Entrepreneurial Taxation with Endogenous Firm Entry and Unemployment

Johan Holmberg*

June 2021

Abstract

This paper deals with optimal nonlinear taxation of labor and entrepreneurial income and extends the recent study of Scheuer (2014) to accommodate equilibrium unemployment. We find that even if employment is endogenous, the government can achieve redistribution of income through taxation without distorting production efficiency. This is possible if the government taxes entrepreneurial and labor income separately. The results also show that including involuntary unemployment creates an incentive to tax entrepreneurial income at lower marginal rates and labor income at higher marginal rates than otherwise.

1 Introduction

This paper deals with the optimal nonlinear taxation of labor earnings and entrepreneurial income in a general equilibrium model in which entrepreneurship is endogenous. We expand the model developed by Scheuer (2014) to include involuntary unemployment. Scheuer (2014) extends the Mirrleesian framework to account for firm and wage formation.[1] The production side of the economy is therefore modeled explicitly, something that many models of optimal taxation abstract from. It is important to consider the interactions between labor supply and labor demand when designing tax systems, as taxes on both labor earnings and firm profits are likely to influence individual behavior — in particular, the choice between employment and entrepreneurship. The demand side of the labor market is important to consider, as it affects both

*Umeå University, e-mail: johan.holmberg@umu.se. First, I wish to express my gratitude to Thomas Aronsson, Gauthier Lanot, David Granlund, and Tomas Sjögren for their valuable input during the writing process. Financial support for the research provided by the Jan Wallander and Tom Hedelius Foundation and the Tore Browaldh Foundation grant no. P2017-0217:1 *Human Capital, Heterogeneity and the Performance of Social Insurance*, is gratefully acknowledged.

[1]Mirrlees (1971) developed a framework for studying optimal redistributive taxation when information between economic agents and the government is asymmetric.

employment level and wage formation within an economy. Separate income tax instruments for entrepreneurs and employees are also relevant to evaluate from a policy perspective: some countries, such as Sweden, have separate income tax treatments for employees and entrepreneurs.[2]

Scheuer uses his model to study optimal redistributive taxation when agents are heterogeneous both in their productivity and in their preferences for entrepreneurship. He finds that optimal marginal taxes depend on the production complementary between labor and entrepreneurship, the distribution of productivity and taste for entrepreneurship, and the elasticities of labor supply and effort relative to after-tax wage rates. If the government cannot discriminate between firm and labor incomes, optimal redistribution requires a distortion of production efficiency and relies on general equilibrium effects for transferring income between entrepreneurs and employees. For example, if production complementarity between entrepreneurial effort and labor supply is strong enough, it is possible to increase the relative wages of employees — compared to the marginal incomes of entrepreneurs — by reducing the marginal tax rate at the top of the income distribution to encourage entrepreneurs to increase their efforts. However, Scheuer also finds that it is possible to achieve redistribution without distorting production efficiency if the government uses two non-linear tax schedules: one for labor earnings and one for entrepreneurial incomes. Thus, direct redistribution is possible using the tax schedule when income types are taxed separately.

The optimal tax policy implications of unemployment have been examined in other contexts, albeit without considering entrepreneurial behavior and the implications thereof.[3] One limitation of Scheuer's analysis is that it relies on the assumption that

[2]In Sweden, among the most debated of these differences are the dividend rules for entrepreneurs with limited liability companies, which yield top marginal tax rates for income declared as capital income and labor income that differ substantially. Alstadsæter and Jacob (2016) noted that if we include payroll taxation (which only contributes to social insurance benefits up to a limit) the top marginal tax rate on labor income was 67% in Sweden in 2011. That same year, the tax rate for entrepreneurial incomes declared as dividends using the so-called "3:12" rules was 41%, which includes the corporate tax rate of 26.3%.

[3]Notable examples are Marceau and Boadway (1994), Aronsson and Sjögren (2004), and Aronsson

there is no unemployment in equilibrium, though we observe unemployment in most countries. Since the Great Depression, unemployment has generally been viewed as the essential signal of economic distress within industrialized economies (Hall et al., 1970). This paper contributes to the literature by relaxing Scheuer's assumption that labor markets clear, thereby allowing for involuntary unemployment. It is important to understand how sensitive policy rules are to this realistic relaxation of the underlying assumptions.

Our analysis shows that it is still possible to achieve redistribution through taxation without distorting production efficiency when employment is endogenous, and therefore when some form of involuntary unemployment can exist. This is also the case when unemployment generates a negative externality on the search cost of employees — i.e., when search costs increase as unemployment grows. The government does not have to distort production efficiency when it can use two nonlinear tax instruments. Our model suggests that we do not have to rely on general equilibrium effects on the labor market to transfer income between entrepreneurs and employees in an efficient way. We also find that the government is unable to perfectly control employees' search effort with a uniform unemployment benefit. Furthermore, we find that the government must compensate employees for their search effort and the unemployment risk they face. This mechanism contributes to reducing the marginal labor income tax rates in comparison to those presented by Scheuer (2014). If unemployment increases the search cost for employees, we find a mechanism that contributes to reducing marginal taxes on entrepreneurial income in comparison to the tax rates from Scheuer.

The model will be presented in the next section. Section 3 provides an analysis of the optimal tax rules and unemployment benefit level. Concluding remarks can be found in Section 4. All calculations for the following analysis can be found in the

and Micheletto (2021). One important takeaway from this literature is that involuntary unemployment can motivate higher marginal tax rates on labor income to increase the employment rate through a division of labor.

mathematical appendix.

2 Baseline model

This section presents the baseline model. First, consider a set of heterogeneous individuals who differ in two dimensions $(\theta, \phi) \in [\underline{\theta}, \bar{\theta}] \times [\underline{\phi}_\theta, \bar{\phi}_\theta]$. We interpret θ as individual productivity and ϕ as the individuals' preference for entrepreneurship. Productivity is assumed to be non-negative, while the preference for entrepreneurship may be positive or negative depending on whether the individual faces an unobserved utility cost or gain from entrepreneurship.[4] Both dimensions of heterogeneity are assumed to be private information. The cumulative distribution function (cdf) for θ is $F(\theta)$, with a probability density function (pdf) $f(\theta)$ while the cdf for ϕ conditional on θ is $G_\theta(\phi)$ with pdf $g_\theta(\phi)$, thus allowing for arbitrary correlation between θ and ϕ.

We normalize the mass of agents to 1. The model is static following Mirrlees (1971), Saez (2001), Lehmann et al. (2011), Scheuer (2014), and much of the modern literature on optimal redistributive taxation. For comparability with Scheuer (2014) we assume that all agents have a utility function that is quasi-linear; i.e., linear in consumption. This linearity means that the marginal utility of consumption will be the same for all individuals independent of their current level of consumption.

Following Scheuer (2014), agents can freely choose between becoming entrepreneurs or employees, and employees are employed by the entrepreneurs. Our model extends this framework by relaxing the assumption that all employees find employment. Agents that choose to become entrepreneurs face the following utility maximization problem:

$$\max_{c,E} U(c, E, \theta) - \phi \equiv c - \psi\Big(\frac{E}{\theta}\Big) - \phi. \tag{1}$$

[4] A positive ϕ can be thought of as the unobserved costs of setting up a firm, such as investment costs. A negative value for ϕ can be interpreted as an unobserved utility gain from entrepreneurship — for instance, due to preferences for self-management or independence.

We assume that consumption c is equal to entrepreneurial income π after tax; i.e., $c = \pi - T_\pi(\pi)$, where $T_\pi(\pi)$ is the tax levied at the given level of entrepreneurial income π.[5] Firm revenue Y is assumed to be a function of efficient labor L and efficient entrepreneurial effort E used in production. The entrepreneurs must pay the market wage w for each efficient unit of labor they hire. Firm profits are therefore $\pi = Y(L, E) - wL$. The production function is assumed to be continuous, increasing, and strictly concave in both inputs, and to exhibit constant returns to scale. Supplying E units of efficient effort induces a utility cost of $\psi\big(\frac{E}{\theta}\big)$, where $\psi'\big(\frac{E}{\theta}\big) > 0$ and $\psi''\big(\frac{E}{\theta}\big) > 0$. Productivity reduces the utility cost of providing any given amount of efficient effort. Note that the preference for entrepreneurship ϕ does not affect the choices entrepreneurs make; it only affects the individual's decision whether or not to become an entrepreneur.

From this optimization problem we can derive the entrepreneurs' private first-order conditions:

$$Y'_L(L, E) = w, \tag{2}$$

$$Y'_E(L, E)\big(1 - T'_\pi(\pi)\big) = \frac{\psi'\big(\frac{E}{\theta}\big)}{\theta}. \tag{3}$$

From these first-order conditions we obtain the optimal demand $L^*(w, \theta)$ of efficient labor and the optimal efficient effort $E^*(w, \theta)$ as well as the indirect utility function for entrepreneurs $v^E(\theta) = c^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)$.

Equilibrium search models (e.g., Mortensen, 1970; Burdett and Mortensen, 1980; Pissarides, 1985; Mortensen and Pissarides, 1994) is the leading theory of equilibrium unemployment. In this type of model, workers lack perfect information and must search for job offers. Following the framework of Scheuer (2014), we use a continuous mea-

[5] It is possible to define the model in terms of expected entrepreneurial incomes instead of using a deterministic function. However, since utility is linear in consumption, this would not affect the decisions the entrepreneurs make.

sure of labor supply. Firms will demand a specific amount of efficient labor instead of posting a set of open positions on the labor market. Thus, unemployment in this model deviates from the traditional search literature as we abstract from vacancies. We also assume that the search is one-sided; i.e., that only employees look for work. A two-sided search would in our continuous setting imply that some firms manage to rent all the efficient labor they need, while others find no labor at all.

Potential employees choose their efficient labor supply l and their efficient search level s. They face the following objective function:

$$\mathbb{E}[U(c,l,\theta)] = H(s)\Big(wl - T_w(wl) - \psi\big(\frac{l}{\theta}\big)\Big) + \big(1 - H(s)\big)b - \varphi\big(\frac{s}{\theta}\big). \tag{4}$$

$H(s)$ is the probability of finding employment given the efficient search level s. We assume that $H'_s(s) > 0$ and $H''_s(s) \leq 0$. Our baseline matching function is similar to Pissarides (1985), where the probability of finding a job depends on search intensity, abstracting from factors that agents cannot control. We later extend the model to include the aggregated unemployment rate in the utility cost of searching to capture the effects of additional competition for jobs.[6] Searching induces a utility cost $\varphi\big(\frac{s}{\theta}\big)$ that is assumed to be increasing, smooth, and convex. An agent who finds employment will have earnings wl, where w is the wage per efficient labor unit. They must also pay a labor income tax $T_w(wl)$. These agents experience the utility cost $\psi\big(\frac{l}{\theta}\big)$ for providing l units of efficient labor. This function shares the same properties as the cost of effort of the entrepreneurs. An agent fails to find employment with probability $1-H(s)$. The unemployed agent receives unemployment benefits b from the government. We assume that individuals use their disposable income for consumption c. From this, we derive the following private first-order conditions for employees:

[6] Adding an externality from the aggregate search level is not as informative, since the government does not have instruments for perfectly controlling search.

$$w\big(1 - T_w'(wl)\big) = \frac{\psi'(\frac{l}{\theta})}{\theta}, \tag{5}$$

$$H_s'(s)\big(wl - T_w(wl) - \psi(\frac{l}{\theta}) - b\big) = \frac{\varphi'(\frac{s}{\theta})}{\theta}. \tag{6}$$

Using these first-order conditions, we find the optimal supply of efficient labor $l^*(w_n, \theta)$, where $w_n = w\big(1 - T_w'(wl)\big)$, and the optimal search level $s^*\big(wl - T_w(wl), \theta, b\big)$. This optimization also characterizes the indirect expected utility function of the employee $v^W(\theta) = c^W(\theta) - H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) - \varphi\big(\frac{s(\theta)}{\theta}\big)$.

The choice of occupation — whether to become an entrepreneur or an employee — depends on the relative utilities of the two options and the individual preferences for entrepreneurship. To facilitate the interpretation we will assume that $\phi > 0$. Let $\tilde{\phi}(\theta)$ be the utility cost of setting up a firm such that an individual of type θ is indifferent to whether they become an entrepreneur or an employee:

$$\tilde{\phi}(\theta) = \begin{cases} \bar{\phi}_\theta & \text{if } v^E(\theta) - v^W(\theta) \geq \bar{\phi}_\theta \\ \underline{\phi}_\theta & \text{if } v^E(\theta) - v^W(\theta) \leq \underline{\phi}_\theta \\ v^E(\theta) - v^W(\theta), & \text{otherwise.} \end{cases} \tag{7}$$

If the unobserved utility cost of setting up a firm is lower than this threshold — i.e., $\phi \leq \tilde{\phi}(\theta)$ — the agent becomes an entrepreneur. The share of entrepreneurs will be equal to $G_\theta(\tilde{\phi}(\theta))$.

The government has access to general nonlinear taxes on income for both employees and entrepreneurs as well as a uniform unemployment benefit, b. Using a type-specific unemployment benefit is impossible here, as individual productivity θ is private information. This unobserved heterogeneity also means that the government faces two incentive compatibility constraints. Additionally, the government faces a labor market clearing condition, a resource constraint, and a non-negativity constraint for the unemployment benefit. The incentive compatibility constraint for entrepreneurs can be written as:

$$v_{\theta}^{E}{}'(\theta) = \psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2} \quad \forall\theta \in [\underline{\theta}, \bar{\theta}], \tag{8}$$

which states that an entrepreneur cannot increase her utility by behaving like an adjacent productivity type. The social planner faces an analogous constraint for potential employees:

$$v_{\theta}^{W}{}'(\theta) = H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^2} \quad \forall\theta \in [\underline{\theta}, \bar{\theta}]. \tag{9}$$

This states that it has to be in the potential employee's self-interest to behave as their true productivity type. Although productivity is private information, we assume that the social planner can observe income and identify individuals as employees or entrepreneurs.[7]

The social planner also faces a labor market clearing condition that states that the amount of labor employed in the economy cannot exceed the total labor demanded by the entrepreneurs. Employees do not create labor demand, and the social planner cannot assign more work than the firms use in their production.

$$\int_{\underline{\theta}}^{\bar{\theta}} G_{\theta}\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta) \geq \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_{\theta}\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta)dF(\theta). \tag{10}$$

In our model, all workers can be employed by any firm, and we abstract from vacancies. An alternative specification can be found in Lehmann et al. (2011), who analyze a model of optimal taxation under friction unemployment with a discrete measure of

[7] Some, such as Scheuer (2014), present the incentive compatibility constraints as $v^E(\theta) \equiv \max_{\tilde{\theta}\forall\theta\in[\underline{\theta},\bar{\theta}]} c^E(\tilde{\theta}) - \psi\big(\frac{E(\tilde{\theta})}{\theta}\big)$, and $v^W(\theta) \equiv \max_{\tilde{\theta}\forall\theta\in[\underline{\theta},\bar{\theta}]} c^W(\tilde{\theta}) - H\big(s(\tilde{\theta})\big)\psi\big(\frac{l(\tilde{\theta})}{\theta}\big) - \varphi\big(\frac{s(\tilde{\theta})}{\theta}\big)$. However, these restrictions are more general than those used to solve the model, as we only allow for mimicking of the adjacent ability type, not across the whole productivity distribution. Using a model with discrete ability types, such as Stiglitz (1982), allows for a more detailed description of mimicking behavior.

labor and a continuum of labor markets for workers with varying ability levels. Our analysis is analogous to theirs in some respects, but uses a continuous measure of labor and includes firm formation. The final constraint for the social planner is the resource constraint.[8]

$$\int_{\underline{\theta}}^{\bar{\theta}} \Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta), E(\theta)\big) - c^E(\theta)\Big) - \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)c^W(\theta)\Big] dF(\theta) \geq 0, \tag{11}$$

and a non-negativity constraint for unemployment benefits:

$$b \geq 0. \tag{12}$$

The Pareto weights of the social planner are indicated by the addition of a tilde to the distribution. For example, the Pareto weights corresponding to $F(\theta)$ are denoted by $\tilde{F}(\theta)$, and so forth. These weights capture the social planner's valuation of the utility for individuals based on their occupation (i.e., entrepreneur or employee), productivity level, and density within the population. If the Pareto weight assigned to a group is larger than its density within the population, the social planner places additional value on the utility of these agents and has an incentive to assign them additional resources.

Following much of the earlier literature on optimal taxation, we solve the social decision problem as a direct problem where l and E are control variables, and $v^E(\theta)$ and $v^W(\theta)$ are state variables. It is possible to solve the model using the indirect utility of employed wage earners as a state variable instead of $v^W(\theta)$. This approach yields the same policy rules for the marginal tax rates. However, the choice between sequential and simultaneous decisions of search effort and labor supply affects the policy rule for the unemployment benefit.[9] The marginal tax rates can be calculated by comparing the private and social first-order conditions. Note that the social planner takes the utility cost of entrepreneurship $\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)} \phi d\tilde{G}_\theta(\phi) d\tilde{F}(\theta)$ into account when formulating

[8]Note that we can find that $c^W(\theta) = v^W(\theta) + H\big(s(\theta)\big)\psi\big(l(\theta)/\theta\big) + \varphi\big(s(\theta)/\theta\big)$ and $c^E(\theta) = v^E(\theta) + \psi\big(E(\theta)/\theta\big)$ from the indirect utility functions.

[9]If search and labor supply are sequential decisions, the unemployment benefit level would not affect the incentive compatibility constraint of the employed. We assume that the decisions are made simultaneously to allow for mimicking in search behavior.

policies. The problem of the social planner is as follows:

$$
\begin{aligned}
&\max_{E(\theta),L(\theta),l(\theta),v^E(\theta),v^W(\theta),\tilde{\phi}(\theta),b} \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big[\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)v^E(\theta) + \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)v^W(\theta)\Big]d\tilde{F}(\theta) \\
&- \int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)} \phi d\tilde{G}_\theta(\phi)d\tilde{F}(\theta), \\
&s.t. \quad v_\theta^{E\prime}(\theta) = \psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2} \quad \forall\theta\in[\underline{\theta},\bar{\theta}], \\
&\qquad v_\theta^{W\prime}(\theta) = H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^2} \quad \forall\theta\in[\underline{\theta},\bar{\theta}], \\
&\int_{\underline{\theta}}^{\bar{\theta}} G_\theta\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta) - \int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta)dF(\theta) \geq 0, \\
&\int_{\underline{\theta}}^{\bar{\theta}} \Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta),E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)\Big) \\
&- \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)\Big)\Big]dF(\theta) \geq 0, \\
&b \geq 0.
\end{aligned}
\tag{13}
$$

3 Results

In this section, we present our results. We begin by presenting the results of the baseline model. At the optimum, the marginal product of efficient labor $Y_L'\big(L(\theta),E(\theta)\big)$ is independent of θ. From this, we know that the marginal productivity of efficient labor will be the same for all firms. Furthermore, the fact that the production function has constant returns to scale also implies that the marginal productivity of efficient effort will be identical for all firms. First, this indicates that we need not distort production efficiency to achieve redistribution using the two nonlinear tax instruments available to the social planner. Second, this means that the wages per unit of efficient labor and marginal profits per unit of efficient effort $\tilde{w}$ will be the same for all individuals. More formally, we find:

$$Y_L'\big(L(\theta),E(\theta)\big) = w \quad \forall\theta\in[\underline{\theta},\bar{\theta}], \tag{14}$$

$$Y_E'\big(L(\theta),E(\theta)\big) = \tilde{w} \quad \forall\theta\in[\underline{\theta},\bar{\theta}]. \tag{15}$$

Note that since we define labor and effort inputs in terms of efficient units, there is no implication that wages or marginal profits per hour worked will be the same for all types. Instead, the implication is that individuals will receive payments equal to their marginal productivity.

Proposition 1:

Define:

$$\begin{aligned}
\epsilon_\pi(\theta) &= \frac{\partial E(\theta)}{\partial \tilde{w}\big(1-T'_\pi(\theta)\big)} \frac{\tilde{w}\big(1-T'_\pi(\theta)\big)}{E(\theta)}, \\
\epsilon_w(\theta) &= \frac{\partial l(\theta)}{\partial w\big(1-T'_w(\theta)\big)} \frac{w\big(1-T'_w(\theta)\big)}{l(\theta)}, \\
\Delta\Pi(\theta) &\equiv G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta) - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\tilde{f}(\theta), \\
\Delta\Omega(\theta) &\equiv \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta), \\
\Delta T(\theta) &\equiv T_\pi\big(\pi(\theta)\big) - T^e_w\big(wl(\theta)\big).
\end{aligned}$$

In the baseline model, the marginal tax policies satisfy the following policy rules:

$$\frac{T'_\pi(\theta)}{1-T'_\pi(\theta)} = \Big(1 + \frac{1}{\epsilon_\pi(\theta)}\Big) \frac{\int_\theta^{\bar{\theta}} \Big[\Delta\Pi(\hat{\theta}) - g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})\Big] d\hat{\theta}}{\theta G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)}, \tag{16}$$

$$\frac{T'_w(\theta)}{1-T'_w(\theta)} = \Big(1 + \frac{1}{\epsilon_w(\theta)}\Big) \frac{\int_\theta^{\bar{\theta}} \Big[\Delta\Omega(\hat{\theta}) + g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})\Big] d\hat{\theta}}{\theta\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)}. \tag{17}$$

$T^e_w\big(wl(\theta)\big)$ refers to the expected tax payments of potential employees with productivity θ.[10] We define the difference between a mass of agents and the Pareto weights assigned to them by the social planner at a given θ as $\Delta\Pi(\theta)$ for entrepreneurs and $\Delta\Omega(\theta)$ for employees, and $\Delta T(\theta)$ as the difference between the tax payment of entrepreneurial and labor income at a given productivity level.

The difference between our results and those of Scheuer (2014) can be found in

[10]Since not all potential employees find employment, the expected tax payments of a potential employee with productivity θ will be $T^e_w\big(wl(\theta)\big) = H\big(s(\theta)\big)T_w\big(wl(\theta)\big) - \Big(1 - H\big(s(\theta)\big)\Big)b$.

the definition of this tax wedge. In our formula, the expected tax paid by employees $T_w^e\big(wl(\theta)\big)$ given their productivity θ accounts for the unemployment risk, unemployment benefits, and search. If the social planner were to not compensate employees for their search effort, more agents would choose to become entrepreneurs. We find that $T_w\big(wl(\theta)\big) = H\big(s(\theta)\big)wl(\theta) - c^W(\theta)$, where the employees' expected consumption includes compensation for their search effort as well as the unemployment benefits for the unemployed.

Using the transversality conditions, we find that $T'_\pi\big(\pi(\underline{\theta})\big) = T'_\pi\big(\pi(\bar{\theta})\big) = T'_w\big(wl(\underline{\theta})\big) = T'_w\big(wl(\bar{\theta})\big) = 0$. This is the standard result that labor supply (in this case also entrepreneurial effort) should not be distorted at the top and bottom of the ability distribution. Marginal tax rate levels will depend on the behavioral response of the agents. The more the marginal tax rates distort the labor or effort supply, the lower these marginal tax rates should be. This effect is captured by the elasticity of labor supply with respect to the after-tax wage rate $\epsilon_w(\theta)$ as well as the elasticity of entrepreneurial effort with respect to the marginal income after tax $\epsilon_w(\theta)$. The higher this elasticity is, the lower the optimal marginal tax rate becomes.

Raising the marginal taxes at productivity level θ will increase the tax revenue generated from the higher-ability groups at the cost of distorting local effort and labor supply. The mass of agents affected by this is represented by $\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)$ for workers and $G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)$ for entrepreneurs. Recall that the Pareto weights capture how the social planner values the utility of particular skill and occupation groups. Thus, the difference between the mass of agents and the Pareto weights the social planner assigns to them[11] captures the redistributive properties of the tax system. If the difference is positive, the social planner will raise the tax for productivity type θ to redistribute these resources to other agents.

Redistribution can either occur between productivity types or between occupations.

[11] $\Delta\Pi(\theta)$ for entrepreneurs and $\Delta\Omega(\theta)$ for employees.

If $f(\hat{\theta}) > \tilde{f}(\hat{\theta})$, the government will assign higher taxes to the more productive agents. In this case, the social value of the additional tax revenue generated by a more progressive tax rate exceeds the loss in utility for the more productive agents. In the case that $G_\theta\big(\tilde{\phi}(\hat{\theta})\big) > \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)$, the government will establish incentives to redistribute income from entrepreneurs to employees, and vice versa if $G_\theta\big(\tilde{\phi}(\hat{\theta})\big) < \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)$. Since we do not include any public goods in this model, the additional tax revenue is used exclusively for income redistribution.

The manner in which the optimal marginal tax rates are affected by occupational choice is captured by the term $g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})$. It is important to consider how differences in taxation between different income types affect agents' choice of occupation. For example, if we increase the marginal tax on entrepreneurial income, some entrepreneurs will choose to become employees. In Equation (16), this mass of agents is captured by the term $-g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})$. If the tax paid on entrepreneurial income at this productivity level is different from the tax paid by equally productive employees, the occupational change will affect the government budget. If entrepreneurs pay higher taxes $\Delta T(\hat{\theta}) > 0$, the government loses income as agents leave entrepreneurship, and vice versa if entrepreneurs pay lower taxes $\Delta T(\hat{\theta}) < 0$. Thus, occupational choices give rise to a tax revenue motive for marginal taxation.

Since the two non-linear tax instruments effectively control labor supply and entrepreneurial effort, we can use the benefit level b to influence the search efforts of the employees. We assume quasi-linear utility functions, meaning that the marginal utility of consumption will be the same for all income groups. This removes one of the main incentives for redistribution of incomes from the employed to the unemployed, as the social planner does not have any preferences for redistribution within specific productivity and occupation groups. In our framework, the benefit level affects search behavior and the incentive compatibility constraint of employees.[12] In this paper, we

[12]In a sequential model, the unemployment benefit would not affect the incentive compatibility constraint of employees.

are interested to see whether these effects are enough by themselves to motivate a positive unemployment benefit level. The policy rule for the unemployment benefit can be written as follows:

$$\begin{aligned} b = & \frac{\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta) \frac{\partial s(\theta)}{\partial b} \Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big) d\theta - \gamma}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big) dF(\theta)} \\ & - \frac{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big) T\big(wl(\theta)\big) dF(\theta)}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big) dF(\theta)}. \end{aligned} \tag{18}$$

We find that $\frac{\partial s(\theta)}{\partial b} < 0$ by applying the implicit function theorem to the private first-order conditions of the employees. Thus, an increase in the unemployment benefit level will reduce the search effort of the potential employees. The denominator in Equation (18) captures the total reduction in employment following an increase in the benefit level. Since we assume that search effort has a positive effect on the probability of finding employment, this term will always be negative.

The numerator of the first term comprises an integral and a Kuhn-Tucker multiplier γ, respectively. From the integral, we find that the benefit level affects the incentive compatibility constraint for employees. An increase in the benefit level reduces the efficient search level, which both decreases the probability of finding work and the utility cost of searching for work. Since the denominator is negative, $\mu^W(\theta) \geq 0 \; \forall \theta \in [\underline{\theta}, \bar{\theta}]$, and as we assume that the marginal utility costs of working and searching are positive and increasing, we know that the first term is positive. If the non-negativity constraint for unemployment benefits is binding, $\gamma > 0$ and $b = 0$.

The numerator of the second term captures the total decrease in tax revenue following an increase in the unemployment benefit. As search effort decreases, unemployment increases, which reduces the total tax revenue collected from the employees. This term will be negative as long as the weighted sum of all taxes paid by employees is positive. The unemployment benefit b will be positive if the social value of relaxing the incentive compatibility constraint for employees dominates the budget effect from reduced search effort.

3.1 Search externality of unemployment

In the search literature, it is common to include a measure of labor market tightness in the matching function (e.g., Shimer, 2005; Lehmann et al., 2011). As the model uses a continuous measure of labor and has a population normalized to 1, it is not possible to define discrete vacancies. Instead, we assume that the utility cost of searching depends on the unemployment level.[13] The cost of searching becomes $\varphi\big(\frac{s(\theta)}{\theta}, u\big)$ were:

$$u = \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(1 - H\big(s(\theta)\big)\Big) dF(\theta). \tag{19}$$

Each individual treats the economy-wide unemployment level as exogenous. If this external effect is negative — i.e., a higher unemployment level increases the individual cost of searching — then $\varphi'_u\big(\frac{s(\theta)}{\theta}, u\big) > 0$. We will assume that the externality affects the marginal cost of searching in a symmetric way to its overall effect on the search cost; i.e., if $\varphi'_u\big(\frac{s(\theta)}{\theta}, u\big) > 0$ then $\varphi''_{s,u}\big(\frac{s(\theta)}{\theta}, u\big) \geq 0$. If a higher unemployment level increases the search cost for employees, this assumption prevents the marginal cost of searching from declining following an increase in unemployment. This is a reasonable assumption that facilitates the interpretation of our results. We add the unemployment externality to the social planner's problem with the Lagrange multiplier κ. All calculations related to the analysis of this version of the model can be found in Appendix B.

We find that the marginal productivity of efficient labor and effort remains equalized across all firms. The optimal benefit rule from this model is similar to Equation (18), with the only difference being that the benefit rule of this model also takes the social

[13] Including a direct effect from unemployment in the probability of finding employment is problematic. These variables would be co-dependent, as the unemployment level is determined by the unemployment risk for employees.

value of decreasing unemployment κ into account. This increases the probability of a binding non-negativity constraint for the benefit level. If we use the social first-order condition for unemployment u to solve for the shadow price of unemployment, we find:

$$\kappa = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\Big[\varphi''_{s,u}\big(\frac{s(\theta)}{\theta},u\big)\frac{s(\theta)}{\theta^2} + \frac{\partial s(\theta)}{\partial u}\Big(H'_s\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \frac{\varphi'_s\big(\frac{s(\theta)}{\theta},u\big)}{\theta^2} + \varphi''_s\big(\frac{s(\theta)}{\theta},u\big)\frac{s(\theta)}{\theta^3}\Big)\Big]d\theta}{1+\int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial u}H'_s\big(s(\theta)\big)dF(\theta)}$$
$$- \frac{\int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big[\varphi'_u\big(\frac{s(\theta)}{\theta},u\big) + \frac{\partial s(\theta)}{\partial u}H'_s\big(s(\theta)\big)\Big(T_w\big(wl(\theta)\big)+b\Big)\Big]dF(\theta)}{1+\int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial u}H'_s\big(s(\theta)\big)dF(\theta)}. \tag{20}$$

The denominator of Equation (20) is positive. To see this, note that the direction of the externality and its effect on search effort will determine the sign of the second term. Since the externality from unemployment increases the search cost (i.e., $\varphi'_u\big(\frac{s(\theta)}{\theta},u\big)>0$) and search effort is affected negatively by unemployment,[14] this integral will be equal to the total reduction in employment among employees resulting from a reduction in search effort as the unemployment level increases. Since this marginal reduction in employment cannot exceed 1, we know that the denominator must be positive.

The numerator of the first term captures the effect of the externality on the incentive compatibility constraint of employees. Following our assumptions regarding the utility cost of labor and search effort, we know that the sign of this term will be determined by the direction of the externality as well as its indirect effect on search effort. If the externality is negative and the incentive compatibility constraint is binding for some ability group, this sum will be negative. If the constraint is not binding anywhere (i.e., when $\mu^W(\theta)=0\,\forall\theta\in[\underline{\theta},\bar{\theta}]$), this term will be 0.

[14]Given the assumptions we make for the utility cost of searching and since we assume that $H''_s\big(s(\theta)\big)\leq 0$, we can use the private first-order condition for search effort and the implicit function theorem to determine that the sign of $\frac{\partial s(\theta)}{\partial u}$ will be the opposite of the sign of $\varphi''_{s,u}\big(\frac{s(\theta)}{\theta},u\big)/\theta$. In the case that $\varphi''_{s,u}\big(\frac{s(\theta)}{\theta},u\big)/\theta=0$, the externality will not have any effect on the search effort of the employees.

The second term will be positive if we assume that unemployment increases search costs. This term captures the effects that increasing search costs have on the government budget and the aggregated utility of the employees. Increasing search costs reduces the utility of the employees $\varphi'_u\big(\frac{s(\theta)}{\theta}, u\big)$. The integral of this term captures the employees' aggregated marginal willingness to pay to reduce the unemployment level. We also find that changes in search behavior affect the government budget. When employees search less intensively, more become unemployed. This reduces the government tax revenue from labor income as the total expenditure on unemployment benefits simultaneously increases. Whenever the effect of the externality on the government budget and the utility of employees is greater than the effect the externality has on the incentive compatibility constraint, κ will have the same sign as the externality. From this, we conclude that if unemployment increases search costs, the social value of reducing unemployment is, under reasonable assumptions, positive.

Proposition 2

By adding a search externality of unemployment to the baseline model, the policy rules for marginal taxation change to read:

$$\frac{T'_\pi}{1-T'_\pi} = \big(1+\frac{1}{\epsilon_\pi(\theta)}\big)\frac{\int_\theta^{\bar\theta}\Big[\Delta\Pi(\hat\theta) - g_\theta\big(\tilde\phi(\hat\theta)\big)f(\hat\theta)\Big(\kappa\Big(1-H\big(s(\hat\theta)\big)\Big)+\Delta T(\hat\theta)\Big)\Big]d\hat\theta}{\theta\Big(1-G_\theta\big(\tilde\phi(\theta)\big)\Big)f(\theta)}, \tag{21}$$

$$\frac{T'_w}{1-T'_w} = \big(1+\frac{1}{\epsilon_w(\theta)}\big)\frac{\int_\theta^{\bar\theta}\Big[\Delta\Omega(\hat\theta) + g_\theta\big(\tilde\phi(\hat\theta)\big)f(\hat\theta)\Big(\kappa\Big(1-H\big(s(\hat\theta)\big)\Big)+\Delta T(\hat\theta)\Big)\Big]d\hat\theta}{\theta\Big(1-G_\theta\big(\tilde\phi(\theta)\big)\Big)f(\theta)}. \tag{22}$$

From these expressions, we find an additional effect from the level of unemployment proportional to $\kappa\Big(1-H\big(s(\hat\theta)\big)\Big)$. Increasing the marginal tax rate on entrepreneurial income at a productivity level θ will induce agents to switch occupations and become employees. This increases unemployment in the economy both by reducing labor demand and by increasing labor supply. By increasing the search cost, higher unemployment also affects the utility of all employees. As unemployment increases, the cost of searching increases as well, and if κ is positive,[15] the optimal marginal tax rate on entrepreneurial income is lower than in the baseline scenario, ceteris paribus. Similarly, the optimal marginal tax rates on labor income will be higher. If unemployment increases the utility cost of searching, it could motivate lower marginal tax rates for entrepreneurs with high incomes.

4 Concluding remarks

In this paper, we analyze the optimal nonlinear taxation of labor and entrepreneurial income and extend the recent study of Scheuer (2014) to accommodate involuntary unemployment. Our analysis shows that redistribution can be achieved through taxation without distorting production efficiency when employment is endogenous to the model, as long as the government can use two non-linear tax instruments. This is also true if we have a negative externality from unemployment on the search cost of employees. Thus, we need not rely on general equilibrium effects on the labor market to transfer income between entrepreneurs and employees; for example, using the complementarity between labor and entrepreneurial effort in the production to increase wages of employees.[16]

The relative sizes of the marginal tax rates for labor and entrepreneurial income will depend on both individual-level characteristics, such as preferences for entrepreneurship, as well as wider labor market implications, such as the manner in which the aggregate unemployment level affects the search cost and utility of employees. We find that if unemployment increases the search cost of employees, the optimal marginal taxes on entrepreneurial income decrease, under reasonable assumptions.

[15] This is true if the effect of the externality on the utility of employees and the government budget is greater than its effect on the incentive compatibility constraint of employees.

[16] If entrepreneurial effort and labor are complements in the firms' production, we can use the tax schedule to incentivize the most productive entrepreneurs to exert more effort, which in turn would increase the equilibrium wage rate for employees.

A limitation of this model is that the quasi-linear utility function makes the social planner indifferent to any income inequality within a given productivity and occupation group, such as inequality resulting from involuntary unemployment. A possible topic for future research could be to study how the tax rules would be affected by a utility function that is concave in consumption. A decreasing marginal utility of consumption would create stronger incentives for the social planner to redistribute income from the employed to the unemployed.

References

Alstadsæter, A. and Jacob, M. (2016). Dividend taxes and income shifting. *The Scandinavian Journal of Economics*, 118(4):693–717.

Aronsson, T. and Micheletto, L. (2021). Optimal redistributive income taxation and efficiency wages. *The Scandinavian Journal of Economics*, 123(1):3–32.

Aronsson, T. and Sjögren, T. (2004). Is the optimal labor income tax progressive in a unionized economy? *Scandinavian Journal of Economics*, 106(4):661–675.

Burdett, K. and Mortensen, D. T. (1980). Search, layoffs, and labor market equilibrium. *Journal of Political Economy*, 88(4):652–672.

Hall, R. E., Gordon, R., and Holt, C. (1970). Why is the unemployment rate so high at full employment? *Brookings papers on economic activity*, 1970(3):369–410.

Lehmann, E., Parmentier, A., and Van der Linden, B. (2011). Optimal income taxation with endogenous participation and search unemployment. *Journal of Public Economics*, 95(11-12):1523–1537.

Marceau, N. and Boadway, R. (1994). Minimum wage legislation and unemployment insurance as instruments for redistribution. *The Scandinavian Journal of Economics*, pages 67–81.

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

Mortensen, D. T. (1970). Job search, the duration of unemployment, and the phillips curve. *The American Economic Review*, 60(5):847–862.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

Pissarides, C. A. (1985). Short-run equilibrium dynamics of unemployment, vacancies, and real wages. *The American Economic Review*, 75(4):676–690.

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

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

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 95(1):25–49.

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

Appendix A:

The problem of the social planner can be stated as follows:

$$
\begin{aligned}
&\max_{E(\theta),L(\theta),l(\theta),v^E(\theta),v^W(\theta),\tilde{\phi}(\theta),b} \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big[\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)v^E(\theta) + \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)v^W(\theta)\Big]d\tilde{F}(\theta) \\
&- \int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)} \phi d\tilde{G}_\theta(\phi)d\tilde{F}(\theta) \\
&s.t. \quad v_\theta^{E\prime}(\theta) = \psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2} \quad \forall\theta\in[\underline{\theta},\bar{\theta}], \\
&\qquad v_\theta^{W\prime}(\theta) = H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^2} \quad \forall\theta\in[\underline{\theta},\bar{\theta}], \\
&\int_{\underline{\theta}}^{\bar{\theta}} G_\theta\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta) - \int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta)dF(\theta) \geq 0, \\
&\int_{\underline{\theta}}^{\bar{\theta}}\Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta),E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)\Big), \\
&-\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)\Big)\Big]dF(\theta) \geq 0, \\
&b \geq 0.
\end{aligned}
\tag{A.1}
$$

From this, we get the following Lagrangian after integrating the incentive compability constraints by parts:

$$
\begin{aligned}
\mathcal{L} &= \int_{\underline{\theta}}^{\bar{\theta}}\Big[\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)v^E(\theta) - \int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)}\phi d\tilde{G}_\theta(\phi) + \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)v^W(\theta)\Big]d\tilde{F}(\theta) \\
&- \int_{\underline{\theta}}^{\bar{\theta}}\Big[\mu_\theta^{E\prime}(\theta)v^E(\theta) + \mu^E(\theta)\psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2}\Big]d\theta \\
&- \int_{\underline{\theta}}^{\bar{\theta}}\Big[\mu_\theta^{W\prime}(\theta)v^W(\theta) + \mu^W(\theta)\Big(H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^2}\Big)\Big]d\theta \\
&- \lambda_L\Big[\int_{\underline{\theta}}^{\bar{\theta}} G_\theta\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta) - \int_{\underline{\theta}}^{\bar{\theta}}\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta)dF(\theta)\Big] \\
&+ \lambda_R\Big[\int_{\underline{\theta}}^{\bar{\theta}}\Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta),E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)\Big) \\
&\qquad - \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)\Big)\Big]dF(\theta)\Big] \\
&+ \gamma b.
\end{aligned}
\tag{A.2}
$$

From $\partial\mathcal{L}/\partial L(\theta) = 0$ we find that:

$$Y'_L\big(L(\theta), E(\theta)\big) = \frac{\lambda_L}{\lambda_R} \quad \forall\theta \in [\underline{\theta}, \bar{\theta}]. \tag{A.3}$$

The production function is assumed to have a constant returns to scale which implies that $Y(L, E) = Y'_L(L, E)L + Y'_E(L, E)E$. From this and from $Y'_L\big(L(\theta), E(\theta)\big) = \lambda_L/\lambda_R$ we find that the marginal productivity of efficient labor and effort must be equal across all firms. This also implies that the payment per efficient unit of labor or profit per efficient unit of effort is the same across the skill distribution:

$$\begin{aligned} Y'_L(L(\theta), E(\theta)) &= w \quad \forall\theta \in [\underline{\theta}, \bar{\theta}], \\ Y'_E(L(\theta), E(\theta)) &= \tilde{w} \quad \forall\theta \in [\underline{\theta}, \bar{\theta}]. \end{aligned} \tag{A.4}$$

From $\partial\mathcal{L}/\partial v^W(\theta) = 0$ we find:

$$\begin{aligned} \mu_\theta^{W\prime}(\theta) = {} & \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) + \lambda_L g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(H\big(s(\theta)\big)l(\theta) + L(\theta)\Big) \\ & - \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(Y\big(L(\theta), E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big) + v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)\Big). \end{aligned} \tag{A.5}$$

Using the constant returns to scale of the production function, that $\lambda_L = w\lambda_R$, and that $c^E(\theta) = v^E(\theta) + \psi\big(\frac{E(\theta)}{\theta}\big)$ and $c^W(\theta) = v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)$ we get:

$$\begin{aligned} \mu_\theta^{W\prime}(\theta) = {} & \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) \\ & - \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(\tilde{w}E(\theta) - c^E(\theta) - \Big(wH\big(s(\theta)\big)l(\theta) - c^W(\theta)\Big)\Big). \end{aligned} \tag{A.6}$$

We can define the tax wedge between entrepreneurs and employees as $\Delta T(\theta) \equiv \tilde{w}E(\theta) - c^E(\theta) - \big(wH\big(s(\theta)\big)l(\theta) - c^W(\theta)\big)$. Using this we get:

$$\mu_\theta^{W\prime}(\theta) = \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\theta). \tag{A.7}$$

Using similar steps, we can from $\partial\mathcal{L}/\partial v^E(\theta) = 0$ derive:

$$\mu_\theta^{E\prime}(\theta) = \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\tilde{f}(\theta) - \lambda_R G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta) + \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\theta). \tag{A.8}$$

From $\int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{E\prime}(\theta)d\theta = 0$ we find that $\lambda_R = 1$. We can thus integrate $\mu^W(\theta)$ and find:

$$\mu^W(\bar{\theta}) = \mu^W(\theta) + \int_\theta^{\bar{\theta}} \Big[\Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)\tilde{f}(\hat{\theta}) - \Big(1 - G_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)f(\hat{\theta}) - g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})\Big]d\hat{\theta}. \tag{A.9}$$

Using the transversality conditions again we get:

$$\mu^W(\theta) = \int_\theta^{\bar{\theta}} \Big[\Big(1 - G_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)f(\hat{\theta}) - \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)\tilde{f}(\hat{\theta}) + g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\hat{\theta})\Big]d\hat{\theta}. \tag{A.10}$$

With the same steps for $\mu^E(\theta)$ we find:

$$\mu^W(\theta) = \int_\theta^{\bar{\theta}} \Big[G_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta}) - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\tilde{f}(\hat{\theta}) - g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\hat{\theta})\Big]d\hat{\theta}. \tag{A.11}$$

If we use that $\lambda_L = w\lambda_R$, that $\lambda_R = 1$, and $\partial\mathcal{L}/\partial l(\theta) = 0$ we get:

$$\mu^W(\theta)H\big(s(\theta)\big)\Big(\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta^2 + \psi''\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^3}\Big) = \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)H\big(s(\theta)\big)\big(w - \psi'\big(\frac{l(\theta)}{\theta}\big)/\theta\big). \tag{A.12}$$

Which we can write as:

$$\frac{\mu^W(\theta)\Big(\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta + \psi''\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2}\Big)}{\theta\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)} = w - \psi'\big(\frac{l(\theta)}{\theta}\big)/\theta. \tag{A.13}$$

We can use the first-order condition of the employee $H\big(s(\theta)\big)\Big(w\big(1 - T_w'(\theta)\big) - \psi'\big(\frac{l(\theta)}{\theta}\big)/\theta\Big) = 0$ and the implicit function theorem to define the elasticity of labor supply with respect to after tax wages:

$$\begin{aligned}
\epsilon_w(\theta) &= \frac{\partial l(\theta)}{\partial w\big(1 - T_w'(\theta)\big)}\frac{w\big(1 - T_w'(\theta)\big)}{l(\theta)} \\
&= -\frac{H\big(s(\theta)\big)}{-H\big(s(\theta)\big)\psi''\big(\frac{l(\theta)}{\theta}\big)/\theta^2}\frac{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{l(\theta)} \\
&= \frac{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{l(\theta)\psi''\big(\frac{l(\theta)}{\theta}\big)/\theta^2}.
\end{aligned} \tag{A.14}$$

From this, we find:

$$\frac{\mu^W(\theta)\big(1+\dfrac{1}{\epsilon_w(\theta)}\big)}{\theta\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)} = \frac{w-\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta} = \frac{w-w\big(1-T'_w(\theta)\big)}{w\big(1-T'_w(\theta)\big)}. \tag{A.15}$$

Combining this with Equation (A.10), and defining $\Delta\Omega(\theta) \equiv \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta)$ we find:

$$\frac{T'_w(\theta)}{1-T'_w(\theta)} = \Big(1+\frac{1}{\epsilon_w(\theta)}\Big)\frac{\int_\theta^{\bar{\theta}}\Big[\Delta\Omega(\hat{\theta})+g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})\Big]d\hat{\theta}}{\theta\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)}, \tag{A.16}$$

which is the expression for the optimal marginal taxation of labor income found in Equation (17). Using similar steps, $\partial\mathcal{L}/\partial E(\theta)$, and defining $\Delta\Pi(\theta) \equiv G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta) - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\tilde{f}(\theta)$ we find that:

$$\frac{T'_\pi(\theta)}{1-T'_\pi(\theta)} = \Big(1+\frac{1}{\epsilon_\pi(\theta)}\Big)\frac{\int_\theta^{\bar{\theta}}\Big[\Delta\Pi(\hat{\theta})-g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Delta T(\hat{\theta})\Big]d\hat{\theta}}{\theta G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)}, \tag{A.17}$$

which gives us Equation 16.

From $\partial\mathcal{L}/\partial b = 0$ we find:

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial b} =& \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial v^W(\theta)}{\partial b} d\tilde{F}(\theta) - \int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta)\frac{\partial v^W(\theta)}{\partial b} d\theta \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big)d\theta \\
&+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big)\frac{\partial v^W(\theta)}{\partial b}\Big(L(\theta) + H\big(s(\theta)\big)l(\theta)\Big)dF(\theta) \\
&+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)l(\theta)dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial v^W(\theta)}{\partial b}dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big)\frac{\partial v^W(\theta)}{\partial b}\Big(Y\big(L(\theta), E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big) \\
&\qquad + v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta\Big)dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta\Big)dF(\theta) \\
&+ \gamma = 0.
\end{aligned} \tag{A.18}$$

Using that $\partial\mathcal{L}/\partial v^W(\theta) = 0\, \forall\theta \in [\underline{\theta}, \bar{\theta}]$ in optimum we find:

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial b} =& - \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big)d\theta \\
&+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)l(\theta)dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta\Big)dF(\theta) \\
&+ \gamma = 0.
\end{aligned} \tag{A.19}$$

If we combine this with the previous findings that $\lambda_R = 1$ and $\lambda_L = w$ we get:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\Big(wl(\theta) - \psi\big(\frac{l(\theta)}{\theta}\big)\Big) - \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta\Big)dF(\theta).
\end{aligned} \tag{A.20}$$

Combining this with the private first-order condition $H'_s\big(s(\theta)\big)\big(wl(\theta) - T_w\big(wl(\theta)\big) - \psi\big(\frac{l(\theta)}{\theta}\big) - b\big) - \psi'\big(\frac{s(\theta)}{\theta}\big)/\theta = 0$ we can rewrite this expression and find:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H'_s\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H'_s\big(s(\theta)\big)\Big(T\big(wl(\theta)\big) + b\Big)dF(\theta).
\end{aligned} \tag{A.21}$$

Solving for b we get:

$$\begin{aligned}
b =& \frac{\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H'_s\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi'\big(\frac{s(\theta)}{\theta}\big)/\theta^2 + \varphi''\big(\frac{s(\theta)}{\theta}\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H'_s\big(s(\theta)\big)dF(\theta)} \\
&- \frac{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H'_s\big(s(\theta)\big)T\big(wl(\theta)\big)dF(\theta)}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H'_s\big(s(\theta)\big)dF(\theta)}.
\end{aligned} \tag{A.22}$$

By using the implicit function theorem on the employees private first-order condition we see that $\partial s(\theta)/\partial b < 0$. We find that $\partial s(\theta)/\partial b = H'_s\big(s(\theta)\big)/\Big(H''_s\big(s(\theta)\big)\big(wl(\theta) - T_w\big(wl(\theta)\big) - \psi\big(\frac{l(\theta)}{\theta}\big) - b\big) - \psi''\big(\frac{s(\theta)}{\theta}\big)/\theta^2\Big)$. Since we assume that $H'_s\big(s(\theta)\big) > 0$, $\psi''\big(\frac{s(\theta)}{\theta}\big)/\theta^2 > 0$ and $H''_s\big(s(\theta)\big) \leq 0$ this implies that $\partial s(\theta)/\partial b < 0$.

Appendix B:

When we include an externality from unemployment in the search cost function, the objective function of the social planner becomes:

$$
\begin{aligned}
&\max_{E(\theta),L(\theta),l(\theta),v^E(\theta),v^W(\theta),\tilde{\phi}(\theta),b,u} \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big[\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)v^E(\theta) + \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)v^W(\theta)\Big] d\tilde{F}(\theta) \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)} \phi d\tilde{G}_\theta(\phi) d\tilde{F}(\theta) \\
&s.t. \quad v_\theta^{E\prime}(\theta) = \psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2} \quad \forall\theta \in [\underline{\theta},\bar{\theta}], \\
&\qquad v_\theta^{W\prime}(\theta) = H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta},u\big)\frac{s(\theta)}{\theta^2} \quad \forall\theta \in [\underline{\theta},\bar{\theta}], \\
&\int_{\underline{\theta}}^{\bar{\theta}} \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta)dF(\theta) - \int_{\underline{\theta}}^{\bar{\theta}} G_\theta\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta) \geq 0, \\
&\int_{\underline{\theta}}^{\bar{\theta}} \Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta),E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)\Big) \\
&- \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta},u\big)\Big)\Big] dF(\theta) \geq 0, \\
&b \geq 0, \\
&u = \int_{\underline{\theta}}^{\bar{\theta}} \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(1-H\big(s(\theta)\big)\Big)dF(\theta).
\end{aligned}
\tag{B.1}
$$

After integrating by parts we get the following Lagrangian,:

$$
\begin{aligned}
\mathcal{L} &= \int_{\underline{\theta}}^{\bar{\theta}} \Big[\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)v^E(\theta) - \int_{\underline{\phi}_\theta}^{\tilde{\phi}(\theta)} \phi d\tilde{G}_\theta(\phi) + \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)v^W(\theta)\Big] d\tilde{F}(\theta) \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \Big[\mu_\theta^{E\prime}(\theta)v^E(\theta) + \mu^E(\theta)\psi'\big(\frac{E(\theta)}{\theta}\big)\frac{E(\theta)}{\theta^2}\Big] d\theta \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \Big[\mu_\theta^{W\prime}(\theta)v^W(\theta) + \mu^W(\theta)\Big(H\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta},u\big)\frac{s(\theta)}{\theta^2}\Big)\Big] d\theta \\
&+ \lambda_L\Big[\int_{\underline{\theta}}^{\bar{\theta}} \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)H\big(s(\theta)\big)l(\theta) - G_\theta\big(\tilde{\phi}(\theta)\big)L(\theta)dF(\theta)\Big] \\
&+ \lambda_R\Big[\int_{\underline{\theta}}^{\bar{\theta}} \Big[G_\theta\big(\tilde{\phi}(\theta)\big)\Big(Y\big(L(\theta),E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)\Big) \\
&\qquad - \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta},u\big)\Big)\Big] dF(\theta)\Big] \\
&+ \gamma b + \kappa\Big[u - \int_{\underline{\theta}}^{\bar{\theta}} \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big(1-H\big(s(\theta)\big)\Big)dF(\theta)\Big].
\end{aligned}
\tag{B.2}
$$

From $\partial\mathcal{L}/\partial L(\theta) = 0$ we find:

$$Y'_L\big(L(\theta), E(\theta)\big) = \frac{\lambda_L}{\lambda_R} \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]. \tag{B.3}$$

Since the production function is assumed to have a constant returns to scale, this implies that the marginal return to labor and effort must be equalized. This implies that:

$$\begin{aligned} &Y'_L\big(L(\theta), E(\theta)\big) = w \quad \forall \theta \in [\underline{\theta}, \bar{\theta}], \\ &\text{and} \\ &Y'_E\big(L(\theta), E(\theta)\big) = \tilde{w} \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]. \end{aligned} \tag{B.4}$$

From $\partial\mathcal{L}/\partial v^W(\theta) = 0$ we find:

$$\begin{aligned} &\mu_\theta^{W\prime}(\theta) = \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) \\ &+ \lambda_L g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(H\big(s(\theta)\big)l(\theta) + L(\theta)\Big) - \kappa g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(1 - H\big(s(\theta)\big)\Big) \\ &- \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(Y\big(L(\theta), E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big) + v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}, u\big)\Big). \end{aligned} \tag{B.5}$$

Using the constant returns to scale of the production function, $\lambda_L = w\lambda_R$, $c^E(\theta) = v^E(\theta) + \psi\big(\frac{E(\theta)}{\theta}\big)$, and $c^W(\theta) = v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}\big)$, we get:

$$\begin{aligned} \mu_\theta^{W\prime}(\theta) =& \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \kappa g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(1 - H\big(s(\theta)\big)\Big) \\ &- \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\bigg(\tilde{w}E(\theta) - c^E(\theta) - \Big(wH\big(s(\theta)\big)l(\theta) - c^W(\theta)\Big)\bigg). \end{aligned} \tag{B.6}$$

We can define the tax wedge between entrepreneurs and employees as $\Delta T(\theta) \equiv \tilde{w}E(\theta) - c^E(\theta) - \big(wH\big(s(\theta)\big)l(\theta) - c^W(\theta)\big)$. Using this we get:

$$\begin{aligned} \mu_\theta^{W\prime}(\theta) =& \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \lambda_R\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \kappa g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(1 - H\big(s(\theta)\big)\Big) \\ &- \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\theta). \end{aligned} \tag{B.7}$$

Using $\partial\mathcal{L}/\partial v^E(\theta) = 0$ and similar steps, we find:

$$\begin{aligned} \mu_\theta^{E\prime}(\theta) =& \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\tilde{f}(\theta) - \lambda_R G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta) + \kappa g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(1 - H\big(s(\theta)\big)\Big) \\ &+ \lambda_R g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Delta T(\theta). \end{aligned} \tag{B.8}$$

By the transversality conditions we know that $\int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{E\prime}(\theta)d\theta = 0$. By adding $\int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta) + \mu_\theta^{E\prime}(\theta)d\theta = 0$ we find that $\lambda_R = 1$. Using this we can rewrite the expression for $\mu_\theta^{W\prime}(\theta)$ to find:

$$\mu_\theta^{W\prime}(\theta) = \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta) - \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - g_\theta\big(\tilde{\phi}(\theta)\big)f(\theta)\Big(\kappa\Big(1 - H\big(s(\theta)\big)\Big) + \Delta T(\theta)\Big). \tag{B.9}$$

By integrating this expression we find:

$$\begin{aligned}\mu^W(\bar{\theta}) = \mu(\theta) + \int_{\theta}^{\bar{\theta}} &\Big[\Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)\tilde{f}(\hat{\theta}) - \Big(1 - G_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)f(\hat{\theta}) \\ &- g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Big(\kappa\Big(1 - H\big(s(\hat{\theta})\big)\Big) + \Delta T(\hat{\theta})\Big)\Big]d\hat{\theta}.\end{aligned} \tag{B.10}$$

By using the transversality condition $\mu\big(\bar{\theta}\big) = 0$ we find that:

$$\begin{aligned}\mu^W(\theta) = \int_{\theta}^{\bar{\theta}} &\Big[\Big(1 - G_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)f(\hat{\theta}) - \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\Big)\tilde{f}(\hat{\theta}) \\ &+ g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Big(\kappa\Big(1 - H\big(s(\hat{\theta})\big)\Big) + \Delta T(\hat{\theta})\Big)\Big]d\hat{\theta}.\end{aligned} \tag{B.11}$$

Using the same steps we find the following from $\partial\mathcal{L}/\partial v^E(\theta) = 0$:

$$\begin{aligned}\mu^E(\theta) = \int_{\theta}^{\bar{\theta}} &\Big[G_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta}) - \tilde{G}_\theta\big(\tilde{\phi}(\hat{\theta})\big)\tilde{f}(\hat{\theta}) \\ &- g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Big(\kappa\Big(1 - H\big(s(\hat{\theta}), u\big)\Big) + \Delta T(\hat{\theta})\Big)\Big]d\hat{\theta}.\end{aligned} \tag{B.12}$$

From $\partial\mathcal{L}/\partial l(\theta) = 0$ we can use that $\lambda_L = w\lambda_R$ and $\lambda_R = 1$ to find:

$$\mu^W(\theta)H\big(s(\theta)\big)\Big(\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta^2 + \psi''\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^3}\Big) = \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)H\big(s(\theta)\big)\big(w - \psi'\big(\frac{l(\theta)}{\theta}\big)/\theta\big). \tag{B.13}$$

This can be rewritten as:

$$\frac{\mu^W(\theta)\Big(\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta + \psi''\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2}\Big)}{\theta\Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)} = w - \psi'\big(\frac{l(\theta)}{\theta}\big)/\theta. \tag{B.14}$$

Using the private first-order condition of the employee $H\big(s(\theta)\big)\Big(w\big(1-T_w'(\theta)\big)-\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta\Big)=0$ and the implicit function theorem, we can define the elasticity of labor supply with respect to after tax wages:

$$
\begin{aligned}
\epsilon_w(\theta) &= \frac{\partial l(\theta)}{\partial w\big(1-T_w'(\theta)\big)}\frac{w\big(1-T_w'(\theta)\big)}{l(\theta)} \\
&= -\frac{H\big(s(\theta)\big)}{-H\big(s(\theta)\big)\psi''\big(\frac{l(\theta)}{\theta}\big)/\theta^2}\frac{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{l(\theta)} \\
&= \frac{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{l(\theta)\psi''\big(\frac{l(\theta)}{\theta}\big)/\theta^2}.
\end{aligned}
\tag{B.15}
$$

From this we find:

$$
\begin{aligned}
\frac{\mu^W(\theta)\big(1+1/\epsilon_w(\theta)\big)}{\theta\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)} &= \frac{w-\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta}{\psi'\big(\frac{l(\theta)}{\theta}\big)/\theta} \\
&= \frac{w-w\big(1-T_w'(\theta)\big)}{w\big(1-T_w'(\theta)\big)} \\
&= \frac{T_w'}{1-T_w'}.
\end{aligned}
\tag{B.16}
$$

Combining this with Equation (B.11), and defining $\Delta\Omega(\theta) \equiv \Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta) - \Big(1-\tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big)\tilde{f}(\theta)$, we find the following rule for the optimal marginal taxation of wage income:

$$
\begin{aligned}
\frac{T_w'}{1-T_w'} &= \big(1+\frac{1}{\epsilon_w(\theta)}\big)\times \\
&\frac{\int_\theta^{\bar{\theta}}\Big[\Delta\Omega(\hat{\theta})+g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Big(\kappa\Big(1-H\big(s(\hat{\theta})\big)\Big)+\Delta T(\hat{\theta})\Big)\Big]d\hat{\theta}}{\theta\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)}.
\end{aligned}
\tag{B.17}
$$

Using similar calculations for $\partial\mathcal{L}/\partial E(\theta)$, defining $\epsilon_\pi(\theta)$ as the marginal elasticity of effort with respect to after-tax profits and $\Delta\Pi(\theta) \equiv G_\theta\big(\tilde{\phi}(\theta)\big)f(\theta) - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\tilde{f}(\theta)$, and using Equation (B.12), we find the following tax rule for firm profits:

$$
\frac{T_\pi'}{1-T_\pi'} = \big(1+\frac{1}{\epsilon_\pi(\theta)}\big)\frac{\int_\theta^{\bar{\theta}}\Big[\Delta\Pi(\hat{\theta})-g_\theta\big(\tilde{\phi}(\hat{\theta})\big)f(\hat{\theta})\Big(\kappa\Big(1-H\big(s(\hat{\theta})\big)\Big)+\Delta T(\hat{\theta})\Big)\Big]d\hat{\theta}}{\theta\Big(1-G_\theta\big(\tilde{\phi}(\theta)\big)\Big)f(\theta)},
\tag{B.18}
$$

From $\partial\mathcal{L}/\partial b = 0$ we find:

$$\frac{\partial\mathcal{L}}{\partial b} = \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial v^W(\theta)}{\partial b} d\tilde{F}(\theta) - \int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta)\frac{\partial v^W(\theta)}{\partial b} d\theta$$
$$- \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big) d\theta$$
$$+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big)\frac{\partial v^W(\theta)}{\partial b}\Big(L(\theta) + H\big(s(\theta)\big)l(\theta)\Big) dF(\theta)$$
$$+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big)l(\theta) dF(\theta)$$
$$- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial v^W(\theta)}{\partial b} dF(\theta)$$
$$- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big)\frac{\partial v^W(\theta)}{\partial b}\Big(Y\big(L(\theta), E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big)$$
$$+ v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}, u\big)\Big) dF(\theta)$$
$$- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta\Big) dF(\theta)$$
$$- \kappa \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big)\frac{\partial v^W(\theta)}{\partial b}\Big(1 - H\big(s(\theta)\big)\Big)$$
$$+ \kappa \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big) dF(\theta)$$
$$+ \gamma = 0. \tag{B.19}$$

Using that $\int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial\mathcal{L}}{\partial v^W(\theta)}\frac{\partial v^W(\theta)}{\partial b} d\theta = 0$ in optimum we find:

$$\frac{\partial\mathcal{L}}{\partial b} = - \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big) d\theta$$
$$+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big)l(\theta) dF(\theta)$$
$$- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta\Big) dF(\theta)$$
$$+ \kappa \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b} H_s'\big(s(\theta)\big) dF(\theta) + \gamma = 0. \tag{B.20}$$

If we combine this with the previous findings that $\lambda_R = 1$ and $\lambda_L = w$ we get:

$$\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma$$
$$= \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\Big(wl(\theta) - \psi\big(\frac{l(\theta)}{\theta}\big)\Big) - \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta\Big)dF(\theta)$$
$$+ \kappa\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)dF(\theta). \tag{B.21}$$

Combining this with the private first-order condition $H_s'\big(s(\theta)\big)\big(wl(\theta) - T_w\big(wl(\theta)\big) - \psi\big(\frac{l(\theta)}{\theta}\big) - b\big) - \psi'\big(\frac{s(\theta)}{\theta}, u\big)/\theta = 0$, we can rewrite this expression and find:

$$\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma$$
$$- \kappa\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)dF(\theta)$$
$$= \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)\Big(T\big(wl(\theta)\big) + b\Big)dF(\theta). \tag{B.22}$$

Solving for b we get:

$$b = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\frac{\partial s(\theta)}{\partial b}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big)d\theta - \gamma}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)dF(\theta)}$$
$$- \frac{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)T\big(wl(\theta)\big)dF(\theta)}{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial b}H_s'\big(s(\theta)\big)dF(\theta)} - \kappa. \tag{B.23}$$

By using the implicit function theorem on the employees private first-order condition we find that $\partial s(\theta)/\partial b < 0$.[17] We can use $\partial \mathcal{L}/\partial u$ to solve for κ:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial u} =& \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - \tilde{G}_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial v^W(\theta)}{\partial u} d\tilde{F}(\theta) \\
& - \int_{\underline{\theta}}^{\bar{\theta}} \mu_\theta^{W\prime}(\theta) \frac{\partial v^W(\theta)}{\partial u} d\theta \\
& - \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta) \Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big) \frac{\partial s(\theta)}{\partial u} d\theta \\
& - \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta) \varphi_{s,u}''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^2} d\theta \\
& + \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big) \frac{\partial v^W(\theta)}{\partial u} \Big(H\big(s(\theta)\big)l(\theta) + L(\theta)\Big) dF(\theta) \\
& + \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big)l(\theta) dF(\theta) \\
& - \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial v^W(\theta)}{\partial u} dF(\theta) \\
& - \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big) \frac{\partial v^W(\theta)}{\partial u} \\
& \times \Big(Y\big(L(\theta), E(\theta)\big) - v^E(\theta) - \psi\big(\frac{E(\theta)}{\theta}\big) + v^W(\theta) + H\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi\big(\frac{s(\theta)}{\theta}, u\big)\Big) dF(\theta) \\
& - \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} \Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta\Big) dF(\theta) \\
& - \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \varphi_u'\big(\frac{s(\theta)}{\theta}, u\big) dF(\theta) \\
& - \kappa \int_{\underline{\theta}}^{\bar{\theta}} g_\theta\big(\tilde{\phi}(\theta)\big) \frac{\partial v^W(\theta)}{\partial u} \Big(1 - H\big(s(\theta)\big)\Big) dF(\theta) \\
& + \kappa \Big(1 + \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) dF(\theta)\Big) = 0.
\end{aligned}
\tag{B.24}
$$

[17]We find that $\partial s(\theta)/\partial b = H_s'\big(s(\theta)\big)/\Big(H_s''\big(s(\theta)\big)\big(wl(\theta) - T_w\big(wl(\theta)\big) - \psi\big(\frac{l(\theta)}{\theta}\big) - b\big) - \psi''\big(\frac{s(\theta)}{\theta}\big)/\theta^2\Big)$. Since we assume that $H_s'\big(s(\theta)\big) > 0$, $\psi''\big(\frac{s(\theta)}{\theta}\big)/\theta^2 > 0$ and $H_s''\big(s(\theta)\big) \le 0$ this implies that $\partial s(\theta)/\partial b < 0$.

Using that $\partial\mathcal{L}/\partial v^W(\theta) = 0\,\forall\theta \in [\underline{\theta}, \bar{\theta}]$ in optimum we get:

$$\begin{aligned}
0 = &- \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta) \frac{\partial s(\theta)}{\partial u} \Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta^2 + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big) d\theta \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\varphi_{s,u}''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^2} d\theta \\
&+ \lambda_L \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) l(\theta) dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} \Big(H_s'\big(s(\theta)\big)\psi\big(\frac{l(\theta)}{\theta}\big) + \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta \Big) dF(\theta) \\
&- \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \varphi_u'\big(\frac{s(\theta)}{\theta}, u\big) dF(\theta) \\
&+ \kappa\Big(1 + \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big) \frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) dF(\theta)\Big),
\end{aligned} \tag{B.25}$$

which can be simplified using $\lambda_R = 1$ and $\lambda_L = w$:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\Big[\varphi_{s,u}''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^2} + \frac{\partial s(\theta)}{\partial u}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \frac{\varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)}{\theta^2} + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big)\Big] d\theta \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big[\varphi_u'\big(\frac{s(\theta)}{\theta}, u\big) + \frac{\partial s(\theta)}{\partial u}\Big(H_s'\big(s(\theta)\big)\Big(wl(\theta) - \psi\big(\frac{l(\theta)}{\theta}\big)\Big) - \frac{\varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)}{\theta}\Big)\Big] dF(\theta) \\
&+ \kappa\Big(1 + \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) dF(\theta)\Big) = 0.
\end{aligned} \tag{B.26}$$

Solving for κ and by using the private fist order condition $H_s'\big(s(\theta)\big)\Big(wl(\theta) - T_w\big(wl(\theta)\big) - \psi\big(\frac{l(\theta)}{\theta}\big) - b\Big) = \varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)/\theta$ we get:

$$\begin{aligned}
\kappa = &\frac{\int_{\underline{\theta}}^{\bar{\theta}} \mu^W(\theta)\Big[\varphi_{s,u}''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^2} + \frac{\partial s(\theta)}{\partial u}\Big(H_s'\big(s(\theta)\big)\psi'\big(\frac{l(\theta)}{\theta}\big)\frac{l(\theta)}{\theta^2} + \frac{\varphi_s'\big(\frac{s(\theta)}{\theta}, u\big)}{\theta^2} + \varphi_s''\big(\frac{s(\theta)}{\theta}, u\big)\frac{s(\theta)}{\theta^3}\Big)\Big] d\theta}{1 + \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) dF(\theta)} \\
&- \frac{\int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\Big[\varphi_u'\big(\frac{s(\theta)}{\theta}, u\big) + \frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big)\Big(T_w\big(wl(\theta)\big) + b\Big)\Big] dF(\theta)}{1 + \int_{\underline{\theta}}^{\bar{\theta}} \Big(1 - G_\theta\big(\tilde{\phi}(\theta)\big)\Big)\frac{\partial s(\theta)}{\partial u} H_s'\big(s(\theta)\big) dF(\theta)}.
\end{aligned} \tag{B.27}$$