scaled by the feedback from the wage to labor supply. If the own-wage effect $\gamma_{\theta, \theta} + \rho_{\theta, \theta}$ is negative (positive), the individual elasticity is scaled down (up), as an increase in labor supply depresses (raises) the wage, which then counteracts (amplifies) the initial labor supply change.
The same scaling factor allows to transform the elasticity of individual labor supply with respect to the wage into an aggregate elasticity:

\[
\tilde{\epsilon}^w (T, l, w) := \frac{\epsilon^w (T, l, w)}{1 - (\gamma_{\theta, \theta} + \rho_{\theta, \theta})\epsilon^w (T, l, w)}.
\]  

(11)

2.3. Special Cases

Parts of the results of the tax analysis require further structural assumptions on production functions, utility functions, or the tax function. I introduce these special cases here and refer to them in the tax analysis whenever needed.

**CES Production** An important special case of the model is obtained when the aggregate production function \( F \) features a constant elasticity of substitution (CES) between worker types. Here I describe the CES version of the model for the case of a type continuum. Results for the case with two labor types are analogous.

The CES case is obtained by assuming the following functional forms for final good firms’ production and for research costs:

\[
\tilde{G}(L, \tilde{\phi}, q) = \left[ \int_{\theta}^{\bar{\theta}} \left( \tilde{\kappa}_{\theta} L_{\theta}^{1-a} \int_{0}^{1} \tilde{\phi}_{\theta, k} q_{\theta, k}^a dk \right)^{\frac{1}{\sigma - 1}} d\theta \right]^{\frac{\sigma - 1}{\sigma - 1}},
\]

\[
C_{\theta}(\tilde{\phi}_{\theta, k}) = \tilde{\phi}_{\theta, k}^{\tilde{\delta}_{\theta, k}}.
\]

Note first that I changed the notation for technology from \( \phi \) to \( \tilde{\phi} \) here (for reasons that will become clear later). Second and more substantially, the CES case equates the set of technology types with the set of worker types, such that technology \( \phi \) and research costs \( C \) are now indexed by \( \theta \). This reflects the assumption that for every worker type \( \theta \) there exists a type of technology, embodied in the intermediate goods \( q_{\theta, k} \), that raises the efficiency of labor of type \( \theta \) in the production process.\(^9\)

Appendix B.3 shows that under these assumptions aggregate production can be written as

\[
F(l, \phi) = \left[ \int_{\theta}^{\bar{\theta}} (\kappa_{\theta} \phi_{\theta} l_{\theta} \phi_{\theta})^{\frac{1}{\sigma - 1}} d\theta \right]^{\frac{\sigma - 1}{\sigma - 1}},
\]

(12)

where \( \kappa \) and \( \sigma \) are functions of the parameters \( \tilde{k}, \eta \) (the unit cost of intermediate goods), \( \alpha \), and \( \tilde{\sigma} \). The new technology variable \( \phi \) is a bijection of the original technology \( \tilde{\phi} \), \( \phi_{\theta} := \tilde{\phi}_{\theta}^{1/(1-a)} \) for all \( \theta \). This substitution is done purely for notational convenience. The new technology \( \phi \) satisfies the reduced form equation (6) with the set of feasible technologies \( \Phi \) given by (see Appendix B.3)

\[
\Phi = \left\{ \phi : \theta \mapsto \phi_{\theta} \in \mathbb{R}_+ \mid \int_{\theta}^{\bar{\theta}} \phi_{\theta}^\tilde{\delta}_{\theta, k} d\theta \leq C \right\}.
\]

\(^9\)Note that the set of technology types is a continuum here, in contrast to the finite set \{1, 2, ..., J\} in the general model above. As noted in footnote 3, the case with a continuum of technology types can be treated analogously to the finite case presented above and is therefore omitted from the presentation of the general model.
where \( \delta := (1 - \alpha) \delta > 1 \).

Moreover, following reduced form equation (5), wages in the CES case are (see Appendix (B.3))

\[
  w_\theta(l, \phi) = (\kappa_\theta \phi_\theta)^{\frac{\sigma - 1}{\sigma}} (l_\theta f_\theta)^{-\frac{1}{\sigma}} F(l, \phi)^{\frac{1}{\sigma}}.
\]

Finally and crucially, the wage elasticities defined in Section 2.2 take a particularly simple form in the CES case (see again Appendix B.3 for the derivations). The own-wage substitution elasticity is given by

\[
  \gamma_{\theta, \theta} = -\frac{1}{\sigma} =: \gamma^{\text{CES}}
\]

and the cross-wage substitution elasticity becomes

\[
  \gamma_{\theta, \tilde{\theta}} = \frac{1}{\sigma} l_{\tilde{\theta}} w_{\tilde{\theta}} f_{\tilde{\theta}}.
\]

The own-wage technical change elasticity is

\[
  \rho_{\theta, \theta} = \frac{(\sigma - 1)^2}{(\delta - 1)\sigma^2 + \sigma} =: \rho^{\text{CES}},
\]

and the cross-wage technical change elasticity is given by

\[
  \rho_{\theta, \tilde{\theta}} = -\frac{(\sigma - 1)^2}{(\delta - 1)\sigma^2 + \sigma} \frac{l_{\tilde{\theta}} w_{\tilde{\theta}} f_{\tilde{\theta}}}{F(l, \phi)}.
\]

**Iso-Elastic Disutility of Labor** When the disutility of labor is iso-elastic, workers’ utility functions take the form

\[
  u_\theta = c_\theta - \frac{e}{e + l_\theta^{\frac{\sigma - 1}{\sigma}}}.
\]

In this case, the hypothetical labor supply elasticity \( e_\theta(l) \) is constant across \( \theta \) and \( l \):

\[
  e_\theta(l) = e \quad \text{for all } \theta, l.
\]

This special case plays an important role in simplifying and clarifying the results of tax analysis by suppressing heterogeneity in labor supply responses to changes in marginal tax rates.

**Constant-Rate-of-Progressivity Taxes** A constant-rate-of-progressivity (CRP) tax function takes the form (e.g. Feldstein, 1969; Heathcote, Storesletten and Violante, 2017)

\[
  T(y) = y - \lambda y^{1-P}.
\]

For any CRP tax schedule \( T \) the rate of progressivity \( P_T \) is constant across income levels:

\[
  P_T(y) = P \quad \text{for all } y.
\]

This special case, when combined with iso-elastic disutility of labor, ensures that the labor supply elasticities \( e^K_\theta \) and \( e^W_\theta \) are constant across workers.
3. Directed Technical Change

Income tax reforms affect technical change via differential changes in labor supply across worker types. An important building block of the tax analysis below is therefore the relationship between the structure of labor supply and technical change. This relationship is studied by the theory of directed technical change.

3.1. General Results

Consider wages as a function of labor inputs and technology as given by the reduced form equation (5) and technology as a function of labor inputs as in equation (6).

Two Types Suppose at first that \( \Theta = \{ \theta, \theta \} \), that is, there are only two types of workers. Then, the theory of directed technical change shows that, without any assumptions beyond those imposed in the presentation of the model in Section 2.1, any increase in the relative input of high-skilled over low-skilled labor induces skill-biased technical change. Inversely, any decrease in the relative input of high-skilled over low-skilled labor induces equalizing technical change, that is, technical change that reduces the skill premium. This result is called weak (relative) bias of technology in the theory of directed technical change (e.g. Acemoglu, 2002, 2007).

Lemma 1. Suppose \( \Theta = \{ \theta, \theta \} \). Consider wages as a function of labor inputs and technology as given by equation (5) and technology as a function of labor inputs as given by equation (6). Take any labor input \( l \) and let \( dl \) be change in labor input such that the relative change \( \hat{l}_{\theta} : = \frac{dl_{\theta}}{l_{\theta}} \) is greater at \( \theta \) than at \( \theta \).

Then, the technical change induced by \( dl \) raises \( w_{\theta} \) relative to \( w_{\theta} \):

\[
\sum_{\theta \in \{ \theta, \theta \}} \rho_{\theta, \theta}(l, \phi^{*}(l)) l_{\theta} \geq \sum_{\theta \in \{ \theta, \theta \}} \rho_{\theta, \theta}(l, \phi^{*}(l)) l_{\theta}.
\]

In words, any decrease in the relative input of high-skilled over low-skilled labor induces technical change that reduces the skill premium.


The intuition behind this result relies on the notion of complementarity between labor and technology. A decrease in the relative input of high-skilled over low-skilled workers raises the profitability of technologies that are (relatively) complementary to low-skilled workers, as these worker types are now more abundant. But since complementarity is a symmetric relation, technologies that are relatively complementary to low-skilled workers also raise the relative productivity of low-skilled workers. The technical change induced by a decrease in the relative input of high-skilled workers will therefore reduce the skill premium.

\(^{10}\)The only important assumption from Section 2.1 for the directed technical change results presented in the following is that final good firms’ production function \( \tilde{G} \) is linear homogeneous in the rival inputs \( l \) and \( q \). This assumption ensures that aggregate production \( F \) is linear homogeneous in \( l \), which in turn is sufficient to apply the directed technical change results from Loebbing (2018) in the two type case.
Type Continuum  To replicate the results from the two type case in the case with a type continuum, the additional assumption that aggregate production $F$ is quasisupermodular in $\phi$ is needed. In particular, we can order technologies according to their skill bias; that is, we say that technology $\phi$ is more skill-biased than $\tilde{\phi}$, and write $\phi \succeq \tilde{\phi}$, if and only if all skill premia are higher at $\phi$ than at $\tilde{\phi}$ at any labor input $l$. This order is clearly only partial, as in general there may exist technologies $\phi$ and $\tilde{\phi}$ one of which induces higher skill premia at the top of the skill distribution while the other induces higher skill premia at the bottom. Quasisupermodularity now says the following. Take any two technologies $\phi$ and $\tilde{\phi}$ and denote by $\phi$ and $\tilde{\phi}$ their supremum and infimum, respectively. Then, if moving from $\phi$ to $\phi$ induces an output gain at a given $l$, $F(l, \phi) \geq F(l, \tilde{\phi})$, then moving from $\phi$ to $\tilde{\phi}$ must induce an output gain at $l$ as well, $F(l, \tilde{\phi}) \geq F(l, \tilde{\phi})$.

If $\phi$ and $\tilde{\phi}$ can be ordered according to their skill bias, the requirement of quasisupermodularity is trivially satisfied, as supremum and infimum are identical to the technologies themselves. If $\phi$ and $\tilde{\phi}$ cannot be ordered, quasisupermodularity rules out a form of substitutability between skill-biased technical changes at different parts of the wage distribution. To see this, consider the case where $\phi$ induces higher skill premia at the top and $\tilde{\phi}$ induces higher skill premia at the bottom of the wage distribution. The supremum then features higher skill premia and the infimum lower skill premia everywhere. Quasisupermodularity now requires that, if raising top skill premia by moving from $\phi$ to $\phi$ (while bottom skill premia are low) generates an output gain, then raising top skill premia by moving from $\phi$ to $\tilde{\phi}$ (while bottom skill premia are high) cannot produce an output loss. Put differently, technical change that raises bottom skill premia (as the change $\phi \rightarrow \tilde{\phi}$) should not offset the output gain from technical change that raises top skill premia (as the changes $\phi \rightarrow \phi$ and $\tilde{\phi} \rightarrow \tilde{\phi}$).

Under quasisupermodularity the weak bias result of Lemma 1 extends to the continuum type case.

**Lemma 2.** Suppose $\Theta = [\underline{\theta}, \overline{\theta}]$. Assume that $(\Phi, \succeq)$ is a lattice under the skill bias order $\succeq$ (defined in the text) and that $F$ is quasisupermodular in $\phi$. Consider wages as a function of labor inputs and technology as given by equation (5) and technology as a function of labor inputs as given by equation (6). Take any labor input $l$ and let $dl$ be a change in the labor input such that the relative change $\hat{l}_\theta := dl / l$ decreases in $\theta$.

Then, the technical change induced by $dl$ raises less skilled workers’ wages relative to more skilled workers’ wages throughout the skill distribution, that is,

$$\rho_{\theta, \theta'} \hat{l}_\theta + \int_{\underline{\theta}}^{\theta} \rho_{\theta, \theta'}(l, \phi^*(l)) \hat{l}_{\theta'} d\theta' \geq \rho_{\theta, \theta'} \hat{l}_\theta + \int_{\underline{\theta}}^{\theta} \rho_{\theta, \theta'}(l, \phi^*(l)) \hat{l}_{\theta'} d\theta' \quad \text{if } \theta \leq \tilde{\theta}.$$

In words, any decrease in the relative input of more skilled versus less skilled workers over the entire skill distribution induces technical change that reduces all skill premia.

**Proof.** By Theorem 5 in Loebbing (2018). \(\Box\)

---

11Existence of supremum and infimum under the skill bias order requires $(\Phi, \succeq)$ to be a lattice, which is assumed henceforth.

12The discussion here ignores the complication that the skill-bias relation $\succeq$ may actually not be a well-defined order but rather a pre-order. See Appendix B.2 in Loebbing (2018) for how to define lattice structure and
While quasisupermodularity is a somewhat abstract concept, it is important to note that it is satisfied in the important benchmark case of CES production. This is because the CES function is agnostic regarding interdependencies between technical changes at different parts of the skill distribution: the relative amount of R&D resources directed at improving the efficiency of labor type \( \theta \) relative to \( \tilde{\theta} \) only depends on the relative labor inputs of these types, but not on the distribution of technology and labor over other parts of the skill distribution. In the absence of empirical evidence on interdependencies between technical improvements that affect different parts of the wage distribution, the agnostic position of the CES function seems a good starting point for the analysis.

### 3.2. Results for the CES Case

While it can be verified that the CES function (12) indeed satisfies quasisupermodularity, there is a more direct way of verifying the prediction of Lemma 2 for the CES case. Using the technical change elasticities from equations (15) and (16), the effect of a labor input change \( \hat{l} \) on the relative wage between types \( \theta \geq \tilde{\theta} \) is obtained as

\[
\rho_{\theta,\tilde{\theta}} \hat{l}_\theta + \int_\theta^{\tilde{\theta}} \rho_{\theta,\tilde{\theta}'} (l, \phi^*(l)) \hat{l}_{\theta'} d\theta' - \rho_{\tilde{\theta},\tilde{\theta}'} \hat{l}_{\tilde{\theta}} - \int_\theta^{\tilde{\theta}} \rho_{\tilde{\theta},\tilde{\theta}'} (l, \phi^*(l)) \hat{l}_{\theta'} d\theta' = \rho^{CES} (\hat{l}_\theta - \hat{l}_{\tilde{\theta}}).
\]

If \( \hat{l} \) represents a decrease in the relative inputs of more versus less skilled workers, the difference \( \hat{l}_\theta - \hat{l}_{\tilde{\theta}} \) is negative and the relative wage of \( \theta \) versus \( \tilde{\theta} \) falls (since \( \rho^{CES} \geq 0 \), see equation (15)), as predicted by Lemma 2.

The weak bias results immediately raise the question whether the induced technical change effects can be strong enough to outweigh the (within-technology) substitution effects, which typically work in the opposite direction of the induced technical change effects. Directed technical change theory shows that this may indeed be the case if the aggregate production function fails to be quasiconcave (cf. Theorem 3 Loebbing, 2018). The model presented above allows for quasiconvexity in aggregate production by departing from perfect competition on technology markets.\(^{13}\) For the CES case there is a sharp parametric condition under which induced technical change effects dominate substitution effects – a case called strong (relative) bias of technology in the literature (e.g. Acemoglu, 2002, 2007). In particular, the total effect of a labor supply change \( \hat{l} \) on the relative wage between types \( \theta \geq \tilde{\theta} \) is

\[
\left( \rho_{\theta,\tilde{\theta}} + \gamma_{\theta,\tilde{\theta}} \right) \hat{l}_\theta + \int_\theta^{\tilde{\theta}} \left( \rho_{\theta,\tilde{\theta}'} + \gamma_{\theta,\tilde{\theta}'} \right) (l, \phi^*(l)) \hat{l}_{\theta'} d\theta' - \left( \rho_{\tilde{\theta},\tilde{\theta}} + \gamma_{\tilde{\theta},\tilde{\theta}} \right) \hat{l}_{\tilde{\theta}} - \int_\theta^{\tilde{\theta}} \left( \rho_{\tilde{\theta},\tilde{\theta}'} + \gamma_{\tilde{\theta},\tilde{\theta}'} \right) (l, \phi^*(l)) \hat{l}_{\theta'} d\theta' = \left( \rho^{CES} + \gamma^{CES} \right) (\hat{l}_\theta - \hat{l}_{\tilde{\theta}}).
\]

Considering again a decrease in the relative input of more skilled workers, \( \hat{l}_\theta \leq \hat{l}_{\tilde{\theta}} \), the effect

---

\(^{13}\)See Appendix B.1 in Loebbing (2018) for a detailed discussion of model features that enable quasiconvexity in the aggregate production function.
on the relative wage is negative if and only if
\[ \rho^{CES} + \gamma^{CES} \geq 0. \]  

Hence, if condition (17) is satisfied, induced technical change effects dominate substitution effects and skill premia fall when the relative input of more skilled workers falls. In the terms of directed technical change theory, if condition (17) is satisfied, there is strong bias of technology.

4. Tax Reforms

Starting from a given tax \( T \) a tax reform is represented by a direction \( \tau: y \mapsto \tau(y) \) and a scaling factor \( \mu > 0 \). The reform is then given by a change in the tax function from \( T \) to \( T + \mu \tau \).

Progressive tax reforms will play a central role in the analysis. I define progressive reforms as those that raise marginal tax rates disproportionately as income increases.

**Definition 1.** Starting from tax \( T \) the tax reform \((\tau, \mu)\) is progressive if and only if

\[ \frac{\tau'(y)}{1 - T'(y)} \geq \frac{\tau'(\tilde{y})}{1 - T'(\tilde{y})} \quad \forall \ y \geq \tilde{y}. \]

The meaning of this definition is clarified by the following observation.

**Observation 1.** Take any tax \( T \) with \( T'(y) < 1 \) for all \( y \), a tax reform \((\tau, \mu)\) such that \( T'(y) + \mu \tau'(y) < 1 \) for all \( y \), and let \( \tilde{T} := T + \mu \tau \) denote the post-reform tax function. Then, the following statements are equivalent.

1. The reform \((\tau, \mu)\) is progressive according to Definition 1.
2. The rate of progressivity of the post-reform tax is higher than that of the pre-reform tax everywhere,

\[ P_{\tilde{T}}(y) \geq P_T(y) \quad \forall \ y. \]
3. The post-reform tax can be obtained by taxing post-tax income under the initial tax in a progressive way, that is, by means of a tax function with increasing marginal tax rates:

\[ R_{\tilde{T}} = r \circ R_T \]

for some concave function \( r \).

**Proof.** See Appendix B.4.

Observation 1 shows that Definition 1 is equivalent to two other natural characterizations of progressive tax reforms. First, progressive reforms according to Definition 1 are equivalent to reforms that raise the rate of progressivity everywhere. Second, progressive reforms can be obtained by augmenting the initial tax by an additional tax on post-tax income that features increasing marginal tax rates.
In the following I focus on the local effects of a reform in the direction of $\tau$, that is, the effects on economic outcomes of changing $T$ to $T + \mu \tau$ as $\mu \to 0$. Note that this does not lead to confusion with the definition of progressivity, as Definition 1 only depends on the direction $\tau$ of a reform. Moreover, I restrict attention to initial tax schedules with marginal tax rates strictly below 100% everywhere. Finally and without loss of generality, suppose that worker types are ordered according to their wages under the initial tax schedule, that is, $w_\theta \leq \hat{w}_\theta$ if $\theta \leq \hat{\theta}$ under the initial tax.

To describe the effects of tax reforms on economic outcomes formally, I write equilibrium variables as a function of the tax, that is, the equilibrium value of a variable $x$ (e.g. wages or labor inputs) under tax $T$ is denoted by $x(T)$.

### 4.1. Direction of Induced Technical Change

A first step in the analysis of the implications of directed technical change for the effects of tax reforms is to analyze the induced technical change effects of tax reforms on the wage distribution. In the notation introduced in Section 2.2 the object of interest is the relative wage change

$$\frac{1}{w_\theta} D_{\phi,\tau} w_\theta(\phi^*(T)),$$

cau sed by the technical change induced by the reform $\tau$.

**Two Types** Consider first the case with two labor types. The induced technical change effect on wages can be written as

$$\frac{1}{w_\theta} D_{\phi,\tau} w_\theta(\phi^*(T)) = \sum_{\hat{\theta} \in \{\theta, \hat{\theta}\}} \rho_{\theta,\hat{\theta}} \hat{l}_{\theta,\tau}(T),$$

where

$$\hat{l}_{\theta,\tau}(T) := \frac{1}{l_{\theta}} D_{\tau} l_{\theta}(T)$$

is the relative change in the labor input of type $\hat{\theta}$ in response to the tax reform. This relative labor input in turn must satisfy

$$\hat{l}_{\theta,\tau}(T) = -\epsilon R^R \tau' \frac{\hat{g}_{\theta}(T)}{1 - T'(y_\theta(T))} + e_{\theta} \sum_{\hat{\theta} \in \{\theta, \hat{\theta}\}} \gamma_{\theta,\hat{\theta}} + \rho_{\theta,\hat{\theta}} \hat{l}_{\theta,\tau}.$$ 

The labor input change is the solution to a fixed point equation because in equilibrium labor supply changes trigger wage adjustments, which feed back to labor supply. Solving for the fixed point and using the result in equation (18) yields an intuitive condition for a tax reform to induce equalizing technical change.

---

14All elasticities in this section are evaluated at the equilibrium under the initial tax $T$. I do not write this dependence explicitly to save on notation.
Proposition 1. Let $\Theta = \{\theta, \bar{\theta}\}$, fix an initial tax $T$, and suppose that \[ \sum_{\theta \in \{\theta, \bar{\theta}\}} \epsilon^{R}_\theta (\gamma_{\theta,\theta} + \rho_{\theta,\theta}) < 1. \] (19)

Then, any reform $\tau$ with
\[ \epsilon^{R}_\theta \tau'(y_\theta(T)) \frac{\tau'(y_\theta(T))}{1 - T'(y_\theta(T))} \leq \epsilon^{R}_\bar{\theta} \tau'(y_{\bar{\theta}}(T)) \frac{\tau'(y_{\bar{\theta}}(T))}{1 - T'(y_{\bar{\theta}}(T))} \]
induces technical change that reduces the skill premium,
\[ \frac{1}{w_\theta} D_{\phi,\tau} w_\theta(T, \phi^*(T)) \geq \frac{1}{w_{\bar{\theta}}} D_{\phi,\tau} w_{\bar{\theta}}(T, \phi^*(T)). \]

If the labor supply elasticity $\epsilon^{R}_\theta$ is constant across types (e.g. because the disutility of labor is iso-elastic and $T$ is CRP), any progressive tax reform induces technical change that reduces the skill premium.

Proof. See Appendix B.5.

Proposition 1 confirms the intuition developed in the introduction. A progressive tax reform raises incentives to work for low-skilled relative to high-skilled workers, which then leads firms to demand technologies that are more complementary to low-skilled workers. This change in production technologies raises low-skilled workers’ relative productivity and hence their relative wage.

Type Continuum In the case of a type continuum the induced technical change effect becomes
\[ \frac{1}{w_\theta} D_{\phi,\tau} w_\theta(T, \phi^*(T)) = \rho_{\theta,\theta} \hat{l}_{\theta,\tau}(T) + \int_{\bar{\theta}}^{\theta} \rho_{\theta,\theta} \hat{l}_{\theta,\tau}(T) d\bar{\theta}. \] (20)

As in the two type case, the labor input changes $\hat{l}_{\theta,\tau}(T)$ solve a fixed point equation:
\[ \hat{l}_{\theta,\tau}(T) = -\epsilon^{R}_\theta \frac{\tau'(y_\theta(T))}{1 - T'(y_\theta(T))} + \epsilon^{w}_\theta (\gamma_{\theta,\theta} + \rho_{\theta,\theta}) \hat{l}_{\theta,\tau}(T) + \epsilon^{w}_\theta \int_{\bar{\theta}}^{\theta} (\gamma_{\theta,\theta} + \rho_{\theta,\theta}) \hat{l}_{\theta,\tau}(T) d\bar{\theta}. \]

Solving for the fixed point and inserting it into equation (20) yields a revealing expression for the induced technical change effects of reform $\tau$, consisting of three terms with intuitive interpretation, two of which can be signed unambiguously.

Proposition 2. Let $\Theta = \{\theta, \bar{\theta}\}$, fix an initial tax $T$, and suppose that
\[ \int_{\theta}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} (\epsilon^{w}_\theta \gamma_{\theta,\theta})^2 d\bar{\theta} d\theta < 1 \quad \text{and} \quad \int_{\theta}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} (\epsilon^{w}_\theta \rho_{\theta,\theta})^2 d\bar{\theta} d\theta < 1. \] (21)

15This condition requires that strong bias of technology is not too strong: the total own-wage elasticity $\gamma_{\theta,\theta} + \rho_{\theta,\theta}$, scaled by the labor supply elasticity with respect to the wage and summed across types, must not exceed one. Otherwise, the equilibrium is unstable, in the sense that an increase in labor supply of type $\theta$ triggers a wage increase for $\theta$ that is more than sufficient to justify the initial labor supply increase.

16These conditions are the high-dimensional counterpart to condition (19) from the two type case. Here, the condition is split into two parts to guarantee that each of the three components of the induced technical change effects exists, and not only their sum. As in the two type case, the conditions essentially constrain the feedback effects from wages on labor supply to prevent instability of equilibrium as described in footnote 15.
Suppose that the elasticities and 
See Appendix B.6. Proof. 
induced technical change effects on the basis of empirical estimates of the relevant quantities 
Note first that Proposition 2 provides a formula for the induced technical change effect of tax 
reforms that only depends on empirically observable quantities. I use this to quantify the 
Then, the induced technical change effect of tax reform $\tau$ on wages can be written as 
\[
\frac{1}{w_{\theta}}D_{\phi,\tau w_{\theta}}(T, \phi^*(T)) = \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} + \int_{\theta}^\pi \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} \\
:= D_{E_{\theta,\tau}(T)} \\
+ \rho_{\theta,\theta} \varepsilon_{\theta}^w \int_{\theta}^\pi \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} + \int_{\theta}^\pi \rho_{\theta,\theta} \varepsilon_{\theta}^w \int_{\theta}^\pi \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} d\hat{\theta} \\
:= T_{E_{\theta,\tau}(T)} \\
+ \rho_{\theta,\theta} \varepsilon_{\theta}^w \int_{\theta}^\pi \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} + \int_{\theta}^\pi \rho_{\theta,\theta} \varepsilon_{\theta}^w \int_{\theta}^\pi \rho_{\theta,\theta}(-\varepsilon_{\theta}^w) \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} d\hat{\theta}, \\
:= S_{E_{\theta,\tau}(T)}
\]
where 
\[
\bar{\rho}_{\theta,\theta} = \sum_{n=1}^{\infty} \rho_{\theta,\theta}^{(n)} \\
\rho_{\theta,\theta}^{(1)} = \rho_{\theta,\theta} \\
\rho_{\theta,\theta}^{(n)} = \int_{\theta}^\pi \varepsilon_{\theta}^w \rho_{\theta,\theta} d\hat{\theta} \quad \forall n > 1
\]
and 
\[
\bar{\gamma}_{\theta,\theta} = \sum_{n=1}^{\infty} \gamma_{\theta,\theta}^{(n)} \\
\gamma_{\theta,\theta}^{(1)} = \gamma_{\theta,\theta} \\
\gamma_{\theta,\theta}^{(n)} = \int_{\theta}^\pi \gamma_{\theta,\theta}^{(n-1)} \varepsilon_{\theta}^w (\gamma_{\theta,\theta} + \rho_{\theta,\theta}) d\hat{\theta} + \int_{\theta}^\pi \rho_{\theta,\theta}^{(n-1)} \varepsilon_{\theta}^w \gamma_{\theta,\theta} d\hat{\theta} \quad \forall n > 1.
\]
Suppose that the elasticities $\varepsilon_{\theta}^w$ and $\varepsilon_{\theta}^g$ are constant across types (e.g. because the disutility of labor is iso-elastic and $T$ is CRP). Then, $DE_{\theta,\tau}(T)$ and $TE_{\theta,\tau}(T)$ are decreasing in $\theta$ for any progressive reform $\tau$.

Proof. See Appendix B.6.

Note first that Proposition 2 provides a formula for the induced technical change effect of tax reforms that only depends on empirically observable quantities. I use this to quantify the induced technical change effects on the basis of empirical estimates of the relevant quantities in Section 6.

Qualitatively, Proposition 2 provides a decomposition of the induced technical change effects of reform $\tau$ into three terms.

1. $DE_{\theta,\tau}(T)$ measures the direct effect of the tax reform on technical change, ignoring the feedback effects on labor supply via changes in wages. This direct effect compresses the wage distribution after a progressive reform (if labor supply elasticities are homogeneous across types). The intuition is the same as in the two type case. The direct
The effect is confounded by the feedback effects from wage adjustments to labor supply and technology. These feedback effects are collected in the two remaining terms.

2. The first of them, $TE_{\theta,\tilde{\theta}}(T)$, isolates the feedback effects that are transmitted via induced technical change. The first summand of the term collects the following chain of effects:

(i) Labor supply of types $\tilde{\theta}$ changes by

$$(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(ii) This induces technical change which changes the wage of type $\theta$ by

$$\bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(iii) This affects labor supply of type $\theta$ by

$$\bar{e}_{\theta} \bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(iv) This again induces technical change which affects the wage of type $\theta$ by

$$\rho_{\theta,\tilde{\theta}} \bar{e}_{\theta} \bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

The second summand captures a similar chain of effects, only that the final effect on the wage of type $\theta$ is transmitted via technical change induced by other types’ labor supply responses rather than by $\theta$’s own labor supply response:

(i) Labor supply of types $\tilde{\theta}$ changes by

$$(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(ii) This induces technical change which changes the wage of types $\tilde{\theta}$ by

$$\bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(iii) This affects labor supply of types $\tilde{\theta}$ by

$$\bar{e}_{\theta} \bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$  

(iv) This again induces technical change which affects the wage of type $\theta$ by

$$\rho_{\theta,\tilde{\theta}} \bar{e}_{\theta} \bar{p}_{\theta,\tilde{\theta}}(-\varepsilon_{\tilde{\theta}}^R) \frac{\tau'(y_{\tilde{\theta}}(T))}{1 - T'(y_{\tilde{\theta}}(T))}.$$
In these expressions, the technical change elasticities $\rho_{\theta, \tilde{\theta}}$ captures all successive rounds of feedback effects (cf. Sachs et al., 2017) from labor supply of type $\tilde{\theta}$ via induced technical change on the wage of type $\theta$. In particular, the first-round effect $\rho_{\theta, \tilde{\theta}}^{(1)}$ is simply the direct technical change effect of $l_{\tilde{\theta}}$ on $w_{\theta}$ given by $\rho_{\theta, \tilde{\theta}}$. The second-round effect $\rho_{\theta, \tilde{\theta}}^{(2)}$ captures the following chain of effects:

(i) Labor supply of type $\tilde{\theta}$ changes by 1 percent.
(ii) This induces technical change that affects wages of all other types $\hat{\theta}$ by $\rho_{\hat{\theta}, \tilde{\theta}}$.
(iii) Type $\hat{\theta}$’s labor supply responds by $\varepsilon_{\hat{\theta}}w_{\hat{\theta}}\rho_{\hat{\theta}, \tilde{\theta}}$.
(iv) This affects the wage of type $\theta$ via first-round effects by $\rho_{\theta, \hat{\theta}}w_{\theta}\rho_{\hat{\theta}, \tilde{\theta}}$.

Third-round effects add yet another layer of induced technical change feedback and this continues ad infinitum.

Since the induced technical change effects follow a clear pattern, as shown by Lemma 2, the slope of the term $TE_{\theta, \tau}(T)$ can be signed for progressive tax reforms. In particular, if labor supply elasticities are constant across types, the feedback effects transmitted by induced technical change amplify the direct effect $DE_{\theta, \tau}(T)$, compressing the wage distribution further. This can be understood as follows. A progressive tax reform causes a direct decrease in the relative supply of skilled workers. This induces technical change that compresses skill premia (by Lemma 2). This technical change in turn leads to a further reduction in the relative supply of skilled labor, as relative wages of skilled workers decrease. This second-round decline in relative skill supply induces a further reduction in skill premia via induced technical change and this process continues until convergence.

3. The third term, $SE_{\theta, \tau}(T)$, has a similar structure to $TE_{\theta, \tau}(T)$. It captures all the feedback effects that contain at least one round of feedback transmitted via within-technology substitution effects. Since in general we have imposed too little structure on aggregate production to identify a clear pattern in the substitution effects, the effect of $SE_{\theta, \tau}(T)$ on the wage distribution cannot be signed.

To sign the total induced technical change effect of progressive tax reforms on the wage distribution, additional structure on substitution effects is needed. A radical way of imposing such structure is to assume that aggregate production is linear in $l$, such that there are no within-technology substitution effects, $\gamma_{\theta, \tilde{\theta}} = 0$ for all $\theta, \tilde{\theta}$. Then, the only remaining effects are the direct effect and the feedback effects via induced technical change, both of which compress the wage distribution.

**Corollary 1.** Suppose the conditions of Proposition 2 are satisfied and aggregate production $F$ is linear in $l$ such that $\gamma_{\theta, \tilde{\theta}} = 0$ for all $\theta, \tilde{\theta}$. Then, $SE_{\theta, \tau}(T) = 0$.

Moreover, if $\varepsilon_{\theta}^W$ and $\varepsilon_{\theta}^R$ are constant across types (e.g. because the disutility of labor is iso-elastic and $T$ is CRP), then any progressive tax reform induces technical change that reduces all skill premia.

A similarly clear pattern emerges when we assume, somewhat more generally, that aggregate production takes the CES form introduced in Section 2.3. In that case the cross-wage
elasticities $\gamma_{\theta,\bar{\theta}}$ and $\rho_{\theta,\bar{\theta}}$ are independent of $\theta$, such that feedback effects affect all wages proportionately (see the proof of Corollary 2 for details). Therefore, the only one of the terms in Proposition 2 that varies with types is the direct effect $DE_{\theta,\tau}(T)$. Since this effect reduces skill premia following progressive reforms, the same holds for the total effect in the CES case.

**Corollary 2.** Suppose the conditions of Proposition 2 are satisfied and aggregate production $F$ takes the CES form introduced in Section 2.3. Then,

$$SE_{\theta,\tau}(T) = TE_{\theta,\tau}(T) = 0$$

and

$$\frac{1}{w_\theta}D_{\phi,\tau}w_\theta(T,\phi^*(T)) = DE_{\theta,\tau}(T).$$

Hence, any reform $\tau$ with

$$R_{\theta,\tau}(y_\theta(T)) \frac{\tau'(y_\theta(T))}{1 - T'(y_\theta(T))}$$

increasing in $\theta$

induces technical change that reduces all skill premia.

If $R_{\theta,\tau}$ is constant across types (e.g. because the disutility of labor is iso-elastic and $T$ is CRP), any progressive tax reform induces technical change that reduces all skill premia.

**Proof.** See Appendix B.7.

Hence, the additional structure imposed on substitution elasticities by the CES function allows to recover the intuitive result from the two type case in an environment with a continuous earnings distribution.

### 4.2. Welfare Implications

Given that under certain conditions progressive tax reforms induce equalizing technical change it is natural to suspect that taking into account the induced technical change effect of tax reforms should raise the expected welfare gains from progressive reforms. This conjecture is examined in the following.

Welfare is measured by a Bergson-Samuelson welfare function $V : \{u_\theta\}_{\theta \in \Theta} \mapsto V(\{u_\theta\}_{\theta \in \Theta})$ that is strictly increasing in all arguments. The marginal welfare weight of an individual worker of type $\theta$ is obtained as

$$g_\theta(\{u_\theta\}_{\theta \in \Theta}) = \frac{1}{h_\theta}D_{u_\theta}V(\{u_\theta\}_{\theta \in \Theta}),$$

where the derivative $D_{u_\theta}$ is defined analogously to the definition of $D_l$ in Section 2.2.

I assume that $V$ is constructed in way that average welfare weight is normalized to one everywhere. More substantially, I assume that the welfare function values equity across workers in the following sense.

**Assumption 1.** For any utility profile $\{u_\theta\}_{\theta \in \Theta}$ such that $u_\theta$ increases in $\theta$, the marginal welfare weights $g_\theta(\{u_\theta\}_{\theta \in \Theta})$ decrease in $\theta$. 


The assumption ensures that redistributing consumption from those with high utility to those with low utility improves welfare.\textsuperscript{17}

To analyze the welfare effects of a tax reform $\tau$, write welfare as a function of the tax system:

$$W(T) := V(\{u_\theta(c_\theta(T), l_\theta(T))\}_{\theta \in \Theta}).$$

**Two Types** Consider first the case with two labor types. The welfare effect of a tax reform can then be decomposed as follows.

**Proposition 3.** Suppose $\Theta = \{\theta_1, \theta_2\}$. The welfare effect of a tax reform $\tau$ can be written as

$$D_\tau W(T) = \sum_{\theta \in \{\theta_1, \theta_2\}} (1 - g_\theta)(y_\theta(T))h_\theta + \sum_{\theta \in \{\theta_1, \theta_2\}} T'(y_\theta(T))y_\theta(T)(-e^R_\theta)\frac{\tau'(y_\theta(T))}{1 - T'(y_\theta(T))}$$

$$= ME_\tau(T) + \sum_{\theta \in \{\theta_1, \theta_2\}} g_\theta(1 - T'(y_\theta(T))) + T'(y_\theta(T))(1 + e^w_\theta) \frac{1}{w_\theta(T)} D_{\phi, \tau} w_\theta(T, \phi^*(T))$$

$$= BE_\tau(T) + \sum_{\theta \in \{\theta_1, \theta_2\}} g_\theta(1 - T'(y_\theta(T))) + T'(y_\theta(T))(1 + e^w_\theta) \frac{1}{w_\theta(T)} D_{\tau} w_\theta(T, \phi^*(T)).$$

Suppose further that $e^R_\theta$ and $e^w_\theta$ are constant across types (e.g. because the disutility of labor is iso-elastic and $T$ is CRP). Then, if $T$ is regressive (i.e., $T'(y)$ is decreasing in $y$),

$$TE^{W}_\tau \geq 0 \quad \text{and} \quad SE^{W}_\tau \leq 0.$$

**Proof.** See Appendix B.8. \qed

Proposition 3 shows that a tax has four distinct effects on welfare. The mechanical effect $ME_\tau(T)$ captures the effect from changing taxes and redistributing revenue in the absence of any behavioral responses. The behavioral effect $BE_\tau(T)$ captures the effect of the partial equilibrium response of labor supply, holding wages constant. Both effects are well known in the literature.

The third term $TE^{W}_\tau(T)$ represents the welfare implications of the technical change induced by the tax reform. The first part,

$$\sum_{\theta \in \{\theta_1, \theta_2\}} g_\theta(1 - T'(y_\theta(T)))y_\theta(T) \frac{1}{w_\theta} D_{\phi, \tau} w_\theta(T, \phi^*(T)),$$

captures the direct effect of the technology-induced wage changes on workers’ utility: from the change in pre-tax income, only the share $1 - T'(y_\theta(T))$ translates directly into a change of

\textsuperscript{17}Note that, since preferences are identical across workers, the interpersonal comparison of utilities inherent in the assumption is not problematic.
utility as the remaining share is taxed away. The second part,

\[ \sum_{\theta \in \Theta(T)} T'(y_\theta(T))(1 + \epsilon_\theta^w)y_\theta(T) \frac{1}{w_\theta} D_{\phi,\tau}w_\theta(T, \phi^s(T)), \]

is the welfare effect of the lump-sum redistribution of the revenue gain or loss induced by the wage adjustments to technical change. Here, the pre-tax income change is scaled by \( 1 + \epsilon_\theta^w \), as the wage change induces a labor supply adjustment of \( \epsilon_\theta^w \).\(^{18}\)

Importantly and somewhat surprisingly, even if the induced technical change reduces the skill premium (e.g., because \( \tau \) is progressive and the conditions of Corollary 2 are satisfied), we cannot sign the induced technical change effects on welfare unambiguously. This is because, when starting from a progressive tax \( T \), the reduction in high-skilled workers’ wages passes through to the government budget to a larger extent than the simultaneous rise in low-skilled workers’ wages, as marginal tax rates are higher for the high-skilled.\(^{19}\) The resulting negative revenue effect then potentially offsets the welfare gain from the shift in wages from high- to low-skilled workers. If, in contrast, the initial tax is regressive, the revenue effect becomes positive and works in the same direction as the direct utility effect of the wage changes. Hence, if the conditions of Corollary 2 are satisfied and the initial tax is regressive, the induced technical change effects of a progressive reform raise welfare.

The final term in Proposition 3, \( SE^W_\tau(T) \) captures the welfare effect of the within-technology substitution effects on wages caused by the tax reform. Its structure is analogous to that of \( TE^W_\tau(T) \), but its sign is typically opposite, because substitution effects work in the opposite direction of induced technical change effects.

**Type Continuum** The decomposition of the welfare effects of tax reforms from the two type case can recovered for a continuum of types as follows.

**Proposition 4.** Suppose \( \Theta = [\underline{\theta}, \overline{\theta}] \). The welfare effect of a tax reform \( \tau \) can be written as

\[
D_{\tau}W(T) = \int_{\underline{\theta}}^{\overline{\theta}} (1 - g_\theta)\tau(y_\theta(T))h_\theta d\theta + \int_{\underline{\theta}}^{\overline{\theta}} T'(y_\theta(T))y_\theta(T)(-\epsilon_\theta^w) \frac{\tau'(y_\theta(T))}{1 - T'(y_\theta(T))}h_\theta d\theta \\
+ \int_{\underline{\theta}}^{\overline{\theta}} [g_\theta(1 - T'(y_\theta(T))) + T'(y_\theta(T))(1 + \epsilon_\theta^w)] y_\theta(T) \frac{1}{w_\theta(T)} D_{\phi,\tau}w_\theta(T, \phi^s(T))h_\theta d\theta \\
+ \int_{\underline{\theta}}^{\overline{\theta}} [g_\theta(1 - T'(y_\theta(T))) + T'(y_\theta(T))(1 + \epsilon_\theta^w)] y_\theta(T) \frac{1}{w_\theta(T)} D_{\tau}w_\theta(T, \phi^s(T))h_\theta d\theta.
\]

**Proof.** See Appendix B.9.
Proposition 4 decomposes the welfare effect of tax reforms into exactly the same terms as in the two type case. The structure of the separate terms is the same as for two types as well. Moreover, as in the two type case, the effects cannot be signed unambiguously because of the potentially diverging effects of wage changes on worker utility (directly) and on government revenue (and hence indirectly on worker utility).

Importantly, however, Proposition 4 can be combined with the expressions for the wage effects \( D_{\phi,\tau}w_\theta(T, \phi^*(T)) \) and \( D_{\tau}w_\theta(T, \phi^*(T)) \) in Proposition 2 and Lemma 3 (in Appendix A.1). This yields a formula for the welfare effects of tax reforms in terms of empirically observable quantities and welfare weights. I use this combination of expressions to quantify the welfare effects of tax reforms and the contribution of induced technical change in Section 6.

For a sharp analytical characterization of the implications of induced technical change for the welfare gains from progressive tax reforms, a different type of comparison is required. Instead of comparing the welfare effects of a given progressive tax reform, comparing the scope for welfare-improvements by means of progressive tax reforms turns out to be insightful.

To obtain clear results, restrict attention to those initial tax schedules for which Corollary 2 gives a clear characterization of the technical change induced by progressive reforms; that is, restrict attention to initial tax schedules that are CRP. Denote by

\[
T := \{ T \mid T \text{ is CRP, } \exists \tau \text{ progressive s.t. } D_\tau W(T) > 0 \}
\]

the set of initial tax schedules (within the CRP class) that can be improved in terms of welfare by means of a progressive reform.

To obtain a benchmark for comparison that does not include the induced technical change effects of tax reforms, let

\[
D_\tau W^{ex}(T) := D_\tau W(T)\big|_{\rho_{ij}=0 \forall \theta, \bar{\theta}}
\]

denote the welfare effect of reform \( \tau \) when counterfactually setting all technical change elasticities to zero. Then, we can define

\[
T^{ex} := \{ T \mid T \text{ is CRP, } \exists \tau \text{ progressive s.t. } D_\tau W^{ex}(T) > 0 \}
\]

as the set of CRP schedules that one would perceive to be improvable by progressive reforms if one were to ignore the induced technical change effects.

The two sets of “progressively improvable” tax functions, \( T \) and \( T^{ex} \) can now be ordered according to set inclusion.

**Proposition 5.** Suppose \( \Theta = [\underline{\theta}, \overline{\theta}] \), \( F \) is CES as introduced in Section 2.3, and the disutility of labor is iso-elastic. Then,

\[
T^{ex} \subseteq T,
\]

that is, the set of initial tax schedules that can be improved by progressive reforms becomes larger when taking into account the induced technical change effects.

**Proof.** See Appendix B.10. \( \square \)

Proposition 5 shows that the scope for welfare improvements by means of progressive tax
reforms is widened when taking into account the induced technical change effects of tax reforms. This aligns neatly with the result from Section 4.1, saying that progressive tax reforms induce equalizing technical change, and with the intuition developed in the introduction.

The idea behind Proposition 5 relies on the mechanism design approach to taxation analysis. It goes as follows. Consider a progressive tax reform which a tax planner who neglects induced technical change effects (the exogenous technology planner, henceforth) expects to raise welfare. For any such reform, a planner who correctly anticipates induced technical change effects (the endogenous technology planner, henceforth) can find another progressive reform that exactly replicates the labor allocation expected by the exogenous technology planner following his reform. But since progressive tax reforms induce equalizing technical change, the endogenous technology planner correctly anticipates that the incentive compatibility constraints across workers allow for more redistribution towards low-skilled workers than expected by the exogenous technology planner. The endogenous technology planner can thus replicate the exogenous technology planner’s expected labor allocation and seize an additional welfare gain through a more equal consumption allocation. So, the endogenous technology planner can find a progressive reform with a higher welfare gain that that expected by the exogenous technology planner from his reform. Since this reasoning holds for any progressive reform of the exogenous technology planner, the endogenous technology planner can find a welfare-improving progressive reform whenever the exogenous technology planner can find one. Hence, the endogenous technology planner perceives the scope for welfare-improvements through progressive tax reforms to be greater.\footnote{The proof of Proposition 5 formalizes this reasoning.}

5. Optimal Taxes

The results from the tax reform analysis suggest that induced technical change should render the optimal tax scheme more progressive. I examine this conjecture in the following.

To derive optimal tax rates, I follow the mechanism design approach to optimal taxation as it provides the fastest route to the central results presented below.\footnote{The alternative approach would be to use the formulas for the welfare effects of tax reforms from Section 4.2 and impose that these effects are zero for all reforms at the optimum. The two approaches yield the same results.} For that, write welfare as a function of consumption and labor allocations instead of utility levels:

\[
\tilde{W}(c, l) := V \left( \{ u_{\theta}(c_{\theta}, l_{\theta}) \}_{\theta \in \Theta} \right).
\]

The goal is to find the consumption-labor allocation that maximizes welfare \( W(c, l) \) subject to the aggregate resource constraint and to incentive compatibility constraints across worker types. The optimal tax schedule is then obtained as the tax that implements the welfare-maximizing allocation.

For the characterization of optimal tax rates it is convenient to denote the average welfare

\[
\overline{W}(c, l) := V \left( \{ u_{\theta}(c_{\theta}, l_{\theta}) \}_{\theta \in \Theta} \right).
\]
weight across all types above a given type $\theta$ as

$$\tilde{g}_{\theta} := \frac{1}{1 - F_{\theta}} \int_{\tilde{\theta}}^{\theta} g_{\tilde{\theta}} h_{\tilde{\theta}} d\tilde{\theta}. $$

Moreover, for a function $x : (\theta, z) \mapsto x_{\theta}(z)$ (e.g., wages or labor inputs) that depends on $\theta$ and potentially further variables $z$, I denote the derivative of $x$ with respect to $\theta$ by

$$x'_{\theta}(z) := \frac{dx_{\theta}(z)}{d\theta}$$

and the corresponding semi-elasticity by

$$\hat{x}_{\theta}(z) := \frac{x'_{\theta}(z)}{x_{\theta}(z)}.$$

I assume throughout the analysis that worker types are ordered according to their wages under the optimum tax schedule, that is, $w_{\theta} \leq w_{\tilde{\theta}}$ if $\theta \leq \tilde{\theta}$ under the optimal tax.

### 5.1. Two Types

Again I start with the analysis of the two type case. Here the aggregate resource constraint is given by

$$\sum_{\theta \in \{\theta, \tilde{\theta}\}} c_{\theta} h_{\theta} = F(l, \phi^*(l))$$

Assumption 1 (the welfare function values equity) guarantees that the low-skilled workers’ incentive compatibility constraint is slack in the optimum, so we can restrict attention to incentive compatibility for high-skilled workers. The incentive compatibility constraint for high-skilled workers will be binding:

$$c_{\tilde{\theta}} - v(l_{\tilde{\theta}}) = c_{\tilde{\theta}} - v \left( \frac{w_{\tilde{\theta}}}{w_{\tilde{\theta}} l_{\tilde{\theta}}} \right).$$

From the incentive compatibility and resource constraints, consumption levels can be derived as a function of labor inputs. I substitute this function into the welfare function $\tilde{W}$ and take first-order conditions with respect to labor inputs. Using workers’ first-order condition to reintroduce marginal tax rates into the problem then yields the following expression for optimal marginal tax rates (see the proof of Proposition 6 for details).
Proposition 6. Suppose $\Theta = \{\theta, \bar{\theta}\}$. Then the optimal tax $T$ satisfies the following conditions.

\[
T'(y_{\theta}) = \left( v'(l_{\theta}) - v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{w_{\theta}}{w_{\bar{\theta}}} \frac{h_{\bar{\theta}}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) \right) + v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{h_{\bar{\theta}}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\gamma_{\theta, \theta} - \gamma_{\theta, \bar{\theta}}) \]
\[
= PE_{\theta}^* \geq 0
\]

\[
T'(y_{\bar{\theta}}) = \left( v'(l_{\bar{\theta}}) - v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{w_{\theta}}{w_{\bar{\theta}}} \frac{h_{\bar{\theta}}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\rho_{\bar{\theta}, \theta} - \rho_{\bar{\theta}, \bar{\theta}}) \right) \]
\[
= SE_{\bar{\theta}}^* \geq 0
\]

\[
T'(y_{\theta}) = \left( v'(l_{\theta}) - v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{w_{\theta}}{w_{\bar{\theta}}} \frac{h_{\theta}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\gamma_{\theta, \theta} - \gamma_{\theta, \bar{\theta}}) \right) + v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{h_{\theta}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\rho_{\theta, \theta} - \rho_{\theta, \bar{\theta}}) \]
\[
= TE_{\theta}^* \leq 0
\]

\[
T'(y_{\bar{\theta}}) = \left( v'(l_{\bar{\theta}}) - v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{l_{\bar{\theta}}}{l_{\theta}} w_{\theta} \frac{h_{\theta}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\rho_{\bar{\theta}, \theta} - \rho_{\bar{\theta}, \bar{\theta}}) \right) \]
\[
= SE_{\theta}^* \leq 0
\]

\[
T'(y_{\theta}) = \left( v'(l_{\theta}) - v' \left( \frac{w_{\theta}}{w_{\bar{\theta}}} \right) \frac{l_{\theta}}{l_{\theta}} w_{\theta} \frac{h_{\theta}}{h_{\theta} w_{\theta}} (1 - g_{\theta}) (\rho_{\theta, \theta} - \rho_{\theta, \bar{\theta}}) \right) \]
\[
= TE_{\theta}^* \geq 0
\]

where all variables are evaluated at equilibrium under the optimal tax $T$.

Proof. See Appendix B.11. □

Proposition 6 provides conditions that decompose the optimal tax rates into three terms. The first term $PE_{\theta}^*$ is the standard expression that would be obtained in a setting with exogenous wages. It is positive for the marginal tax rate of low-skilled workers and negative for high-skilled workers, reflecting the well-known result that the optimal marginal tax rate for the highest income earners is zero when wages are exogenous.

The second term, $SE_{\theta}^*$, captures the impact of substitution effects between workers’ wages on the optimal tax rates. As discussed by Stiglitz (1987), this term is positive for low-skilled workers and negative for high-skilled workers, pushing for a negative marginal tax rate at the top. This is because substitution effects imply that an increase in high-skilled workers’ labor supply reduces the skill premium, which relaxes high-skilled workers’ incentive compatibility constraint. The optimal tax schedule capitalizes on this effect by distorting high-skilled workers’ labor supply upwards (imposing a negative marginal tax rate on them) and using the resulting slack in the incentive compatibility constraint for redistribution towards low-skilled workers.

The third term, $TE_{\theta}^*$, is novel. It captures the induced technical change effects. Since the induced technical change effects counteract substitution effects, the sign of $TE_{\theta}^*$ is opposite to $SE_{\theta}^*$: it pushes for lower marginal taxes at the bottom and higher marginal taxes at the top. The intuition is closely related to the intuition for the positive influence of induced technical change effects on the scope for welfare improvements through progressive tax reforms described in Section 4.2. By raising marginal tax rates at the top and lowering them at the bottom, the optimal tax schedule reduces the relative supply of skilled labor, thus inducing firms to operate technologies with a higher relative productivity for low-skilled workers. This
reduces the skill premium, slackens high-skilled workers’ incentive compatibility constraint, and widens the scope for redistribution. The combined impact of induced technical change and substitution effects depends on the condition for strong bias. In particular, if

\[ \gamma_{\theta} + \rho_{\theta} - \gamma_{\theta} > 0, \]

the induced technical change effects dominate and the skill premium increases in the supply of skilled labor (i.e., there is strong bias). In this case, the marginal tax rate at the top becomes positive. The optimal tax scheme distorts high-skilled workers’ labor supply downwards to reduce the skill premium via induced technical change effects. In the opposite case where substitution effects dominate, the marginal tax rate at the top is negative, following the logic of Stiglitz (1987).

While the results from the two type case are simple and instructive regarding the new effects introduced by directed technical change, it is well-known that many insights from the two type case are of little practical relevance. This is because they focus on workers who are either the highest or the lowest income earners in the economy, while almost all workers in reality are somewhere in between. The prescriptions regarding optimal marginal tax rates can differ strongly between such interior workers and those at either end of the earnings distribution. Hence, I proceed by analyzing optimal taxes in the setting with a continuous earnings distribution.

5.2. Type Continuum

With a continuum of labor types the resource constraint becomes

\[ \int_{\theta}^{\tilde{\theta}} c_{\theta} h_{\theta} d\theta = F(l, \phi^*(l)). \] (22)

Incentive compatibility constraints are given by

\[ u_{\theta} = \max_{\tilde{\theta} \in \Theta} \left\{ c_{\tilde{\theta}} - v \left( \frac{w_{\theta} l_{\theta}}{w_{\tilde{\theta}}} \right) \right\} \forall \theta. \]

According to standard results in taxation theory, this is equivalent to the following conditions:

\[ c'_{\theta} = v'(l_{\theta}) \left( w_{\theta} l_{\theta} + w_{\theta} l'_{\theta} \right) \frac{1}{w_{\theta}} \forall \theta, \] (23)

\[ y'_{\theta} \geq 0 \quad \text{and} \quad w'_{\theta} \geq 0 \quad \forall \theta. \] (24)

As is usual in the literature, I drop the monotonicity requirement (24) and study the relaxed problem of maximizing welfare subject to (22) and (23). In all quantitative simulations of optimal taxes, I verify that the monotonicity condition (24) holds at the optimum.
serting this function into the welfare objective and optimizing over \( l \) yields a high-dimensional counterpart of the optimality conditions for two types in Proposition 6.\(^{23}\)

**Proposition 7.** Suppose \( \Theta = [\theta, \bar{\theta}] \). Then the optimal tax \( T \) satisfies the following conditions.

\[
T'(y_{\theta}) = PE_{\theta}^* + SE_{\theta}^* + TE_{\theta}^*,
\]

for all \( \theta \), where

\[
\frac{PE_{\theta}^*}{1 - PE_{\theta}} = \left( 1 + \frac{1}{e_{\theta}} \right) \frac{1 - H_{\theta}}{h_{\theta}} (1 - \tilde{g}_{\theta}) \tilde{w}_{\theta}
\]

\[
SE_{\theta}^* - \theta = \int_{\theta}^{\bar{\theta}} (1 - T'(y_{\tilde{\theta}})) y_{\tilde{\theta}} \frac{1 - H_{\tilde{\theta}}}{h_{\tilde{\theta}}w - \tilde{\theta}} (1 - \tilde{g}_{\tilde{\theta}}) D_{l_{\tilde{\theta}}} \tilde{w}_{\tilde{\theta}} (l, \phi * (l)) \, d\tilde{\theta}
\]

\[
TE_{\theta}^* = \int_{\theta}^{\bar{\theta}} (1 - T'(y_{\tilde{\theta}})) y_{\tilde{\theta}} \frac{1 - H_{\tilde{\theta}}}{h_{\tilde{\theta}}w - \tilde{\theta}} (1 - \tilde{g}_{\tilde{\theta}}) D_{\phi, l_{\tilde{\theta}}} \tilde{w}_{\tilde{\theta}} (l, \phi * (l)) \, d\tilde{\theta}
\]

and all variables are evaluated at equilibrium under the optimal tax \( T \).

Moreover, \( TE_{\theta}^* \leq 0 \) and \( TE_{\bar{\theta}}^* \geq 0 \).

**Proof.** See Appendix B.12. \( \square \)

The conditions in Proposition 7 closely resemble those from the two type case. Again, optimal marginal tax rates are decomposed into three terms, each of them with an interpretation analogous to its two type counterpart. The sign of the induced technical change effects on optimal marginal tax rates, which is the main object of interest here, is the same as for two types as well. Induced technical change effects call for an upwards adjustment of marginal tax rates at the top and for a downwards adjustment at the bottom of the income distribution. Without further assumptions, however, unambiguous results are again limited to marginal tax rates at the highest and lowest income levels, respectively. To gain insights into the impact of directed technical change on optimal marginal tax rates on interior income levels I turn to the CES case.\(^{24}\)

In the CES case, the condition for optimal marginal tax rates simplifies as follows.

**Proposition 8.** Suppose \( \Theta = [\theta, \bar{\theta}] \) and aggregate production \( F \) takes the CES form introduced in Section 2.3. Then the optimal tax \( T \) satisfies the following conditions.

\[
\frac{T'(y_{\theta})}{1 - T'(y_{\theta})} = \left( 1 + \frac{1}{e_{\theta}} \right) \frac{1 - B_{\beta_{\theta}}}{b_{\beta_{\theta}}} (1 - \tilde{g}_{\theta}) + \gamma_{CBS} (1 - g_{\theta}) + \rho_{CBS} (1 - g_{\theta})
\]

for all \( \theta \), where all variables are evaluated at equilibrium under the optimal tax \( T \), the function \( \beta : \theta \mapsto \beta_{\theta} \) is given by

\[
\beta_{\theta} := \kappa_{\theta}^{1+\gamma_{CBS}+\rho_{CBS}} h_{\theta}^{\gamma_{CBS}+\rho_{CBS}} \forall \theta
\]

\(^{23}\)This procedure adapts the approach of Diamond (1998) to an environment with endogenous wages.

\(^{24}\)Another limitation of Proposition 7 is that it does not provide a closed-form solution for optimal marginal tax rates, not even for exogenous welfare weights and iso-elastic disutility of labor. This precludes sharp analytical comparisons between optimal marginal tax rates with and without induced technical change effects. The CES case overcomes this limitation as well.
while $B$ and $b$ are the cumulative distribution and the density function of $\beta$.

Alternatively, the conditions for optimal marginal tax rates can be written as

$$T'(y) \frac{1 - T'(y)}{1 - y} = 1 - M_y y (1 - \bar{g}_y) + \gamma^{CES} (1 - g_y) + \rho^{CES} (1 - g_y),$$

(26)

where all variables are evaluated at equilibrium under the optimal tax $T$, $M$ and $m$ denote the cumulative distribution and the density function of $y$ at the optimum, and $\theta_y$ denotes the type of workers who earn income $y$ at the optimum.

Proof. See Appendix B.13. \qed

Proposition 8 provides two complementary conditions for optimal marginal tax rates. Equation (26) is a direct extension of the corresponding result in Sachs et al. (2017, p. 39). It extends their formula to include the induced technical change adjustment term $\rho^{CES} (1 - g_y)$. Equation (25) is more novel. It provides a Diamond (1998) style expression for optimal marginal tax rates, taking into account the endogeneity of wages through substitution and induced technical change effects.

Remarkably, when welfare weights are exogenous and the disutility of labor is iso-elastic (such that $e_\theta$ is exogenous), equation (25) offers a closed-form solution for optimal marginal tax rates. Hence, once the model is calibrated appropriately, optimal marginal tax rates can be simulated directly from equation (25). I do so in the quantitative assessment of optimal taxes with directed technical change in Section 6.

From an analytical point of view, the fact that equation (25) has closed form means that it can be used to precisely identify the role of directed technical change in shaping the optimal tax schedule. The alternative equation (26) will then be useful to develop intuition for the insights derived from (25).

To identify the role of directed technical change precisely, I use again the concepts of endogenous and exogenous technology planners introduced in Section 4.2. The endogenous technology planner fully understands how the economy works and computes optimal taxes according to equation (25). The exogenous technology planner knows about all fundamentals of the model but mistakenly believes that technology is exogenous. In particular, he observes the economy under a given tax $T$ and believes that technology remains fixed at its current state $\phi^*(T)$, independently of the tax schedule.25 Lemma 4 in Appendix A.2 shows that the tax $T_T^e$ perceived as optimal by the exogenous technology planner satisfies the following condition:

$$T'(y_\theta) \frac{1 - T'(y_\theta)}{1 - y_\theta} = \left(1 + \frac{1}{e_\theta}\right) \frac{1 - \bar{p}_\theta}{\bar{p}_\theta \bar{g}_\theta} (1 - \bar{g}_\theta) + \gamma^{CES} (1 - g_\theta) + \rho^{CES} (1 - g_\theta),$$

(27)

25Formally, the exogenous technology planner bases his computation of optimal taxes on the reduced form equations (2), (1), (5), and (7), but replaces the equilibrium technology equation (6) by the “wrong” equation

$$\phi^*(l) = \phi^*(l(T)) = \arg\max_{\phi \in \Phi} F(l(T), \phi) \quad \forall l.
where the function $\beta : \theta \mapsto \beta_\theta$ is given by

$$\beta_\theta := \kappa_\theta^{1+r^{CES}} h_\theta^{r^{CES}} (\phi' (T))^{1+r^{CES}} \forall \theta$$

while $B$ and $b$ are the cumulative distribution and the density function of $\beta$.

Comparing the exogenous technology planner’s tax rates (27) with those of the endogenous technology planner in equation (25), there are two differences. First, the endogenous technology planner takes into account the directed technical change adjustment $\rho CES (1 - g_\theta)$. This term is increasing in $\theta$ (as welfare weights are decreasing in $\theta$ at the optimum) and in this sense necessitates a progressive adjustment of the tax schedule. The intuition for this adjustment is the same as in the two type case. Lowering marginal tax rates at the bottom and raising them at the top induces technical change that compresses the wage distribution and hence improves equity.

The second difference is that the endogenous technology planner uses the hazard rate of $\beta$ whereas the exogenous technology planner uses that of $\bar{\beta}$. The function $\beta$ can be interpreted as the exogenous part of inequality in the model: if the supply of types of labor was identical, wages would be proportional to $\beta$. $\bar{\beta}$ instead is the exogenous technology planner’s wrong inference about the exogenous part of inequality. The exogenous technology planner believes that, if all labor types’ supply was identical, wages would be proportional to $\bar{\beta}$ instead of $\beta$. It can be shown that the hazard rate of $\beta$ is smaller than the hazard rate of $\bar{\beta}$ everywhere,

$$\frac{1 - B_{\beta_\theta}}{b_{\beta_\theta} \beta_\theta} \leq \frac{1 - B_{\bar{\beta}_\theta}}{b_{\bar{\beta}_\theta} \bar{\beta}_\theta} \forall \theta,$$

if the tax schedule $T$ under which the exogenous technology planner observes the economy has a rate of progressivity smaller than one. I assume this to be the case henceforth. The second adjustment therefore reduces marginal tax rates everywhere.

To understand this adjustment intuitively, consider equation (26). There, induced technical change effects enter the formula for marginal tax rates via the labor supply elasticity $\epsilon_{R}^{\theta}$. In particular, when labor supply of type $\theta$ falls due to an increase in the marginal tax rate, induced technical change amplifies the fall in labor supply by reducing the wage of type $\theta$. Therefore, the endogenous technology planner perceives the equilibrium elasticity $\epsilon_{R}^{\theta}$ to be higher for all types than the exogenous technology planner, which leads to downwards adjustment of the endogenous technology planner’s preferred marginal tax rates.

To summarize, both directed technical change adjustments reduce marginal tax rates in the lower part of the income distribution, whereas they work in opposite directions in the upper part. I next consider the lower and upper tail of the income distribution and show that the directed technical change adjustments can be signed unambiguously in both tails.

---

26 If $P_T (y) > 1$, the labor supply elasticity with respect to the wage, $\epsilon_{w}^{\theta}$, becomes negative, which produces some counterintuitive results. Assuming that $P_T (y) < 1$ everywhere is reasonable, as empirical tax schedules typically have rates of progressivity way below one.
Optimal Taxes in the Lower Tail  
For low income levels both directed technical change adjustments reduce marginal tax rates, so we directly arrive at the following comparison.

Corollary 3. Suppose $\Theta = [\underline{\theta}, \overline{\theta}]$, aggregate production $F$ is CES, the disutility of labor is iso-elastic, welfare weights are exogenous, and the tax $T$ satisfies $P_T(y) < 1$ for all $y$. Let $T$ denote the optimal tax schedule as characterized by Proposition 8 and $T^T_{\text{ex}}$ the tax schedule perceived as optimal by the exogenous technology planner as given by equation (27).

Then, we have
$$T'(\theta(T)) \leq T^T_{\text{ex}}(\theta(T^T_{\text{ex}})) \quad \forall \theta \leq \tilde{\theta}_1$$
for some $\tilde{\theta}_1 > \bar{\theta}$. Moreover, if $\gamma_{\text{CES}} + \rho_{\text{CES}} > 0$,
$$T'(\theta(T)) < 0 \quad \forall \theta < \tilde{\theta}_2$$
for some $\tilde{\theta}_2 > \theta$.

Proof. See Appendix B.14.

Corollary 3 establishes that accounting for directed technical change unambiguously lowers marginal tax rates for a non-degenerate interval of types at the lower end of the income distribution. If there is strong bias, that is, if $\gamma_{\text{CES}} + \rho_{\text{CES}} > 0$, directed technical change effects may even call for negative marginal tax rates in the lower tail of the income distribution. The idea behind this result is that negative marginal tax rates at the bottom stimulate the labor supply of low-skilled workers. This in turn induces firms to operate technologies with a higher complementarity with low-skilled workers, which raises their wages. In the case of strong bias, this induced technical change effect is strong enough to raise low-skilled workers’ wages above their initial level despite them being in greater supply. This then provides utility gains for the low-skilled and improves equity.

Optimal Taxes in the Upper Tail  
For high incomes the two directed technical change adjustments work in opposite directions. Yet, when assuming realistically that the upper tail of the income distribution resembles the tail of a Pareto distribution, the ambiguity resolves.

In particular, suppose that we observe the economy at a tax $\tilde{T}$ with a constant marginal top tax rate. If under this tax the hazard rate of the income distribution converges to a constant for high incomes (as is the case for a Pareto distribution), one can show that also the hazard rate of the exogenous inequality measure $\beta$ converges to a constant. This then yields a parametric expression for the optimal marginal tax rate in the upper tail.

Corollary 4. Let $\Theta = [\underline{\theta}, \overline{\theta}]$ and aggregate production be CES. Suppose at a tax $\tilde{T}$, with $\tilde{T}'(y) = \tau_{\text{top}}$ for all $y \geq \tilde{y}$ and some threshold $\tilde{y}$, the hazard rate of the income distribution satisfies
$$\lim_{\theta \to \overline{\theta}} \frac{1 - M_{\theta \tilde{y}}}{m_{\theta \tilde{y}} y_{\theta \tilde{y}}} = \frac{1}{\alpha}$$
for some $\alpha > 1$. Moreover, let the disutility of labor be iso-elastic and welfare weights satisfy
$$\lim_{\theta \to \overline{\theta}} g_{\theta} = \delta_{\text{top}}$$
at the optimal tax.

Then, the optimal tax $T$ satisfies

$$\lim_{\theta \to \bar{\theta}} \frac{T'(y_{\theta})}{1 - T'(y_{\theta})} = \frac{1 - g_{\text{top}}}{\alpha e} + \frac{\alpha - 1}{\alpha} \gamma_{\text{CES}} (1 - g_{\text{top}}) + \frac{\alpha - 1}{\alpha} \rho_{\text{CES}} (1 - g_{\text{top}}).$$

(29)

Proof. See Appendix 4. □

Complementing Corollary 4, Lemma 5 in Appendix A.2 shows that under the conditions of Corollary 4 the exogenous technology planner’s preferred tax satisfies

$$\lim_{\theta \to \bar{\theta}} \frac{T'(y_{\theta})}{1 - T'(y_{\theta})} = \frac{1 - g_{\text{top}}}{\alpha e} + \frac{\alpha - 1}{\alpha} \gamma_{\text{CES}} (1 - g_{\text{top}}).$$

(30)

This expression is smaller than the endogenous technology planner’s preferred rate (29).

Hence, taking into account directed technical change when designing optimal taxes leads to unambiguously higher marginal tax rates in the upper (Pareto) tail of the income distribution.

6. Quantitative Analysis

IN PROGRESS

7. Conclusion

I investigate the implications of directed technical change for the design of non-linear labor income taxes. First, I develop a model with directed technical change and endogenous labor supply, in which the structure of labor supply determines the direction of technical change. Tax reforms affect the direction of technical change by altering the structure of labor supply. I derive conditions under which any progressive income tax reform induces technical change that compresses the wage distribution. Relatedly, using a welfare measure that values equity across workers, I show that accounting for directed technical change unambiguously increases the set of tax schedules that can be improved in terms of welfare by means of progressive tax reforms. Finally, when directed technical change is taken into account – as opposed to treating technology as exogenous – optimal marginal tax rates are higher in the upper tail and lower in the lower tail of the income distribution. Optimal marginal tax rates in the lower tail may even become negative.

In deriving these results, I adapt the approach of Diamond (1998) in the analysis of optimal non-linear labor income taxes to environments with an endogenous wage distribution. Thereby, I derive a closed-form expression for optimal marginal tax rates when the aggregate production function takes the CES form. Moreover, I propose a new tool for a clean analytical comparison between optimal tax schedules that do and do not account for endogeneity of wages. These innovations can be useful more generally in the analysis of non-linear taxation in models of general equilibrium.
References


A. Additional Results

A.1. Substitution Effects of Tax Reforms on Wages

Proposition 2 provides a general formula for the induced technical change effects of tax reforms on wages. Here, I state the counterpart of this proposition for the within-technology substitution effects. The formulae have exactly the same structure. The only difference is that here technical change wage elasticities are replaced by substitution elasticities.

**Lemma 3.** Let $\Theta = [\underline{\theta}, \overline{\theta}]$, fix an initial tax $T$, and suppose that

$$
\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \epsilon_{\theta, \hat{\theta}}^2 d\hat{\theta} d\theta < 1 \quad \text{and} \quad \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \epsilon_{\theta, \hat{\theta}}^2 d\hat{\theta} d\theta < 1. \quad (31)
$$

Then, the within-technology substitution effect of tax reform $\tau$ on wages can be written as

$$
\frac{1}{w_{\theta} \phi_{\theta}(T, \phi^*(T))} = \gamma_{\theta, \hat{\theta}} \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} + \int_{\underline{\theta}}^{\overline{\theta}} \gamma_{\theta, \hat{\theta}} \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} \\
+ \gamma_{\theta, \hat{\theta}} \epsilon_{\theta} \rho_{\theta, \hat{\theta}} \left[ \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} + \int_{\underline{\theta}}^{\overline{\theta}} \gamma_{\theta, \hat{\theta}} \epsilon_{\theta} \rho_{\theta, \hat{\theta}} d\hat{\theta} \right] \\
+ \gamma_{\theta, \hat{\theta}} \epsilon_{\theta} \rho_{\theta, \hat{\theta}} \left[ \frac{\tau'(y_{\theta}(T))}{1 - T'(y_{\theta}(T))} d\hat{\theta} + \int_{\underline{\theta}}^{\overline{\theta}} \gamma_{\theta, \hat{\theta}} \epsilon_{\theta} \rho_{\theta, \hat{\theta}} d\hat{\theta} \right],
$$

where

$$
\overline{\rho}_{\theta, \hat{\theta}} = \sum_{n=1}^{\infty} \rho_{\theta, \hat{\theta}}^{(n)}
$$

$$
\rho_{\theta, \hat{\theta}}^{(1)} = \rho_{\theta, \hat{\theta}}
$$

$$
\rho_{\theta, \hat{\theta}}^{(n)} = \int_{\underline{\theta}}^{\overline{\theta}} \rho_{\theta, \hat{\theta}}^{(n-1)} \epsilon_{\theta} \rho_{\theta, \hat{\theta}} d\hat{\theta} \quad \forall n > 1
$$
\[
\overline{T}_{\theta, \beta} = \sum_{n=1}^{\infty} \gamma_{\theta, \beta}^{(n)}
\]
\[
\gamma_{\theta, \beta}^{(1)} = \gamma_{\theta, \beta}
\]
\[
\gamma_{\theta, \beta}^{(n)} = \int_{\theta}^{\beta} \gamma_{\theta, \beta}^{(n-1)} \epsilon_{\beta}^{\theta} (\gamma_{\theta, \beta} + \rho_{\theta, \beta}) d\theta + \int_{\theta}^{\beta} \rho_{\theta, \beta}^{(n-1)} \epsilon_{\beta}^{\theta} \gamma_{\theta, \beta} d\theta \quad \forall n > 1.
\]

**Proof.** The proof is analogous to the proof of Proposition 2 and therefore omitted. \(\square\)

**A.2. Optimal Taxes under CES Production for the Exogenous Technology Planner**

As described in the main text, the exogenous technology planner believes that the economy works according to all reduced form equations from Section 2.1 with the exception of the condition for the equilibrium technology (6). Instead of following equation (6), the exogenous technology planner believes that technology remains fixed at its equilibrium value under a given tax \(T\), \(\phi^*(l(T))\). The idea is that the planner observes the economy under the tax \(T\) when computing optimal taxes and believes technology to be exogenous.

The exogenous technology planner’s optimal tax \(T_{\theta}^{ex}\) then satisfies the conditions provided by the following Lemma.

**Lemma 4.** Suppose \(\Theta = [\theta, \beta]\) and aggregate production \(F\) takes the CES form introduced in Section 2.3. Suppose equilibrium variables are determined according to the reduced form equations (2), (1), (5), and (7), plus the (exogenous) technology equation

\[
\phi^*(l) = \phi^*(l(T)) = \arg\max_{\phi \in \Phi} F(l(T), \phi) \quad \forall l.
\]

Then the exogenous technology planner’s preferred tax \(T_{\theta}^{ex}\) satisfies the following conditions.

\[
\frac{T'(y_\theta)}{1 - T'(y_\theta)} = \left(1 + \frac{1}{\epsilon_\theta}\right) \frac{1 - \overline{b}_\theta}{\overline{b}_\theta} \left(1 - g_\theta\right) + \gamma_{\theta}^{CES} \left(1 - g_\theta\right) + \rho_{\theta}^{CES} \left(1 - g_\theta\right)
\]

for all \(\theta\), where all variables satisfy the equations listed above under the tax \(T_{\theta}^{ex}\), the function \(\overline{b} : \theta \mapsto \overline{b}_\theta\) is given by

\[
\overline{b}_\theta := \kappa_{\theta}^{1+\gamma_{\theta}^{CES}} h_{\theta}^{\gamma_{\theta}^{CES}} \left(\phi^*(T)\right)^{1+\gamma_{\theta}^{CES}} \quad \forall \theta
\]

while \(\overline{B}\) and \(\overline{b}\) are the cumulative distribution and the density function of \(\overline{B}\).

**Proof.** ...

Assuming a Pareto tail for the upper part of the income distribution at the tax rate \(\overline{T}\) yields the following result for the exogenous technology planner’s preferred marginal tax rate in the upper tail.

**Lemma 5.** Let \(\Theta = [\theta, \beta]\) and aggregate production be CES. Suppose equilibrium variables are determined according to the reduced form equations (2), (1), (5), and (7), plus the (exogenous) technology equation

...
equation

$$\phi^*(l) = \phi^*(l(T)) = \arg \max_{\phi \in \Phi} \mathcal{F}(l(T), \phi) \quad \forall l.$$ 

Suppose that the tax $T$ satisfies $T'(y) = \tau^{\text{top}}$ for all $y \geq \bar{y}$ and some threshold $\bar{y}$ and that the hazard rate of the income distribution under tax $T$ satisfies

$$\lim_{\theta \to \bar{y}} \frac{1 - M_{y\theta}}{m_{y\theta} y\theta} = \frac{1}{\alpha}$$

for some $\alpha > 1$. Moreover, let the disutility of labor be iso-elastic and welfare weights satisfy

$$\lim_{\theta \to \bar{y}} g_{\theta} = g^{\text{top}}$$

at the exogenous technology planner’s preferred tax.

Then, the exogenous technology planner’s preferred tax $T^c_T$ satisfies

$$\lim_{\theta \to \bar{y}} \frac{T'(y_{\theta})}{1 - T'(y_{\theta})} = \frac{1 - g^{\text{top}}}{\alpha e} + \frac{\alpha - 1}{\alpha} \gamma^{\text{CES}} (1 - g^{\text{top}}).$$

(32)

Proof. ...

B. Omitted Proofs and Derivations

This section collects all proofs and derivations omitted from the main text.

B.1. Derivation of the labor demand equation (2.1)

Final good firm profits are given by

$$\bar{G}(L, \phi, q) = \int_{\theta}^{\bar{\theta}} w_{\theta} L_{\theta} d\theta - \sum_{j=1}^{I} \int_{0}^{1} p_{j,k} q_{j,k} dk.$$ 

Taking the derivative $D_{L\theta}$ as defined in Section 2.2 and equating it with zero yields:

$$D_{L\theta} \bar{G}(L, \phi, q) = D_{L\theta} \int_{\theta}^{\bar{\theta}} w_{\theta} L_{\theta} d\bar{\theta}$$

The remaining task is to show that

$$D_{L\theta} \int_{\theta}^{\bar{\theta}} w_{\theta} L_{\theta} d\bar{\theta} = w_{\theta}.$$ 

I derive this equality for interior types $\theta \in (\theta, \bar{\theta})$ in detail to demonstrate the working of the functional derivative $D_{L\theta}$. The derivations for the highest and lowest types $\bar{\theta}$ and $\bar{\theta}$ work analogously and are therefore omitted.
By definition:

\[ D_{l_{\theta}} \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} L_{\theta} \, d\vartheta = \lim_{\Delta \to 0} \frac{1}{\Delta} \frac{d}{d\mu} \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} \left( L_{\theta} + \mu L_{\theta, \theta} \right) \, d\vartheta \bigg|_{\mu=0}. \]

Moreover, by definition of \( L_{\theta} \):

\[ \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} \left( L_{\theta} + \mu L_{\theta, \theta} \right) \, d\vartheta = \int_{\theta - \Delta}^{\theta} w_{\theta} \left( L_{\theta} + \mu \frac{\vartheta - \theta + \Delta}{\Delta} \right) \, d\vartheta + \int_{\theta}^{\theta + \Delta} w_{\theta} \left( L_{\theta} + \mu \frac{\theta - \vartheta + \Delta}{\Delta} \right) \, d\vartheta. \]

Hence:

\[ \frac{d}{d\mu} \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} \left( L_{\theta} + \mu L_{\theta, \theta} \right) \, d\vartheta \bigg|_{\mu=0} = \int_{\theta - \Delta}^{\theta} \frac{\vartheta - \theta + \Delta}{\Delta} \, d\vartheta + \int_{\theta}^{\theta + \Delta} \frac{\theta - \vartheta + \Delta}{\Delta} \, d\vartheta. \]

Then, by L'Hôpital's rule:

\[ \lim_{\Delta \to 0} \frac{1}{\Delta} \frac{d}{d\mu} \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} \left( L_{\theta} + \mu L_{\theta, \theta} \right) \, d\vartheta \bigg|_{\mu=0} = \lim_{\Delta \to 0} \frac{1}{2\Delta} \int_{\theta - \Delta}^{\theta} w_{\theta} \, d\vartheta + \lim_{\Delta \to 0} \frac{1}{2\Delta} \int_{\theta}^{\theta + \Delta} w_{\theta} \, d\vartheta. \]

Applying L'Hôpital's rule again, we obtain:

\[ \lim_{\Delta \to 0} \frac{1}{\Delta} \frac{d}{d\mu} \int_{\frac{\vartheta}{2}}^{\vartheta} w_{\theta} \left( L_{\theta} + \mu L_{\theta, \theta} \right) \, d\vartheta \bigg|_{\mu=0} = \lim_{\Delta \to 0} \frac{w_{\theta - \Delta}}{2} + \lim_{\Delta \to 0} \frac{w_{\theta + \Delta}}{2} = w_{\theta}, \]

which is the desired result.

**B.2. Labor Supply Elasticities**

Here I derive expressions (8) and (9) for the labor supply elasticities with respect to the marginal retention rate and the wage. The starting point is workers’ first-order condition (2):

\[ v'(I_{\theta}(T, w_{\theta})) = R_{T}(w_{\theta}l_{\theta}(T, w_{\theta}))w_{\theta}. \]

Taking the derivative \( D_{\tau} \) on both sides of the equation yields:

\[ v''(I_{\theta}(T, w_{\theta}))D_{\tau}I_{\theta}(T, w_{\theta}) = w_{\theta} \frac{d}{d\mu} \left( 1 - T'(w_{\theta}l_{\theta}(T, w_{\theta})) + \mu \right) \bigg|_{\mu=0} - T''(w_{\theta}l_{\theta}(T, w_{\theta}))w_{\theta}^{2}D_{\tau}I_{\theta}(T, w_{\theta}) \]

and hence:

\[ D_{\tau}I_{\theta}(T, w_{\theta}) = \frac{w_{\theta}}{v''(I_{\theta}(T, w_{\theta})) + T''(w_{\theta}l_{\theta}(T, w_{\theta}))w_{\theta}^{2}}. \]

By definition of \( \epsilon_{\theta}^{R} \) we obtain

\[ \epsilon_{\theta}^{R} = \frac{w_{\theta}(1 - T'(w_{\theta}l_{\theta}(T, w_{\theta})))}{v''(I_{\theta}(T, w_{\theta}))T'(w_{\theta}l_{\theta}(T, w_{\theta}))w_{\theta}^{2}}. \]

Again using the first-order condition to replace \( 1 - T'(wl)w \) by \( v'(l) \), we obtain equation (8).
For equation (9) differentiate the first-order condition with respect to \( w_\theta \) on both sides,
\[
v''(l_\theta(T,w_\theta)) \frac{\partial l_\theta(T,w_\theta)}{\partial w_\theta} = 1 - T'(w_\theta l_\theta(T,w_\theta)) - T''(w_\theta l_\theta(T,w_\theta))w_\theta^2 \frac{\partial l_\theta(T,w_\theta)}{\partial w_\theta} - T''(w_\theta l_\theta(T,w_\theta))w_\theta l_\theta(T,w_\theta),
\]
and rearrange it to obtain
\[
\frac{\partial l_\theta(T,w_\theta)}{\partial w_\theta} = \frac{1 - T'(w_\theta l_\theta(T,w_\theta)) - T''(w_\theta l_\theta(T,w_\theta))w_\theta}{v''(l_\theta(T,w_\theta)) + T''(w_\theta l_\theta(T,w_\theta))w_\theta^2}.
\]
Then, use the definition of \( e_\theta^w \) to get
\[
e_\theta^w = \frac{1 - \frac{T''(w_\theta l_\theta(T,w_\theta))w_\theta l_\theta(T,w_\theta)}{1 - T'(w_\theta l_\theta(T,w_\theta))} \left( 1 - \frac{T'(w_\theta l_\theta(T,w_\theta))w_\theta}{v''(l_\theta(T,w_\theta)) l_\theta(T,w_\theta)} \right)}{1 + \frac{T''(w_\theta l_\theta(T,w_\theta))w_\theta l_\theta(T,w_\theta)}{1 - T'(w_\theta l_\theta(T,w_\theta))} \left( 1 - \frac{T'(w_\theta l_\theta(T,w_\theta))w_\theta}{v''(l_\theta(T,w_\theta)) l_\theta(T,w_\theta)} \right)}.
\]
Replacing \( (1 - T'(wl))w \) by \( v'(l) \) yields equation (9).

**B.3. CES Production Function**

Derivations of reduced form equations and wage elasticities for the special case with CES production function...

**B.4. Proof of Observation 1**

The strategy of the proof is to show that statement 1 implies statement 2, statement 2 implies statement 3, and statement 3 implies statement 1.

(1 \(\Rightarrow\) 2) We can transform Definition 1 into
\[
\frac{R_T'(y) - \mu r'(y)}{R_T'(y)} \leq \frac{R_T'(^y) - \mu r'(^y)}{R_T'(^y)} \quad \forall y \geq ^y.
\]
Taking logs and rearranging yields
\[
\log R_T'(y) - \log R_T'(^y) \leq \log R_T'(^y) - \log R_T'(^y) \quad \forall y \geq ^y.
\]
Setting \( ^y = y - d \), dividing both sides of the equation by \( d \), and taking the limit as \( d \to 0 \), we obtain
\[
-\frac{1}{y} P_T(y) \leq -\frac{1}{y} P_T(y) \quad \forall y
\]
and hence
\[
P_T(y) \geq P_T(y) \quad \forall y.
\]

(2 \(\Rightarrow\) 3) Take a function \( r \) such that \( R_T(y) = r(R_T(y)) \) for all \( y \). Differentiating both sides with respect to \( y \) yields
\[
r'(R_T(y)) R_T'(y) = R_T'(y) \quad \forall y,
\]
and after taking logs and rearranging:

$$\log r'(R_T(y)) = \log R'_T(y) - \log R'_T(\tilde{y}) \quad \forall y.$$ 

Differentiating again with respect to $y$ and multiplying through by $y$ gives

$$\frac{r''(R_T(y)) R'_T(\tilde{y})}{r'(R_T(y))} R'_T(y) y = -(P_T(y) - P_T(\tilde{y})) < 0 \quad \forall y,$$

where the inequality is statement 2. The assumption that $T'(y) < 1$ and $T'(y) + \mu \tau'(y) < 1$ for all $y$ implies that $r'(R_T(y)) > 0$ and $R'_T(y) > 0$, such that $r''(R_T(y)) < 0$ for all $y$, which is statement 3.

$(3 \Rightarrow 1)$ Statement 3 implies

$$\frac{R'_T(y)}{R'_T(\tilde{y})} = \frac{r'(R_T(y)) R'_T(y)}{r'(R_T(\tilde{y})) R'_T(\tilde{y})} \leq \frac{R'_T(y)}{R'_T(\tilde{y})} \quad \forall y \geq \tilde{y},$$

because $r$ is concave and $R_T$ strictly increasing. Replacing $R'_T(y)$ by $R'_T(y) - \mu \tau'(y)$ and rearranging yields

$$\frac{R'_T(y) - \mu \tau'(y)}{R'_T(y)} \leq \frac{R'_T(\tilde{y}) - \mu \tau'(\tilde{y})}{R'_T(\tilde{y})} \quad \forall y \geq \tilde{y}$$

and hence:

$$\frac{\tau'(y)}{R'_T(y)} \geq \frac{\tau'(\tilde{y})}{R'_T(\tilde{y})} \quad \forall y \geq \tilde{y},$$

which is statement 1.

**B.5. Proof of Proposition 1**

(ALL PROOFS ARE AVAILABLE UPON REQUEST.)

(They are not yet tidied up.)
B.6. Proof of Proposition 2
B.7. Proof of Corollary 2
B.8. Proof of Proposition 3
B.9. Proof of Proposition 4
B.10. Proof of Proposition 5
B.11. Proof of Proposition 6
B.12. Proof of Proposition 7
B.13. Proof of Proposition 8
B.14. Proof of Corollary 3
B.15. Proof of Corollary 4