A Proofs and Technical Details

A.1 Proofs of Section 1

Concavity of the Utility of Earnings v. Our analysis requires that the utility of pre-tax earnings $z \mapsto v(z) \equiv u(R(z))$ is concave on $\mathbb{R}_+^2$. It is easy to show that this is equivalent to $p_1(z)p_2(z) > -\gamma(z)$ where $\gamma(z) \equiv -R(z)u''(R(z))/u'(R(z))$ is the agent's coefficient of relative risk aversion, and $p_1(z) \equiv (1 - T(z)/z)R'(z)$, $p_2(z) \equiv zT''(z)/R'(z)$ are two measures of the local rate of progressivity of the tax schedule. If the tax schedule is CRP with parameter p, these variables are respectively equal to $\frac{1}{1-p}$ and p. If $u(c) = \log c$, then $\gamma(z) = -1$ and the above restriction is always satisfied regardless of the value of p. ■

Lemma 3 (First-Order Approach.) *Suppose that the equilibrium effort level is interior. The firm's problem is equivalent to maximizing (1) subject to (3) and the local incentive constraint:*

$$h'(\ell) \;=\; v(\bar{z}) - v(\underline{z}). \tag{18}$$

Proof of Lemma 3. Denote the agent's expected utility of effort ℓ by

$$V(\ell) \equiv (1-\ell)u(R(\underline{z}(\theta))) + \ell u(R(\bar{z}(\theta))) - h(\ell).$$

The first-order condition reads $V'(\ell) = 0$, where

$$V'(\ell) = u(R(\bar{z}(\theta))) - u(R(\underline{z}(\theta))) - h'(\ell).$$

We then have

$$V''(\ell) = -h''(\ell) < 0,$$

where the inequality follows from the convexity of the disutility of effort. Thus, the agent's problem is concave and, as long as the effort choice is interior, the first-order condition is necessary and sufficient. ■

Lemma 4 (Optimal Contract: General Preferences and Taxes.) *The base pay*

$\underline{z}(\theta)$ and high-level pay $\bar{z}(\theta)$ satisfy

$$v(\underline{z}) - h(\ell) \;=\; U(\theta) - \ell h'(\ell) \tag{19}$$

$$v(\bar{z}) - h(\ell) \;=\; U(\theta) + (1-\ell)h'(\ell). \tag{20}$$

The effort level $\ell(\theta)$ exerted by the worker satisfies

$$\theta \;=\; b + \left[\frac{1}{v'(\bar{z})} - \frac{1}{v'(\underline{z})}\right]\ell(1-\ell)h''(\ell). \tag{21}$$

Expected utility $U(\theta)$ satisfies $\underline{z} + \ell b = \theta\ell$.

Proof of Lemma 4. The participation constraint reads:

$$(1-\ell)v(\underline{z}) + \ell v(\bar{z}) - h(\ell) = U(\theta),$$

and the local incentive constraint reads:

$$v(\bar{z}) - v(\underline{z}) = h'(\ell).$$

Solving this linear system of equations for $v(\underline{z})$ and $v(\bar{z})$ as functions of ℓ and $U(\theta)$ immediately delivers equations (19) and (20). The optimal effort level $\ell(\theta)$ maximizes the firm's profit $\ell\theta - (\underline{z} + \ell b)$ subject to the participation and incentive constraints, taking the reservation value $U(\theta)$ as given. The first-order condition reads:

$$\theta = b + \frac{\partial \underline{z}}{\partial \ell} + \ell\frac{\partial b}{\partial \ell}.$$

Applying the implicit function theorem to equations (19) and (20) leads to

$$v'(\underline{z})\frac{\partial \underline{z}}{\partial \ell} = -\ell h''(\ell)$$

and

$$v'(\bar{z})\left(\frac{\partial \underline{z}}{\partial \ell} + \frac{\partial b}{\partial \ell}\right) = (1-\ell)h''(\ell).$$

Solving for $\frac{\partial \underline{z}}{\partial \ell}, \frac{\partial b}{\partial \ell}$ and substituting into the first-order condition yields

$$\theta = b - (1-\ell)\frac{1}{v'(\underline{z})}\ell h''(\ell) + \frac{1}{v'(\bar{z})}\ell(1-\ell)h''(\ell).$$

Rearranging terms leads to equation (21). Finally, the zero-profit condition $\underline{z}+\ell b = \ell\theta$ pins down the equilibrium reservation utility $U(\theta)$. ■

Proof of Lemma 1. Defining β by $\bar{z} = e^{\beta}\underline{z}$, the free-entry condition (4) can be expressed as

$$(1-\ell)\underline{z} + \ell e^{\beta}\underline{z} = \theta\ell.$$

This immediately leads to the solution (6) for the equilibrium values of the base pay $\underline{z}$ and high-performance pay $\bar{z}$. From a firm's viewpoint, the participation constraint (3) determines the base pay $\underline{z}$ given ℓ and the reservation value U. With log utility and a CRP tax schedule, it reads:

$$\log\frac{1-\tau}{1-p} + (1-p)[(1-\ell)\log(\underline{z}) + \ell\log(\bar{z})] - h(\ell) = U.$$

The pass-through $\beta \equiv \log(\bar{z}) - \log(\underline{z})$ is given as a function of the desired effort level ℓ by the incentive constraint (7). Substituting $\beta = h'(\ell)/(1-p)$ into the previous equation, we get:

$$\log\frac{1-\tau}{1-p} + (1-p)\left[\log(\underline{z}) + \ell\frac{h'(\ell)}{1-p}\right] - h(\ell) = U.$$

Solving for $\underline{z}$ leads to:

$$\underline{z} = \left(\frac{1-\tau}{1-p}\right)^{-1/(1-p)} e^{\frac{1}{1-p}U} e^{\frac{1}{1-p}[h(\ell)-\ell h'(\ell)]} \tag{22}$$

and hence

$$\bar{z} = \left(\frac{1-\tau}{1-p}\right)^{-1/(1-p)} e^{\frac{1}{1-p}U} e^{\frac{1}{1-p}[h(\ell)+(1-\ell)h'(\ell)]}.$$

Substituting these expressions into the free-entry condition determines the reservation value U as a function of labor effort:

$$e^{\frac{1}{1-p}U} = \left(\frac{1-\tau}{1-p}\right)^{1/(1-p)} e^{-\frac{1}{1-p}[h(\ell)-\ell h'(\ell)]} \frac{\theta\ell}{1-\ell+\ell e^{\frac{1}{1-p}h'(\ell)}},$$

i.e.,

$$U = \log\left(\frac{1-\tau}{1-p}\right) + (1-p)\log(\theta\ell) - h(\ell) + (1-p)[\ell\beta - \log(1-\ell+\ell e^{\beta})].$$

Noting that $u(\mathbb{E}[z|\theta]) = \log(\theta\ell)$ and $\mathbb{E}[u(z)|\theta] = (1-\ell)\log\underline{z} + \ell\log(e^{\beta}\underline{z}) = \log\underline{z} + \ell\beta$, and using expression (6), we obtain:

$$U = v(\theta\ell) - h(\ell) + (1-p)\{\mathbb{E}[u(z)|\theta] - u(\mathbb{E}[z|\theta])\}.$$

Finally, the first-order condition for effort is obtained by differentiating the firm's expected profit $\theta\ell - [\underline{z} + \ell b]$ and equating it to zero:

$$\theta = b + \frac{\partial\underline{z}}{\partial\ell} + \ell\frac{\partial b}{\partial\ell} = b + \frac{1}{1-p}\ell h''(\ell)\underline{z}e^{\frac{1}{1-p}h'(\ell)} + [1-\ell+\ell e^{\frac{1}{1-p}h'(\ell)}]\frac{\partial\underline{z}}{\partial\ell},$$

where the second equality follows from the fact that $b = (e^{\beta}-1)\underline{z}$ by definition, with $\beta = \frac{1}{1-p}h'(\ell)$ since the contract must respect the incentive constraint (7). Since the firm takes as given the worker's reservation utility U, we differentiate equation (22) to obtain:

$$\frac{\partial\underline{z}}{\partial\ell} = -\underline{z}\frac{1}{1-p}\ell h''(\ell).$$

Substituting into the previous expression gives

$$\theta = b + [e^{\frac{1}{1-p}h'(\ell)} - 1]\underline{z}\frac{1}{1-p}\ell(1-\ell)h''(\ell),$$

which leads to equation (8) for the CRP environment:

$$\theta \;=\; b + b\,\ell(1-\ell)\frac{h''(\ell)}{1-p}. \tag{23}$$

Note that in equilibrium, we can use (6) to rewrite this equation as

$$1 = \frac{\ell(e^{\beta}-1)}{1+\ell(e^{\beta}-1)}\left[1 + \frac{1}{1-p}\ell(1-\ell)h''(\ell)\right],$$

and hence

$$\ell(1-\ell) = \frac{1}{\beta(e^{\beta}-1)}\frac{h'(\ell)}{\ell h''(\ell)}.$$

This expression implies that the optimal effort level ℓ is independent of the worker's productivity θ. ■

A.2 Proofs of Section 2

Proof of Lemma 2. The first order condition (23) for labor effort can be rewritten as $1-p=\ell^2(1-\ell)h''(\ell)[e^{h'(\ell)/(1-p)}-1]$. Apply the implicit function theorem to get:

$$\varepsilon_{\ell,1-p}\equiv\frac{1-p}{\ell}\frac{\partial\ell}{\partial(1-p)}=\frac{1+\frac{\beta e^{\beta}}{e^{\beta}-1}}{\frac{2-3\ell}{1-\ell}+\frac{\beta e^{\beta}}{e^{\beta}-1}\frac{\ell h''(\ell)}{h'(\ell)}+\frac{\ell h'''(\ell)}{h''(\ell)}}.$$

Recall that

$$\frac{\partial\Pi}{\partial\ell}=\theta-b\Big[1+\frac{\ell(1-\ell)h''(\ell)}{1-p}\Big].$$

Differentiating this expression using $\frac{\partial z}{\partial\ell}=-\ell\underline{z}\frac{h''(\ell)}{1-p}$ and $\frac{\partial b}{\partial\ell}=[\underline{z}+(1-\ell)b]\frac{h''(\ell)}{1-p}$ leads to

$$\frac{\partial^2\Pi}{\partial\ell^2}=-[\underline{z}+(2-3\ell)b]\frac{h''(\ell)}{1-p}-\ell(1-\ell)[\underline{z}+(1-\ell)b]\Big(\frac{h''(\ell)}{1-p}\Big)^2-b\ell(1-\ell)\frac{h'''(\ell)}{1-p}.$$

The second-order condition for optimal labor effort, $\frac{\partial^2\Pi}{\partial\ell^2}\leq 0$, can therefore be expressed as

$$\Big[\ell\beta\frac{\ell h''(\ell)}{h'(\ell)}-\frac{1}{1-\ell}\frac{1}{e^{\beta}-1}\Big]-\Big[\frac{2-3\ell}{1-\ell}+\frac{\beta e^{\beta}}{e^{\beta}-1}\frac{\ell h''(\ell)}{h'(\ell)}+\frac{\ell h'''(\ell)}{h''(\ell)}\Big]\leq 0,$$

where we used the fact that $\underline{z}/b=1/(e^{\beta}-1)$. The first-order condition for labor effort implies that the first square bracket is equal to zero. Therefore, we obtain $\varepsilon_{\ell,1-p}>0$. Now, suppose that the disutility of effort is isoelastic, $h(\ell)=\frac{\ell^{1+1/\varepsilon_\ell^F}}{1+1/\varepsilon_\ell^F}$ with ε_ℓ^F constant. We can then rewrite the labor effort elasticity as

$$\varepsilon_{\ell,1-p}=\frac{\varepsilon_\ell^F}{1+\frac{1-\ell/(1-\ell)}{1+\beta e^{\beta}/(e^{\beta}-1)}\varepsilon_\ell^F}.$$

This expression implies that $\varepsilon_{\ell,1-p}>\frac{\varepsilon_\ell^F}{1+\varepsilon_\ell^F}$, and that $\varepsilon_{\ell,1-p}>\varepsilon_\ell^F$ if and only if $1-\frac{\ell}{1-\ell}<0$, i.e., $\ell>\frac{1}{2}$. ■

Proof of Proposition 1. Equation (10) follows immediately from equation (7) and Lemma 2. ■

A.3 Proofs of Section 3.1

Let $\mathcal{U}(\theta) \equiv v(\theta\ell) - h(\ell) = \log \frac{1-\tau}{1-p} + (1-p)\log(\ell\theta) - h(\ell)$ be the worker's expected utility in the full-insurance setting. By the envelope theorem, we have $\frac{d\mathcal{U}(\theta)}{dp} = \frac{1}{1-p} - \log(\mathbb{E}z)$. The next lemma computes the impact of tax progressivity on the equilibrium reservation utility in our framework.

Lemma 5 *The impact of tax progressivity on expected utility is given by*

$$\frac{dU(\theta)}{dp} = \frac{d\mathcal{U}(\theta)}{dp} + (\log(\mathbb{E}[z|\theta]) - \mathbb{E}[\log z|\theta]) + \frac{b}{\beta\mathbb{E}[z|\theta]}\mathbb{V}(\log z|\theta)\ \varepsilon_{\beta,1-p}, \tag{24}$$

where $\mathbb{V}(\log z|\theta) = \beta^2\ell(1-\ell)$ *is the variance of log-earnings conditional on ability* θ*.*

Proof of Lemma 5. Differentiating equation (9) gives

$$\frac{\partial U(\theta)}{\partial(1-p)} = \left[-\frac{1}{1-p} + \log(\ell\theta) + \left(\frac{1-p}{\ell} - h'(\ell)\right)\frac{\partial \ell}{\partial(1-p)}\right] + (\log(\mathbb{E}[z|\theta]) - \mathbb{E}[\log z|\theta])$$
$$- (1-p)\left\{\left[\frac{e^\beta - 1}{1+\ell(e^\beta-1)} - \beta\right]\frac{\partial \ell}{\partial(1-p)} + \frac{(e^\beta-1)\ell(1-\ell)}{1+\ell(e^\beta-1)}\frac{\partial\beta}{\partial(1-p)}\right\}.$$

But recall that

$$\frac{d\mathcal{U}(\theta)}{d(1-p)} = -\frac{1}{1-p} + \log(\ell\theta)$$

and that

$$\frac{\partial\beta}{\partial(1-p)} = \frac{\beta}{1-p}[\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p} + \varepsilon_{\beta,1-p}] = \frac{\partial\beta}{\partial\ell}\frac{\partial\ell}{\partial(1-p)} + \frac{\beta}{1-p}\varepsilon_{\beta,1-p}.$$

Substituting into the previous equation and using $\frac{\partial\beta}{\partial\ell} = \frac{\beta}{\ell}\frac{\ell h''(\ell)}{h'(\ell)}$ and $\frac{e^\beta-1}{1+\ell(e^\beta-1)} = \frac{b}{\mathbb{E}[z|\theta]}$ leads to

$$\frac{\partial U(\theta)}{\partial(1-p)} = \frac{d\mathcal{U}(\theta)}{d(1-p)} + (\log(\mathbb{E}[z|\theta]) - \mathbb{E}[\log z|\theta]) - \frac{b}{\mathbb{E}[z|\theta]}\beta\ell(1-\ell)\varepsilon_{\beta,1-p}$$
$$+ \frac{1-p}{\ell}\left[1 - \frac{\ell h'(\ell)}{1-p} - \frac{\ell(e^\beta-1)}{1+\ell(e^\beta-1)} + \beta\ell - \frac{\beta(e^\beta-1)\ell(1-\ell)}{1+\ell(e^\beta-1)}\frac{\ell h''(\ell)}{h'(\ell)}\right]\frac{\partial\ell}{\partial(1-p)}.$$

Using the first-order condition for labor effort $\beta(e^{\beta}-1)\ell(1-\ell)\frac{\ell h''(\ell)}{h'(\ell)}=1$ derived in the proof of Lemma 1, we obtain that the term in square brackets that multiplies $\frac{\partial \ell}{\partial(1-p)}$ is equal to zero; this is a manifestation of the envelope theorem in our setting. This easily yields expression (24). ■

The interpretation of equation (24) is that a higher rate of progressivity p raises expected utility when earnings are uncertain by reducing the consumption spread that workers face. This is captured by the second term (in square brackets) on the right-hand side of the equation. In addition, higher progressivity raises the dispersion of pre-tax earnings via the crowd-out elasticity $\varepsilon_{\beta,1-p}=-1$. This causes a loss in expected utility proportional to the conditional variance of log-earnings $\mathbb{V}(\log z|\theta)$. Finally, notice that the crowding-in $\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}$ does not appear in formula (24) because this effect operates via optimal labor effort choices; the envelope theorem implies that its impact on welfare is only of second-order.[28]

Proof of Proposition 2. The proof proceeds in several steps. We first derive the effect of a change in progressivity on the social objective. Second, we evaluate its impact on government revenue by decomposing it into a statutory effect, a standard behavioral effect with exogenous private insurance, and fiscal externalities due to crowd-out and crowd-in. Third, we compute the marginal value of public funds. Finally, we equate the sum of these effects to zero to obtain our characterization of optimal tax progressivity.

Social Welfare Effect. Denote the change in the social welfare objective resulting from a change in tax progressivity by

$$WE=\int \alpha(\theta)\frac{\partial U(\theta)}{\partial(1-p)}dF(\theta).$$

By equation (24), we have

$$WE=-\frac{1}{1-p}+\int \alpha(\theta)\log(\ell\theta)dF(\theta)+\log(\mathbb{E}[z|\theta])-\mathbb{E}[\log z|\theta]-\frac{b}{\mathbb{E}z}\beta\ell(1-\ell)\varepsilon_{\beta,1-p},$$

[28] More precisely: Consider an (equivalent) dual formulation of the firm's problem, which consists of maximizing the worker's expected utility subject to making non-negative profits. The envelope theorem applied to this problem implies that changes in labor effort do not have first-order effects on expected utility.

where $\log(\mathbb{E}[z|\theta]) - \mathbb{E}[\log z|\theta] = \log(1 + \ell(e^{\beta} - 1)) - \beta\ell$. Now suppose that $\log\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$ and $\alpha(\theta) \propto e^{-a\log\theta}$. Note that, if a random variable x is normally distributed with mean μ and variance σ^2, we have $\mathbb{E}[e^{-ax}] = e^{-a\mu + \frac{1}{2}a^2\sigma^2}$. Moreover, letting φ denote the pdf of x, we have $\varphi'(x) = -\frac{x-\mu}{\sigma^2}\varphi(x)$, so that

$$\mathbb{E}[(x-\mu)e^{-ax}] = \int (x-\mu)e^{-ax}\varphi(x)dx = -\sigma^2 \int e^{-ax}\varphi'(x)dx.$$

An integration by parts implies that this expression is equal to $-a\sigma^2 \int e^{-ax}\varphi(x)dx = -a\sigma^2 e^{-a\mu+\frac{1}{2}a^2\sigma^2}$. Therefore, we obtain $\mathbb{E}[xe^{-ax}] = (\mu - a\sigma^2)e^{-a\mu+\frac{1}{2}a^2\sigma^2}$. Hence,

$$\int \alpha(\theta)\log(\ell\theta)dF(\theta) = \log\ell + \frac{\int e^{-a\log\theta}\log\theta f(\theta)d\theta}{\int e^{-a\log\theta}f(\theta)d\theta} = \log\ell + \mu_\theta - a\sigma_\theta^2.$$

Statutory Revenue Effect. Government revenue is equal to

$$\int \mathbb{E}[T(z)|\theta]dF(\theta) = \int [(1-\ell)T(\underline{z}) + \ell T(\bar{z})]dF(\theta) = Z - C,$$

where $Z \equiv \int \mathbb{E}[z|\theta]dF(\theta)$ is aggregate income, and $C \equiv \int \mathbb{E}[R(z)|\theta]dF(\theta)$ is aggregate consumption. Under the assumptions of Proposition 2, we have

$$C = \frac{1-\tau}{1-p}\frac{1+\ell(e^{(1-p)\beta}-1)}{[1+\ell(e^{\beta}-1)]^{1-p}}\int(\theta\ell)^{1-p}dF(\theta), \tag{25}$$

with

$$\int(\theta\ell)^{1-p}dF(\theta) = \left[(1-p)\mu_\theta + \frac{1}{2}(1-p)^2\sigma_\theta^2\right]\ell^{1-p}.$$

The statutory (or mechanical) effect is obtained by evaluating the change in government revenue following a change in progressivity keeping the contract $(\ell, \underline{z}, \bar{z})$ and hence β fixed, that is,

$$ME = \int \left.\frac{\partial\mathbb{E}[T(z)|\theta]}{\partial(1-p)}\right|_{\ell,\underline{z},\bar{z}} dF(\theta)$$

We obtain:

$$ME = \left[\frac{1}{1-p} - \frac{\beta\ell e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)} + \log[1+\ell(e^{\beta}-1)] - \frac{\partial \log \int (\theta\ell)^{1-p} dF(\theta)}{\partial(1-p)}\right] C,$$

with

$$\begin{aligned}\frac{\partial}{\partial(1-p)} \log \int (\theta\ell)^{1-p} dF(\theta) &= \log \ell + \frac{\partial}{\partial(1-p)} \log \int \theta^{1-p} dF(\theta) \\ &= \log \ell + \mu_\theta + (1-p)\sigma_\theta^2.\end{aligned}$$

Behavioral Effect with Exogenous Private Insurance. By equation (6), in response to a change in progressivity, the income levels change (in percentage terms) by

$$\frac{\partial \log \underline{z}}{\partial \log(1-p)} = \frac{\underline{z}}{\mathbb{E}[z|\theta]}\varepsilon_{\ell,1-p} - \beta\ell \frac{\bar{z}}{\mathbb{E}[z|\theta]}(\varepsilon_{\beta,1-p} + \varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}),$$

where we used the fact that $1 - \frac{\ell b}{\mathbb{E}[z|\theta]} = \frac{\underline{z}}{\mathbb{E}[z|\theta]}$, and

$$\frac{\partial \log \bar{z}}{\partial \log(1-p)} = \frac{\underline{z}}{\mathbb{E}[z|\theta]}\varepsilon_{\ell,1-p} + \beta(1-\ell)\frac{\underline{z}}{\mathbb{E}[z|\theta]}(\varepsilon_{\beta,1-p} + \varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}).$$

The standard behavioral effect of an increase in $1-p$ is equal to the change in government revenue triggered by labor effort responses ℓ only – that is, keeping the bonus rate β fixed. We get[29]

$$\begin{aligned}BE =& \frac{1}{1-p}\int \left(\mathbb{E}\left[T'(z)z\, \frac{\partial \log \underline{z}}{\partial \log(1-p)}\bigg|_{\beta} \bigg|\theta\right] + (T(\bar{z}) - T(\underline{z}))\ell\varepsilon_{\ell,1-p}\right) dF(\theta) \\ =& \frac{1}{1-p}\int \mathbb{E}[T'(z)z|\theta]\frac{\underline{z}}{\mathbb{E}[z|\theta]}\varepsilon_{\ell,1-p} dF(\theta) + \int \ell(T(\bar{z}) - T(\underline{z}))\varepsilon_{\ell,1-p} dF(\theta).\end{aligned}$$

Since $\varepsilon_{\ell,1-p}$ and $\frac{\underline{z}}{\mathbb{E}[z|\theta]}$ are constant (independent of θ), this expression can be rewritten

[29]Note that, in a model with only intensive-margin responses to taxes, i.e., with an exogenous probability π of earning the bonus, the free-entry condition $\mathbb{E}z \equiv (1-\pi)\underline{z} + \pi\bar{z} = \ell\theta$ would imply $\frac{\partial \log \underline{z}}{\partial \log(1-p)} = \frac{\partial \log \bar{z}}{\partial \log(1-p)} = \varepsilon_{\ell,1-p}$, and the change in government revenue caused by a change in progressivity would be equal to $\varepsilon_{\ell,1-p}\int \mathbb{E}[T'(z)z|\theta]dF(\theta)$. This is the expression we would obtain, for instance, in the full-insurance benchmark.

as:

$$BE = \frac{1}{1-p}\left[\frac{\underline{z}}{\mathbb{E}[z|\theta]}\int \mathbb{E}[T'(z)z|\theta]dF(\theta) + \ell\int(T(\bar{z}) - T(\underline{z}))dF(\theta)\right]\varepsilon_{\ell,1-p}.$$

With a CRP tax schedule, we can write

$$\int \mathbb{E}[T'(z)z|\theta]dF(\theta) = \int \mathbb{E}[z - (1-\tau)z^{1-p}|\theta]dF(\theta) = Z - (1-p)C.$$

The post-tax bonus rate is equal to $\log\frac{\bar{c}}{\underline{c}} = \log\frac{\frac{1-\tau}{1-p}\bar{z}^{1-p}}{\frac{1-\tau}{1-p}\underline{z}^{1-p}} = (1-p)\beta$. Hence, writing $\mathbb{E}[c|\theta] = (1-\ell)\underline{c} + \ell e^{(1-p)\beta}\underline{c}$ leads to $\frac{1}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\underline{c}}{\mathbb{E}[c|\theta]}$ and $\frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\bar{c}}{\mathbb{E}[c|\theta]}$. Therefore, $\frac{b}{\mathbb{E}[z|\theta]}$ and $\frac{\gamma}{\mathbb{E}[c|\theta]}$ are constant, where $\gamma \equiv \bar{c} - \underline{c}$. We can thus write the contribution of extensive margin adjustments to the excess burden of the rise in progressivity as follows:

$$\int(T(\bar{z}) - T(\underline{z}))dF(\theta) = \int\left[\left(\bar{z} - \frac{1-\tau}{1-p}\bar{z}^{1-p}\right) - \left(\underline{z} - \frac{1-\tau}{1-p}\underline{z}^{1-p}\right)\right]dF(\theta)$$
$$= \int bdF(\theta) - \int \gamma dF(\theta) = \frac{b}{\mathbb{E}[z|\theta]}\int \mathbb{E}[z|\theta]dF(\theta) - \frac{\gamma}{\mathbb{E}[c|\theta]}\int \mathbb{E}[c|\theta]dF(\theta) = \frac{b}{\mathbb{E}[z|\theta]}Z - \frac{\gamma}{\mathbb{E}[c|\theta]}C.$$

Collecting terms, and using the fact that $\ell\frac{\gamma}{\mathbb{E}[c|\theta]} = 1 - \frac{\underline{c}}{\mathbb{E}[c|\theta]}$, we obtain

$$\begin{aligned} BE =& \frac{1}{1-p}\left[\frac{\underline{z}}{\mathbb{E}[z|\theta]}Z - (1-p)\frac{\underline{z}}{\mathbb{E}[z|\theta]}C + \ell\frac{b}{\mathbb{E}[z|\theta]}Z - \ell\frac{\gamma}{\mathbb{E}[c|\theta]}C\right]\varepsilon_{\ell,1-p} \\ =& -\frac{1}{1-p}\left[1 - \frac{Z}{C} + (1-p)\frac{\underline{z}}{\mathbb{E}[z|\theta]} - \frac{\underline{c}}{\mathbb{E}[c|\theta]}\right]\varepsilon_{\ell,1-p}C. \end{aligned}$$

Fiscal Externalities from Crowd-Out and Crowd-In. Finally, the change in government revenue due to the endogeneity of the bonus rate β, keeping effort ℓ fixed, is given by

$$\begin{aligned} FE =& \frac{1}{1-p}\int\left[(1-\ell)T'(\underline{z})\underline{z}\left.\frac{\partial\log\underline{z}}{\partial\log(1-p)}\right|_{\ell} + \ell T'(\bar{z})\bar{z}\left.\frac{\partial\log\bar{z}}{\partial\log(1-p)}\right|_{\ell}\right]dF(\theta) \\ =& \frac{1}{1-p}[\varepsilon_{\beta,1-p} + \varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}]\beta\ell(1-\ell)\left[\frac{\underline{z}}{\mathbb{E}[z|\theta]}\int T'(\bar{z})\bar{z}dF(\theta) - \frac{\bar{z}}{\mathbb{E}[z|\theta]}\int T'(\underline{z})\underline{z}dF(\theta)\right], \end{aligned}$$

where the second equality uses the expressions derived above for the earnings elasticities. The term in square brackets can be rewritten as

$$\begin{aligned}
&\frac{\underline{z}}{\mathbb{E}[z|\theta]}\int[\bar{z}-(1-\tau)\bar{z}^{1-p}]dF(\theta)-\frac{\bar{z}}{\mathbb{E}[z|\theta]}\int[\underline{z}-(1-\tau)\underline{z}^{1-p}]dF(\theta)\\
=&(1-\tau)\frac{1}{1+\ell(e^{\beta}-1)}\frac{e^{\beta}-e^{(1-p)\beta}}{[1+\ell(e^{\beta}-1)]^{1-p}}\int(\theta\ell)^{1-p}dF(\theta)\\
=&\frac{1}{1-\ell}(1-p)\left[\frac{e^{\beta}}{1+\ell(e^{\beta}-1)}-\frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)}\right]C,
\end{aligned}$$

where the last equality follows from the expression (25) for C derived above. Thus, we obtain

$$FE=\beta\ell\left[\frac{e^{\beta}}{1+\ell(e^{\beta}-1)}-\frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)}\right][\varepsilon_{\beta,1-p}+\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}]C.$$

Marginal Value of Public Funds. The marginal value of public funds λ, when the tax code is restricted to the CRP class, is defined by the effect on social welfare of an increase the tax parameter τ, normalized to raise government revenue by 1 dollar. At the optimum tax schedule, λ is the Lagrange multiplier of the government budget constraint (12). We have

$$\frac{\partial\int\mathbb{E}[T(z)|\theta]dF(\theta)}{\partial\tau}=\frac{\partial Z}{\partial\tau}-\frac{\partial C}{\partial\tau}.$$

The first-order condition for effort (23) implies that $\frac{\partial\ell}{\partial\tau}=0$. Thus, $\frac{\partial Z}{\partial\tau}=0$ and, using expression (25), $\frac{\partial C}{\partial\tau}=-\frac{C}{1-\tau}$. Hence, we obtain

$$\frac{\partial\int\mathbb{E}[T(z)|\theta]dF(\theta)}{\partial\tau}=\frac{C}{1-\tau}.$$

The impact on social welfare of the normalized tax change is given by

$$\lambda=\left(\frac{C}{1-\tau}\right)^{-1}\int\alpha(\theta)\frac{\partial U(\theta)}{\partial\tau}dF(\theta)=\left(\frac{C}{1-\tau}\right)^{-1}\int\alpha(\theta)\frac{1}{1-\tau}dF(\theta)=\frac{1}{C}.$$

<u>*Optimal Rate of Progressivity.*</u> The optimal rate of progressivity is the solution to

$$\begin{aligned} 0 &= \frac{\partial \int \alpha(\theta)U(\theta)dF(\theta)}{\partial(1-p)} + \lambda \frac{\partial \int \mathbb{E}[T(z)|\theta]dF(\theta)}{\partial(1-p)} \\ &= WE + \frac{1}{C}[ME + BE + FE]. \end{aligned}$$

That is, the optimal level of p satisfies

$$\begin{aligned} 0 =& -\frac{1}{1-p} + \log \ell + \mu_\theta - a\sigma_\theta^2 - \log(1+\ell(e^\beta - 1)) + \beta\ell - \frac{b}{\mathbb{E}[z|\theta]}\beta\ell(1-\ell)\varepsilon_{\beta,1-p} \\ &+ \frac{1}{1-p} - \frac{\beta\ell e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)} + \log[1+\ell(e^\beta - 1)] - (\log \ell + \mu_\theta + (1-p)\sigma_\theta^2) \\ &- \frac{1}{1-p}\left[1 - \frac{Z}{C} + (1-p)\frac{\underline{z}}{\mathbb{E}[z|\theta]} - \frac{\underline{c}}{\mathbb{E}[c|\theta]}\right]\varepsilon_{\ell,1-p} \\ &+ \beta\ell\left[\frac{e^\beta}{1+\ell(e^\beta-1)} - \frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)}\right][\varepsilon_{\beta,1-p} + \varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}]. \end{aligned}$$

Rearranging terms, this formula can be rewritten as

$$\begin{aligned} 0 =& -(1-p+a)\sigma_\theta^2 - \beta\ell(1-\ell)\frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)} \\ &- \frac{1}{1-p}\left[1 - \frac{Z}{C} + (1-p)\frac{\underline{z}}{\mathbb{E}[z|\theta]} - \frac{\underline{c}}{\mathbb{E}[c|\theta]}\right]\varepsilon_{\ell,1-p} - \beta\ell(1-\ell)\frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)}\varepsilon_{\beta,1-p} \\ &+ \beta\ell\left[\frac{e^\beta}{1+\ell(e^\beta-1)} - \frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)}\right]\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}. \end{aligned}$$

We saw that $\frac{1}{1+\ell(e^\beta-1)} = \frac{\underline{z}}{\mathbb{E}[z|\theta]}$, $\frac{e^\beta}{1+\ell(e^\beta-1)} = \frac{\bar{z}}{\mathbb{E}[z|\theta]}$, $\frac{1}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\underline{c}}{\mathbb{E}[c|\theta]}$, and $\frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\bar{c}}{\mathbb{E}[c|\theta]}$. We can therefore rewrite the optimal tax equation as (recall that $\gamma \equiv \bar{c} - \underline{c}$)

$$\begin{aligned} &(1-p+a)\sigma_\theta^2 + \beta\ell(1-\ell)\frac{\gamma}{\mathbb{E}[c|\theta]}(1+\varepsilon_{\beta,1-p}) \\ =& \left[\frac{1}{1-p}\frac{Z}{C} - \left(\frac{\underline{z}}{\mathbb{E}[z|\theta]} + \frac{\ell}{1-p}\frac{\gamma}{\mathbb{E}[c|\theta]}\right)\right]\varepsilon_{\ell,1-p} + \beta\ell\left(\frac{\bar{z}}{\mathbb{E}[z|\theta]} - \frac{\bar{c}}{\mathbb{E}[c|\theta]}\right)\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}. \end{aligned}$$

Recall that $\mathbb{V}(\log z|\theta) = \beta^2\ell(1-\ell)$, $\frac{\bar{z}}{\mathbb{E}[z|\theta]} - \frac{\bar{c}}{\mathbb{E}[c|\theta]} = (1-\ell)(\frac{b}{\mathbb{E}[z|\theta]} - \frac{\gamma}{\mathbb{E}[c|\theta]})$, and $\frac{Z}{C} = 1 + \frac{G}{Z-G} \equiv 1 + \frac{s}{1-s}$ where s is the ratio of public expenditures S to output $Z = C + S$.

We thus get

$$
\begin{aligned}
&(1-p+a)\sigma_\theta^2 + \mathbb{V}(\log z|\theta)\frac{1}{\beta}\frac{\gamma}{\mathbb{E}[c|\theta]}(1+\varepsilon_{\beta,1-p}) \\
&= \left[\frac{s/(1-s)+p}{1-p} + \ell\left(\frac{b}{\mathbb{E}[z|\theta]} - \frac{1}{1-p}\frac{\gamma}{\mathbb{E}[c|\theta]}\right)\right]\varepsilon_{\ell,1-p} + \mathbb{V}(\log z|\theta)\frac{1}{\beta}\left(\frac{b}{\mathbb{E}[z|\theta]} - \frac{\gamma}{\mathbb{E}[c|\theta]}\right)\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}.
\end{aligned}
$$

Dividing through by $(1-p)$ and rearranging terms leads to

$$
\frac{p}{(1-p)^2} = \frac{(1+\frac{a}{1-p})\sigma_\theta^2 + \mathbb{V}(\log z|\theta)\frac{1}{(1-p)\beta}\frac{\gamma}{\mathbb{E}[c|\theta]}(1+\varepsilon_{\beta,1-p})}{\varepsilon_{\ell,1-p}\left[(1+\frac{s}{(1-s)p}) + \frac{1-p}{p}\ell(\frac{b}{\mathbb{E}[z|\theta]} - \frac{1}{1-p}\frac{\gamma}{\mathbb{E}[c|\theta]})\right] + \mathbb{V}(\log z|\theta)\frac{1-p}{\beta p}(\frac{b}{\mathbb{E}[z|\theta]} - \frac{\gamma}{\mathbb{E}[c|\theta]})\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}}
$$

Note that, to a second order as $\beta \to 0$ (keeping ℓ fixed), we get

$$
\begin{aligned}
\kappa_1\mathbb{V}(\log z|\theta) &= \frac{\beta\ell(1-\ell)}{1-p}\frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)} \sim \beta^2\ell(1-\ell) = \mathbb{V}(\log z|\theta) \\
\kappa_3\mathbb{V}(\log z|\theta) &= \beta\ell(1-\ell)\left[\frac{e^\beta-1}{1+\ell(e^\beta-1)} - \frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)}\right] \sim p\mathbb{V}(\log z|\theta).
\end{aligned}
$$

Note that $\kappa_3 > 0$ if and only if $\frac{e^\beta-1}{1+\ell(e^\beta-1)} = \frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)}$, or equivalently $p > 0$.

Extension to a model with fixed-pay jobs. Let s_{pp} be the fraction of performance-pay (or "moral-hazard") jobs, and s_{fp} the fraction of fixed-pay jobs in the economy. The welfare effect becomes:

$$
\begin{aligned}
WE = &-\frac{1}{1-p} + s_{pp}(\log\ell_{pp} + \mu_{\theta,pp}) + (1-s_{pp})(\log\ell_{fp} + \mu_{\theta,fp}) \\
&+ s_{pp}\left[\log\mathbb{E}[z_{pp}|\theta] - \mathbb{E}[\log z_{pp}|\theta] - \frac{b_{pp}}{\mathbb{E}[z_{pp}|\theta]}\beta\ell_{pp}(1-\ell_{pp})\varepsilon_{\beta,1-p}\right].
\end{aligned}
$$

Aggregate consumption is equal to $C = s_{pp}C_{pp} + (1-s_{pp})C_{fp}$, with

$$
\begin{aligned}
C_{pp} =& \frac{1-\tau}{1-p}\frac{1+\ell_{pp}(e^{(1-p)\beta}-1)}{[1+\ell_{pp}(e^\beta-1)]^{1-p}}\left[(1-p)\mu_{\theta,pp} + \frac{1}{2}(1-p)^2\sigma_{\theta,pp}^2\right]\ell_{pp}^{1-p} \\
C_{fp} =& \frac{1-\tau}{1-p}\left[(1-p)\mu_{\theta,fp} + \frac{1}{2}(1-p)^2\sigma_{\theta,fp}^2\right]\ell_{fp}^{1-p}.
\end{aligned}
$$

The mechanical effect can then be written as

$$ME = \frac{1}{1-p}C + \left[-\log \ell_{fp} - \mu_{\theta,fp} - (1-p)\sigma^2_{\theta,fp}\right](1-s_{pp})C_{fp}$$
$$+ \left[-\frac{\beta \ell_{pp} e^{(1-p)\beta}}{1+\ell_{pp}(e^{(1-p)\beta}-1)} + \log(1+\ell_{pp}(e^{\beta}-1)) - \log \ell_{pp} - \mu_{\theta,pp} - (1-p)\sigma^2_{\theta,pp}\right] s_{pp} C_{pp}.$$

The behavioral effect of the perturbation is equal to

$$BE = \frac{1}{1-p}\left[\frac{Z_{fp}}{C_{fp}} - (1-p)\right]\varepsilon_{\ell_{fp},1-p}(1-s_{pp})C_{fp}$$
$$- \frac{1}{1-p}\left[1 - \frac{Z_{pp}}{C_{pp}} + (1-p)\frac{\underline{z}_{pp}}{\mathbb{E}z_{pp}} - \frac{\underline{c}_{pp}}{\mathbb{E}c_{pp}}\right]\varepsilon_{\ell_{pp},1-p} s_{pp} C_{pp}$$

where Z_i is the aggregate output of jobs of type i, and $\underline{z}_{pp}/\mathbb{E}z_{pp}$ and $\underline{c}_{pp}/\mathbb{E}c_{pp}$ are constants defined as above. Finally, the fiscal externalities amount to

$$FE = \beta \ell_{pp}\left[\frac{e^{\beta}}{1+\ell_{pp}(e^{\beta}-1)} - \frac{e^{(1-p)\beta}}{1+\ell_{pp}(e^{(1-p)\beta}-1)}\right][\varepsilon_{\beta,1-p} + \varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}] s_{pp} C_{pp}.$$

The optimal rate of progressivity satisfies $0 = WE + \frac{1}{C}[ME + BE + FE]$. Using the previous expressions and rearranging terms following the same steps as above leads to

$$\frac{p}{(1-p)^2} = \frac{\Sigma^2_{\theta} + \frac{s_{pp}C_{pp}}{C}\kappa_1(1+\varepsilon_{\beta,1-p})\mathbb{V}(\log z_{pp}|\theta) - \kappa_4}{E_{\ell,1-p} + \frac{s_{pp}C_{pp}}{C}\ell_{pp}\kappa_2\varepsilon_{\ell_{pp},1-p} + \frac{s_{pp}C_{pp}}{C}\kappa_3\varepsilon_{\beta,\ell_{pp}}\varepsilon_{\ell_{pp},1-p}\mathbb{V}(\log z_{pp}|\theta)}$$

where we denote the average variance of abilities by

$$\Sigma^2_{\theta} = \frac{s_{pp}C_{pp}}{C}\sigma^2_{\theta,pp} + \frac{(1-s_{pp})C_{fp}}{C}\sigma^2_{\theta,fp},$$

the average labor supply elasticity by

$$E_{\ell,1-p} = \frac{s_{pp}C_{pp}}{C}\left(1+\frac{1}{p}\left(\frac{Z_{pp}}{C_{pp}}-1\right)\right)\varepsilon_{\ell_{pp},1-p} + \frac{(1-s_{pp})C_{fp}}{C}\left(1+\frac{1}{p}\left(\frac{Z_{fp}}{C_{fp}}-1\right)\right)\varepsilon_{\ell_{fp},1-p},$$

the constants $\kappa_1, \kappa_2, \kappa_3$ by

$$\kappa_1 = \frac{1}{\beta(1-p)} \left(\frac{\bar{c}_{pp} - \underline{c}_{pp}}{\mathbb{E}[c_{pp}|\theta]} + \frac{1 - \frac{C_{pp}}{C}}{\frac{C_{pp}}{C}} \frac{\bar{z}_{pp} - \underline{z}_{pp}}{\mathbb{E}[z_{pp}|\theta]} \right)$$
$$\kappa_2 = \frac{1-p}{p} \left(\frac{\bar{z}_{pp} - \underline{z}_{pp}}{\mathbb{E}[z_{pp}|\theta]} - \frac{1}{1-p} \frac{\bar{c}_{pp} - \underline{c}_{pp}}{\mathbb{E}[c_{pp}|\theta]} \right)$$
$$\kappa_3 = \frac{1-p}{\beta p} \left(\frac{\bar{z}_{pp} - \underline{z}_{pp}}{\mathbb{E}[z_{pp}|\theta]} - \frac{\bar{c}_{pp} - \underline{c}_{pp}}{\mathbb{E}[c_{pp}|\theta]} \right)$$

and the constant κ_4 is given by

$$\begin{aligned}\kappa_4 = \; & \frac{1}{1-p}(1 - s_{pp}) \left(1 - \frac{C_{fp}}{C}\right) [\mu_{\theta,fp} + \log \ell_{fp}] \\ & + \frac{1}{1-p} s_{pp} \left(1 - \frac{C_{pp}}{C}\right) \left[\mu_{\theta,pp} + \log \ell_{pp} + \log \frac{\underline{z}_{pp}}{\mathbb{E}[z_{pp}|\theta]} + \beta \ell_{pp} \frac{\bar{z}_{pp}}{\mathbb{E}[z_{pp}|\theta]}\right].\end{aligned}$$

This concludes the proof. ■

A.4 Proofs of Section 3.2

Lemma 6 *Suppose that the tax schedule is affine with tax rate τ and lump-sum transfer R_0. The optimal contract satisfies*

$$\underline{z} = \frac{1 - (e^{h'(\ell)} - 1)\frac{R_0}{(1-\tau)\theta}}{1 + \ell(e^{h'(\ell)} - 1)} \ell\theta \quad \textit{and} \quad \bar{z} = \frac{e^{h'(\ell)} + \frac{1-\ell}{\ell}(e^{h'(\ell)} - 1)\frac{R_0}{(1-\tau)\theta}}{1 + \ell(e^{h'(\ell)} - 1)} \ell\theta. \tag{26}$$

Labor effort, when interior, satisfies

$$1 = \frac{(e^{h'(\ell)} - 1)(\ell + \frac{R_0}{(1-\tau)\theta})}{1 + \ell(e^{h'(\ell)} - 1)} (1 + \ell(1-\ell)h''(\ell)). \tag{27}$$

As $\theta \to \infty$, labor effort converges to $\ell^* \in (0, 1)$ and $\beta = \log(\bar{z}/\underline{z})$ converges to $h'(\ell^*)$.

Proof of Lemma 6. The local incentive constraint (5) reads

$$\log((1-\tau)\bar{z} + R_0) - \log((1-\tau)\underline{z} + R_0) = h'(\ell).$$

It implies that $\bar{z} = e^{h'(\ell)}\underline{z} + (e^{h'(\ell)} - 1)\tilde{R}_0$, where we let $\tilde{R}_0 = \frac{R_0}{1-\tau}$. Plugging this expression into the zero profit condition $\ell\bar{z} + (1-\ell)\underline{z} = \theta\ell$ yields equations (26).

In particular, we have $b = \bar{z} - \underline{z} = (e^{h'(\ell)} - 1)(1 + \frac{\tilde{R}_0}{\theta}\frac{1}{\ell})\frac{\theta\ell}{1+\ell(e^{h'(\ell)-1})}$. The first-order condition for effort, derived in Lemma 9 in Appendix B, reads

$$\theta = b + \left[\frac{1}{v'(\bar{z})} - \frac{1}{v'(\underline{z})}\right]\ell(1-\ell)h''(\ell).$$

Note that $v'(z) = \frac{1-\tau}{(1-\tau)z+R_0} = (z + \tilde{R}_0)^{-1}$, so that $\frac{1}{v'(\bar{z})} - \frac{1}{v'(\underline{z})} = \bar{z} - \underline{z} = b$. Use this relation, plug the expression for b and divide both sides by θ to get equation (27). To see that ℓ^* is interior, consider the limit of the right-hand side of (27) when $\theta \to \infty$. This limit is 0 when $\ell^* = 0$ and $(e^{h'(\ell^*)} - 1)/e^{h'(\ell^*)} < 1$ if $\ell^* = 1$. Thus, (27) can be satisfied in the limit only if $\ell^* \in (0, 1)$. ■

Proof of Proposition 3. Suppose that the optimal marginal tax rate is constant above some earnings $\tilde{z}$. For agents with earnings sufficiently above $\tilde{z}$, the tax schedule is effectively affine, with marginal tax rate τ and lump-sum transfer $R_0 = \tau\tilde{z} - T(\tilde{z})$. We consider a uniform increase in the marginal tax rate τ in the interval $[z^*, \infty)$ by a small $\delta > 0$, where $z^* \gg \tilde{z}$. This perturbation is represented by the functional $\hat{R}(z) = -(z - z^*)\mathbb{I}_{\{z \geq z^*\}}$ and $\hat{R}'(z) = -\mathbb{I}_{\{z \geq z^*\}}$; see Section B. Thus, the tax liability levied after the reform on workers with income $z > z^*$ is $T(z^*) + (\tau + \delta)(z - z^*)$. Let θ^* denote the threshold ability type above which an agent earns income above z^* with positive probability, and θ^{**} be the threshold ability type above which the base pay and the high-level pay are both above z^*.

The first-order effect as $\delta \to 0$ (Gateaux derivative) of the perturbation on government revenue is given by

$$\hat{R} = \int_{z^*}^{\infty}(z - z^*)dF_z(z) + \tau\int_{\theta^*}^{\infty}\theta\hat{\ell}dF(\theta),$$

where F_z denotes the distribution of earnings in the economy. The first integral in the right-hand side is the mechanical effect of the perturbation, i.e., the impact of the tax reform on government revenue keeping individual earnings and behavior fixed. The second integral is the behavioral effect of the reform, which accounts for the endogenous responses of labor contracts. The impact of these responses can be summarized by the change in average earnings $\theta\hat{\ell}$ of all individuals whose income is larger than z^* with positive probability. This is true even for workers in $[\theta^*, \theta^{**}]$, because their base pay $\underline{z} < z^*$, is initially taxed at the same rate as their high-level pay $\bar{z} > z^*$. Now, we can rewrite this expression as

$$\hat{R} = (1 - F_z(z^*))(Z - z^*) - \frac{\tau}{1-\tau}(1 - F(\theta^*))\,\bar{Z}\,\mathcal{E},$$

where Z is average earnings in $[z^*, \infty)$, $\overline{Z}$ is average earnings among workers with ability in $[\theta^*, \infty)$, and

$$\mathcal{E} = \int_{\theta^*}^{\infty} \frac{\theta\ell}{\overline{Z}} \left(-\frac{1-\tau}{\theta\ell}\theta\hat{\ell}\right) \frac{dF(\theta)}{1 - F(\theta^*)}$$

is the (income-weighted) average elasticity of mean earnings among types $[\theta^*, \infty)$ with respect to the tax rate on incomes $[z^*, \infty)$. Defining

$$\psi_\rho = \lim_{z^* \to \infty} \frac{(1 - F(\theta^*))\overline{Z}}{(1 - F(z^*))Z},$$

we obtain

$$\lim_{z^* \to \infty} \frac{\hat{R}}{(1 - F_z(z^*))(Z - z^*)} = 1 - \frac{\tau}{1-\tau}\,\psi_\rho\,\rho^*\,\mathcal{E}^*,$$

where $\rho^* = \lim_{z^* \to \infty} \frac{Z}{Z - z^*}$ is the Pareto coefficient of the income distribution, and $\mathcal{E}^*$ is the limit of $\mathcal{E}$ as $z^* \to \infty$.

To compute the first-order effect of the perturbation of the social welfare objective, use Lemma 8 in Section B to show that the Gateaux derivative of individual utility in response to the tax reform is given by

$$\hat{U} = \ell\mu(\bar{z} \mid \theta)u'(R(\bar{z}))\hat{R}(\bar{z}) + (1 - \ell)\mu(\underline{z} \mid \theta)u'(R(\underline{z}))\hat{R}(\underline{z}),$$

where we denote

$$\mu(z \mid \theta) \equiv \frac{1/v'(\bar{z})}{\mathbb{E}[1/v'(z) \mid \theta]} = \frac{z + \frac{R_0}{1-\tau}}{\ell\bar{z} + (1-\ell)\underline{z} + \frac{R_0}{1-\tau}}.$$

Thus, the first-order effect of the tax reform on social welfare is given by

$$\widehat{SW} = -\int_{\theta^*}^{\infty} \left\{\ell(\bar{z} - z^*)\tilde{g}(\bar{z} \mid \theta) + (1 - \ell)(\underline{z} - z^*)\mathbb{I}_{\{\underline{z} \geq z^*\}}\tilde{g}(\underline{z} \mid \theta)\right\} dF(\theta),$$

where $\tilde{g}(\cdot \mid \theta)$ is defined in equation (17). We can rewrite the previous expression as

$$\widehat{SW} = -(1 - F_z(z^*))(Z - z^*)(G(z^*) + \Delta(z^*)),$$

where $G(z^*) = \mathbb{E}\big[\frac{z - z^*}{Z - z^*} g(z \mid \theta) \mid z > z^*\big]$ is the (income-weighted) average marginal social welfare weight above z^* in the full-insurance setting, and

$$\Delta(z^*) = \int_{\theta^*}^{\infty} \bigg[\ell \frac{\bar{z} - z^*}{Z - z^*} (\mu(\bar{z} \mid \theta) - 1) g(\bar{z} \mid \theta) + (1 - \ell) \frac{\underline{z} - z^*}{Z - z^*} \mathbb{I}_{\{\underline{z} \geq z^*\}} (\mu(\underline{z} \mid \theta) - 1) g(\underline{z} \mid \theta) \bigg] \frac{dF(\theta)}{1 - F_z(z^*)}$$

is the additional correction due to the crowd-out of within-firm insurance. Defining

$$\psi_g = \lim_{z^* \to \infty} 1 + G(z^*)/\Delta(z^*),$$

we obtain

$$\lim_{z^* \to \infty} \frac{\widehat{SW}}{(1 - F_z(z^*))(Z - z^*)} = -\psi_g \, g^*,$$

where g^* is the limit of $G(z^*)$ as $z^* \to \infty$.

Since the tax schedule is optimal, this perturbation cannot yield any welfare gains. In particular, as $z^* \to \infty$, we must have

$$\lim_{z \to z^*} \frac{1}{(1 - F_z(z^*))(Z - z^*)} \big[\hat{R} + \widehat{SW}\big] = 0.$$

This implies

$$1 - \frac{\tau^*}{1 - \tau^*} \psi_\rho \, \rho^* \, \mathcal{E}^* - \psi_g \, G^* = 0,$$

which easily leads to formula (16).

We now derive explicit expressions for ψ_ρ and ψ_g. We will repeatedly invoke the mean value theorem for integrals (MVT), according to which, for any continuous function $X(\cdot)$ and any θ_1, θ_2, there exists $t \in (\theta_1, \theta_2)$ such that $\int_{\theta_1}^{\theta_2} X(\theta) dF(\theta) = X(t) \int_{\theta_1}^{\theta_2} dF(\theta)$.

Regarding ψ_ρ, note that

$$\frac{1 - F(\theta^*)}{1 - F_z(z^*)} \cdot \frac{\overline{Z}}{Z} = \frac{\int_{\theta^*}^{\infty} \theta \ell \, dF(\theta)}{\int_{z^*}^{\infty} z \, dF_z(z)} = \frac{\int_{\theta^*}^{\infty} \theta \ell \, dF(\theta)}{\int_{\theta^*}^{\infty} \theta \ell \, dF(\theta) - \int_{\theta^*}^{\infty} \mathbf{1}_{\underline{z} \leq z^*} (1 - \ell) \underline{z} \, dF(\theta)}.$$

Since $z^* = \underline{z}(\theta^{**})$, the second term in the denominator becomes $\int_{\theta^*}^{\theta^{**}}(1-\ell)\underline{z}dF(\theta)$ and we can write

$$\frac{1-F(\theta^*)}{1-F_z(z^*)}\cdot\frac{\overline{Z}}{Z} = \left[1 - \frac{\int_{\theta^*}^{\theta^{**}}(1-\ell)\underline{z}dF(\theta)}{\int_{\theta^*}^{\infty}\ell\theta dF(\theta)}\right]^{-1}.$$

The MVT implies that there exists $t_1 \in (\theta^*, \theta^{**})$ and $t_2 \geq \theta^*$ such that

$$\frac{\int_{\theta^*}^{\theta^{**}}(1-\ell)\underline{z}dF(\theta)}{\int_{\theta^*}^{\infty}\ell\theta dF(\theta)} = \frac{(1-\ell(t_1))\underline{z}(t_1)/t_1}{\ell(t_2)}\cdot\frac{\int_{\theta^*}^{\theta^{**}}\theta dF(\theta)}{\int_{\theta^*}^{\infty}\theta dF(\theta)}$$

By Lemma 6, the first ratio converges to $\frac{1-\ell^*}{1+\ell^*(e^{h'(\ell^*)}-1)}$ as $z^* \to \infty$. Since θ is Pareto distributed, the second ratio is equal to $1 - (\frac{\theta^*}{\theta^{**}})^{\rho^*-1}$, where $\rho^* > 1$ is the Pareto coefficient; it is easy to show that it coincides with that of the tail of the earnings distribution. Using the identity $\bar{z}(\theta^*) = \underline{z}(\theta^{**}) = z^*$ and Lemma 6, we find $\lim_{z^*\to\infty}\frac{\theta^*}{\theta^{**}} = e^{-h'(\ell^*)}$. Thus,

$$\psi_\rho = \left[1 - \frac{(1-\ell^*)}{1+\ell^*(e^{h'(\ell^*)}-1)}\left(1-e^{-(\rho^*-1)h'(\ell^*)}\right)\right]^{-1} > 1.$$

Regarding ψ_g, let $\bar{\mu}^* = \lim_{\theta\to\infty}\mu(\bar{z}|\theta)$ and $\underline{\mu}^* = \lim_{\theta\to\infty}\mu(\underline{z}|\theta)$. By Lemma 6, we have

$$\bar{\mu}^* = \frac{e^{h'(\ell^*)}}{\ell^* e^{h'(\ell^*)} + 1 - \ell^*} \quad \text{and} \quad \underline{\mu}^* = \frac{\bar{\mu}^*}{e^{h'(\ell^*)}}.$$

Next, define

$$A(z^*) \equiv \int_{\theta^*}^{\infty}\frac{\bar{z}-z^*}{Z-z^*}\ell\frac{dF(\theta)}{1-F_z(z^*)} = \frac{\int_{\theta^*}^{\infty}(\bar{z}-z^*)\ell dF(\theta)}{\int_{\theta^*}^{\infty}(\bar{z}-z^*)\ell dF(\theta) + \int_{\theta^{**}}^{\infty}(\underline{z}-z^*)(1-\ell)dF(\theta)}$$
$$= \left[1 + \frac{\int_{\theta^{**}}^{\infty}(\underline{z}-z^*)(1-\ell)dF(\theta)}{\int_{\theta^*}^{\infty}(\bar{z}-z^*)\ell dF(\theta)}\right]^{-1},$$

and let A^* denote the limit of $A(z^*)$ as $z^* \to \infty$, if it exists. By the MVT, there exist

$t_1, t_1' > \theta^{**}$ and $t_2, t_2' > \theta^*$ such that

$$\frac{\int_{\theta^{**}}^{\infty}(\underline{z}-z^*)(1-\ell)dF(\theta)}{\int_{\theta^*}^{\infty}(\bar{z}-z^*)\ell dF(\theta)} = \frac{\frac{\underline{z}(t_1)}{t_1}(1-\ell(t_1))\int_{\theta^{**}}^{\infty}\theta dF(\theta) - z^*(1-\ell(t_1'))(1-F(\theta^{**}))}{\frac{\bar{z}(t_2)}{t_2}\ell(t_2)\int_{\theta^*}^{\infty}\theta dF(\theta) - z^*\ell(t_2')(1-F(\theta^*))}$$
$$= \frac{\theta^{**}}{\theta^*}\frac{1-F(\theta^{**})}{1-F(\theta^*)}\frac{\frac{\underline{z}(t_1)}{t_1}(1-\ell(t_1))\frac{\rho^*}{\rho^*-1} - \frac{\underline{z}(\theta^{**})}{\theta^{**}}(1-\ell(t_1'))}{\frac{\bar{z}(t_2)}{t_2}\ell(t_2)\frac{\rho^*}{\rho^*-1} - \frac{\bar{z}(\theta^*)}{\theta^*}\ell(t_2')}$$

where in the second line we used the following property of Pareto distributions: $\int_x^{\infty}\theta\frac{dF(\theta)}{1-F(x)} = x\frac{\rho^*}{\rho^*-1}$. The first two terms of this expression are equal to $(\frac{\theta^*}{\theta^{**}})^{\rho^*-1}$ and converge to $e^{-(\rho^*-1)h'(\ell^*)}$ as $z^* \to \infty$. The last term converges to $\frac{1-\ell^*}{\ell^*}e^{-h'(\ell^*)}$. As a result,

$$A^* = \left[1 + e^{-\rho^* h'(\ell^*)}\cdot\frac{1-\ell^*}{\ell^*}\right]^{-1} = \frac{\ell^* e^{\rho^* h'(\ell^*)}}{\ell^* e^{\rho^* h'(\ell^*)} + 1 - \ell^*}.$$

Now, applying the MVT to the expression $\Delta(z^*)$ dervived above implies that there exist $t_1 > \theta^*$ and $t_2 > \theta^{**}$ such that

$$\Delta = (\mu(\bar{z} \mid t_1) - 1)g(\bar{z} \mid t_1)A(z^*) + (\mu(\underline{z} \mid t_2) - 1)g(\underline{z} \mid t_2)(1 - A(z^*))$$
$$\xrightarrow[z^*\to\infty]{} (\bar{\mu}^* - 1)g^* A^* + (\underline{\mu}^* - 1)g^*(1 - A^*).$$

Therefore, we obtain

$$\psi_g = 1 + \frac{\Delta^*}{g^*} = 1 + (\bar{\mu}^* - 1)A^* + (\underline{\mu}^* - 1)(1 - A^*)$$
$$= 1 + \ell^*(1-\ell^*)\frac{e^{\rho^* h'(\ell^*)} - 1}{\ell^* e^{\rho^* h'(\ell^*)} + 1 - \ell^*}\cdot\frac{e^{h'(\ell^*)} - 1}{\ell^* e^{h'(\ell^*)} + 1 - \ell^*} > 1.$$

This concludes the proof. ∎

B General Tax Incidence Analysis

In this section, we evaluate the impact of tax reforms on workers' labor contracts and welfare. Consider a given baseline retention schedule $R : \mathbb{R}_+ \to \mathbb{R}$ and another function ("tax reform") $\hat{R} : \mathbb{R}_+ \to \mathbb{R}$. The first-order change in a functional $\Psi(R)$ in response to the reform $\hat{R}$ is given by the Gateaux derivative $\hat{\Psi}(R, \hat{R}) \equiv \lim_{\delta\to 0}(\Psi(R + \delta\hat{R}) - \Psi(R))/\delta$. The proofs of the following results are gathered in Appendix B.1.

Lemma 7 *The first-order impact of a tax reform $\hat{R}$ on earnings $\underline{z}, \bar{z}$ is given by*

$$\hat{\underline{z}} = \left[-\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \frac{\hat{U}}{v'(\underline{z})}\right] - \frac{\ell h''(\ell)}{v'(\underline{z})}\hat{\ell} \tag{28}$$

$$\hat{\bar{z}} = \left[-\frac{\hat{R}(\bar{z})}{R'(\bar{z})} + \frac{\hat{U}}{v'(\bar{z})}\right] + \frac{(1-\ell)h''(\ell)}{v'(\bar{z})}\hat{\ell}, \tag{29}$$

where $\hat{U}$ and $\hat{\ell}$ denote the first-order changes in the reservation utility and labor effort due to the reform. The terms in square brackets constitute the crowd-out, while the terms multiplied by $\hat{\ell}$ constitute the crowd-in.

The first term in equations (28) and (29), $-\hat{R}(z)/R'(z)$, implies that, *ceteris paribus*, the agent's consumption $c = R(z)$ remains unchanged despite the tax change. Indeed, this term implies $\hat{c} = \hat{R}(z) + R'(z)\hat{z} = 0$. Thus, absent any change in the reservation utility and optimal labor effort, the firm would simply adjust pre-tax earnings so as to keep the worker's disposable income levels $\underline{c}$ and $\bar{c}$ fixed. In other words, any attempt by the government to affect consumption insurance would be fully offset by the firm in order to preserve incentives.

Second, the tax reform affects the earnings contract via its impact on the equilibrium reservation utility. The increase in income z resulting from an increase in the reservation value $\hat{U} > 0$ is inversely proportional to the marginal utility $v'(z)$. Thus, the earnings of the high-performers increase by a larger amount than those of the low-performers with the same ability. This ensures that the utility gain $\hat{U}$ is distributed uniformly across agents regardless of their performance, thus preserving incentive compatibility (18).

Third, the tax reform modifies the desired effort level. Recall that, by equation (18), eliciting higher effort $\hat{\ell} > 0$ requires widening the gap between the utility of high- and low-performers by $\Delta h'(\ell) = h''(\ell)\hat{\ell}$. The implied change in earnings—the crowd-in effect—is given by the third term in equations (28) and (29).

Lemma 8 *The first-order impact of a tax reform $\hat{R}$ on expected utility $U(\theta)$ is given by*

$$\hat{U} = \mathbb{E}\left[\mu(z)u'(R(z))\hat{R}(z)|\theta\right] \text{ where } \mu(z) = \frac{\frac{1}{v'(z)}}{\mathbb{E}\left[\frac{1}{v'}|\theta\right]} \text{ for } z \in \{\underline{z}, \bar{z}\}, \tag{30}$$

In the standard (full-insurance) model, a tax cut $\hat{R}(z) > 0$ affects the worker's utility in proportion to their marginal utility of consumption $u'(R(z))$, and the envelope theorem ensures that the endogenous behavioral responses have no first-order impact on utility. In the model with performance pay, the envelope theorem still applies to the endogenous effort (i.e., crowd-in) responses: Equation (30) shows that the worker's expected utility is unaffected by changes in labor effort. However, the earnings adjustments caused by the crowd-out of private insurance have a first-order impact on welfare. The additional factor μ present in equation (30) accounts for these welfare effects: To keep effort incentive-compatible, earnings z must change in proportion to $\frac{1}{v'(z)}$ so that the utility difference $v(\bar{z}) - v(\underline{z})$ remains unchanged.

Lemma 9 *Suppose that the base pay $\underline{z}$ and high-level pay $\bar{z}$ are both located in brackets where the marginal tax rate is locally constant. The effect of a tax reform $\hat{R}$ on labor effort ℓ is given by:*

$$\frac{\hat{\ell}}{\ell} = \varepsilon_{\ell,R'(\underline{z})} \frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \varepsilon_{\ell,R'(\bar{z})} \frac{\hat{R}'(\bar{z})}{R'(\bar{z})} + \varepsilon_{\ell,R(\underline{z})} \frac{\hat{R}(\underline{z})}{R'(\underline{z})\underline{z}} + \varepsilon_{\ell,R(\bar{z})} \frac{\hat{R}(\bar{z})}{R'(\bar{z})\bar{z}} \tag{31}$$

where the elasticities of labor effort with respect to the marginal tax rates at $\underline{z}$ and $\bar{z}$ are respectively given by:

$$\varepsilon_{\ell,R'(\underline{z})} = -\frac{1}{D}\left(\frac{\ell b}{\underline{z}}\varepsilon_{in}\right) \quad \text{and} \quad \varepsilon_{\ell,R'(\bar{z})} = \frac{1}{D}\left(\frac{\ell b}{\underline{z}}\varepsilon_{in} + 1\right)$$

and the elasticities of labor effort with respect to the average tax rates at $\underline{z}$ and $\bar{z}$ are respectively given by:

$$\varepsilon_{\ell,R(\underline{z})} = -\frac{1}{D}\left(1 - \varepsilon_{out} + \frac{(1-\ell)\underline{z}}{\mathbb{E}[1/v'|\theta]}E\right) \quad \text{and} \quad \varepsilon_{\ell,R(\bar{z})} = -\frac{1}{D}\left(-\varepsilon_{out} + \frac{(1-\ell)\underline{z}}{\mathbb{E}[1/v'|\theta]}E\right),$$

where we denote:

$$D \equiv -\frac{\ell^2}{\underline{z}}\frac{\partial^2\Pi(\theta)}{\partial \ell^2} > 0 \quad \text{and} \quad E \equiv \left(\frac{\ell b}{\underline{z}}\varepsilon_{in} + 1\right)\frac{-u''(\bar{c})}{(u'(\bar{c}))^2} - \left(\frac{\ell b}{\underline{z}}\varepsilon_{in}\right)\frac{-u''(\underline{c})}{(u'(\underline{c}))^2}.$$

Thus, a higher marginal tax rate on base (resp. high-level) pay raises (resp., reduces) labor effort: $\varepsilon_{\ell,R'(\underline{z})} < 0 < \varepsilon_{\ell,R'(\bar{z})}$. Suppose moreover that the utility function is logarithmic, which implies $E = 1$. A higher average tax rate on base pay raises labor effort: $\varepsilon_{\ell,R(\underline{z})} < 0$. A higher average tax rate on high-level pay reduces labor effort, so that $\varepsilon_{\ell,R(\bar{z})} > 0$, if and only if $R'(\underline{z}) < R(\underline{z})/\underline{z}$.

In the standard model, effort responds negatively to the marginal tax rate (MTR) due to the substitution effect and positively to the average tax rate (ATR) due to the income effect. In our model, a higher ATR levied on the base pay increases effort, but a higher ATR levied on the high-level pay reduces effort (when the tax schedule is progressive). Intuitively, adjusting effort in this way allows the worker to escape some of the increased tax burden by reducing the probability of receiving the income level that is taxed more heavily.

Moreover, a higher MTR levied on high-level pay reduces effort, but a higher MTR on base pay, perhaps surprisingly, *increases* effort. Note that a change in marginal rates alone does not modify workers' incentives for effort, since the incentive constraint (18) is unaffected. What are modified, however, are the firm's incentives to offer different pay structures. Intuitively, a higher MTR gives firms incentives to reduce gross pay and save on payroll, since the implied reduction in after-tax income—which matters for workers—is now smaller. A higher MTR at the high-level pay $\bar{z}$ then leads to a lower value of $\bar{z}$ and, via the incentive constraint, to a lower effort level. Conversely, a higher MTR on base pay $\underline{z}$ leads to a lower value of $\underline{z}$ and, hence, a higher effort level.

Lemma 10 *The first-order impact of a tax reform $\hat{R}$ on earnings $\underline{z}, \bar{z}$ is given by*

$$(1-\ell)\hat{\underline{z}} \;=\; -\,(1-\varepsilon_{out})(1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} \;+\; \varepsilon_{out}\,\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})} \;-\; \ell\bar{z}\,\varepsilon_{in}\,\frac{\hat{\ell}}{\ell} \tag{32}$$

$$\ell\hat{\bar{z}} \;=\; (1-\varepsilon_{out})(1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} \;-\; \varepsilon_{out}\,\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})} \;+\; \left[\ell\bar{z}\varepsilon_{in}+\underline{z}\right]\frac{\hat{\ell}}{\ell}, \tag{33}$$

where the crowd-out parameter $\varepsilon_{out} \in (0,1)$ and the crowd-in parameter $\varepsilon_{in} > 0$ are respectively given by

$$\varepsilon_{out} \;=\; \frac{1}{\mathbb{E}\left[\frac{1}{v'(z)}|\theta\right]}\,\frac{1-\ell}{v'(\underline{z})} \qquad \textit{and} \qquad \varepsilon_{in} \;=\; \frac{\ell(1-\ell)h''(\ell)}{\bar{z}\,v'(\underline{z})},$$

and where the labor supply response $\hat{\ell}/\ell$ is given by Lemma 9.

Lemma 10 is obtained by substituting expression (30) into equations (28) and (29) and gives a full characterization of the incidence of tax reforms on the earnings contract. It generalizes the crowding-in and crowding-out forces highlighted in Section

2 to arbitrary tax systems and reforms. Our model gives simple analytic expressions for the crowd-out and crowd-in parameters $\varepsilon_{out} \in (0,1)$ and ε_{in}. For instance, if the utility of consumption is CRRA with risk aversion σ and the tax schedule is CRP, we have $\varepsilon_{out} = (1-\ell)/[1-\ell+\ell e^{(p+(1-p)\sigma)\beta}]$, where β is the pass-through defined in Section 2.

B.1 Proofs of Section B

Proof of Lemma 7. Consider a reform $\delta\hat{R} : \mathbb{R}_+ \to \mathbb{R}_+$ of the tax schedule, where $\delta \in \mathbb{R}$. Denote by $\hat{\underline{z}}$ and $\hat{\bar{z}}$ the Gateaux derivatives of base pay and high-performance pay following this reform, and by $\hat{\ell}$ and $\hat{U}$ those of labor effort and reservation utility. To a first order as $\delta \to 0$, the values of $\hat{\underline{z}}$ and $\hat{\bar{z}}$ are the solution to the following system:

$$u[R(\underline{z}+\delta\hat{\underline{z}})+\delta\hat{R}(\underline{z})]-h(\ell+\delta\hat{\ell}) = U(\theta)+\delta\hat{U}-(\ell+\delta\hat{\ell})h'(\ell+\delta\hat{\ell})$$
$$u[R(\bar{z}+\delta\hat{\bar{z}})+\delta\hat{R}(\bar{z})]-h(\ell+\delta\hat{\ell}) = U(\theta)+\delta\hat{U}+(1-\ell-\delta\hat{\ell})h'(\ell+\delta\hat{\ell}).$$

Linearizing this system around the initial equilibrium leads to

$$u'(R(\underline{z}))\hat{R}(\underline{z})+R'(\underline{z})u'(R(\underline{z}))\hat{\underline{z}}-h'(\ell)\hat{\ell} = \hat{U}-[h'(\ell)+\ell h''(\ell)]\hat{\ell}$$
$$u'(R(\bar{z}))\hat{R}(\bar{z})+R'(\bar{z})u'(R(\bar{z}))\hat{\bar{z}}-h'(\ell)\hat{\ell} = \hat{U}+[-h'(\ell)+(1-\ell)h''(\ell)]\hat{\ell}.$$

Rearranging terms and noting that $R'u' = v'$ leads to equations (28) and (29). ■

Proof of Lemma 8. The perturbed free-entry condition reads:

$$(\underline{z}+\delta\hat{\underline{z}})+(\ell+\delta\hat{\ell})(b+\delta\hat{b}) = \theta(\ell+\delta\hat{\ell}).$$

Linearizing this system around the initial equilibrium as $\delta \to 0$ leads to $\hat{\underline{z}}+\ell\hat{b}+b\hat{\ell} = \theta\hat{\ell}$ or, since $\hat{b} = \hat{\bar{z}} - \hat{\underline{z}}$,

$$(1-\ell)\hat{\underline{z}}+\ell\hat{\bar{z}} = (\theta-b)\hat{\ell}.$$

Substituting expressions (28) and (29) into this equation and rearranging terms, we obtain

$$\left[\frac{1-\ell}{v'(\underline{z})}+\frac{\ell}{v'(\bar{z})}\right]\hat{U}=\frac{1-\ell}{R'(\underline{z})}\hat{R}(\underline{z})+\frac{\ell}{R'(\bar{z})}\hat{R}(\bar{z})$$
$$+\left[\theta-b-\left(\frac{1}{v'(\bar{z})}-\frac{1}{v'(\underline{z})}\right)\ell(1-\ell)h''(\ell)\right]\hat{\ell}.$$

But the first-order condition for labor effort (21) when taxes are levied on total earnings can be written as

$$\theta=b+\left[\frac{1}{v'(\bar{z})}-\frac{1}{v'(\underline{z})}\right]\ell(1-\ell)h''(\ell). \tag{34}$$

Thus, the Gateaux derivative of expected utility is given by

$$\hat{U}=\frac{(1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})}+\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})}}{(1-\ell)\frac{1}{v'(\underline{z})}+\ell\frac{1}{v'(\bar{z})}},$$

which is equal to expression (30). ∎

Proof of Lemma 9. The first-order condition (34) for labor effort, expressed at the perturbed tax schedule and to a first order as $\delta\to 0$, reads:

$$\theta=b+\delta\hat{b}+\left[\frac{1}{[R'(\bar{z})+\delta(\hat{R}'(\bar{z})+R''(\bar{z})\hat{\bar{z}})]u'[R(\bar{z})+\delta(\hat{R}(\bar{z})+R'(\bar{z})\hat{\bar{z}})]}-\right.$$
$$\left.\frac{1}{[R'(\underline{z})+\delta(\hat{R}'(\underline{z})+R''(\underline{z})\hat{\underline{z}})]u'[R(\underline{z})+\delta(\hat{R}(\underline{z})+R'(\underline{z})\hat{\underline{z}})]}\right]$$
$$\times(\ell+\delta\hat{\ell})(1-\ell-\delta\hat{\ell})h''(\ell+\delta\hat{\ell}).$$

Suppose for simplicity that the tax schedule is piecewise linear, so that $R''(\underline{z})=R''(\bar{z})=0$. A first-order Taylor expansion of this expression around the initial equilibrium leads to

$$
\begin{aligned}
0 = & \left[\frac{\ell(1-\ell)h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))}\right]\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} - \left[\frac{\ell(1-\ell)h''(\ell)}{R'(\bar{z})u'(R(\bar{z}))}\right]\frac{\hat{R}'(\bar{z})}{R'(\bar{z})} \\
& + \left[\frac{\ell(1-\ell)h''(\ell)}{u'(R(\underline{z}))}\frac{u''(R(\underline{z}))}{u'(R(\underline{z}))}\right]\frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \left[\frac{\ell(1-\ell)h''(\ell)}{u'(R(\bar{z}))}\frac{u''(R(\bar{z}))}{u'(R(\bar{z}))}\right]\frac{\hat{R}(\bar{z})}{R'(\bar{z})} \\
& + \left[\frac{\ell(1-\ell)h''(\ell)}{u'(R(\underline{z}))}\frac{u''(R(\underline{z}))}{u'(R(\underline{z}))}\right]\hat{\underline{z}} - \left[\frac{\ell(1-\ell)h''(\ell)}{u'(R(\bar{z}))}\frac{u''(R(\bar{z}))}{u'(R(\bar{z}))}\right]\hat{\bar{z}} + \hat{b} \\
& + \ell(1-\ell)h''(\ell)\left(\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}\right)\left(\frac{1-2\ell}{1-\ell} + \frac{\ell h'''(\ell)}{h''(\ell)}\right)\frac{\hat{\ell}}{\ell}.
\end{aligned}
$$

Recall that, by the first-order condition for effort (34) and the zero-profit condition (4),

$$
\left[\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}\right]\ell(1-\ell)h''(\ell) = \theta - b = \frac{\underline{z}}{\ell}.
$$

Moreover, we saw that

$$
\hat{\underline{z}} = -\frac{\frac{\ell}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \frac{\frac{\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\bar{z})}{R'(\bar{z})} - \frac{\ell}{1-\ell}\frac{\ell(1-\ell)h''(\ell)}{v'(\underline{z})}\frac{\hat{\ell}}{\ell}
$$

and

$$
\hat{\bar{z}} = \frac{\frac{1-\ell}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \frac{\frac{1-\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\bar{z})}{R'(\bar{z})} + \frac{\ell(1-\ell)h''(\ell)}{v'(\bar{z})}\frac{\hat{\ell}}{\ell}
$$

and hence

$$
\hat{b} = \frac{\frac{1}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \frac{\frac{1}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\bar{z})}{R'(\bar{z})} + \left[\frac{\underline{z}}{\ell} + \frac{\ell h''(\ell)}{v'(\underline{z})}\right]\frac{\hat{\ell}}{\ell}.
$$

We thus obtain

$$
\begin{aligned}
D\frac{\hat{\ell}}{\ell} = & -\frac{\ell}{\underline{z}}\left[\frac{\ell(1-\ell)h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))}\right]\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \frac{\ell}{\underline{z}}\left[\frac{\ell(1-\ell)h''(\ell)}{R'(\bar{z})u'(R(\bar{z}))}\right]\frac{\hat{R}'(\bar{z})}{R'(\bar{z})} \\
& - \frac{\ell}{\underline{z}}\left[\frac{\frac{\frac{1}{1-\ell}}{R'(\bar{z})u'(R(\bar{z}))} + \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right](1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} \\
& + \frac{\ell}{\underline{z}}\left[\frac{\frac{\frac{1}{\ell}}{R'(\underline{z})u'(R(\underline{z}))} - \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right]\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})}
\end{aligned}
$$

where

$$D = \frac{1-2\ell}{1-\ell} + \frac{\ell h'''(\ell)}{h''(\ell)} + \ell h''(\ell)\times$$
$$+\left(\frac{\ell}{\underline{z}}\frac{1-\ell}{R'(\bar{z})u'(R(\bar{z}))} + \frac{\ell}{\underline{z}}\frac{\ell}{R'(\underline{z})u'(R(\underline{z}))} - \frac{\ell\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} + (1-\ell)\frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}}{\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}}\right).$$

Now, recall that the firm's profit is equal to $\Pi(\theta) = \ell\theta - \underline{z} - \ell b$. Thus, we can write

$$\begin{aligned}\frac{\partial \Pi(\theta)}{\partial \ell} &= \theta - b - \frac{\partial \underline{z}}{\partial \ell} - \ell\frac{\partial b}{\partial \ell}\\ &= \theta - b - \left[\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}\right]\ell(1-\ell)h''(\ell).\end{aligned}$$

The second-order condition to the firm's maximization problem reads:

$$\frac{\partial^2 \Pi(\theta)}{\partial \ell^2} \leq 0. \tag{35}$$

Differentiating the previous expression leads to

$$\begin{aligned}\frac{\partial^2 \Pi(\theta)}{\partial \ell^2} = &-\frac{\partial b}{\partial \ell} - \left[\frac{u''(R(\underline{z}))}{(u'(R(\underline{z})))^2}\frac{\partial \underline{z}}{\partial \ell} - \frac{u''(R(\bar{z}))}{(u'(R(\bar{z})))^2}\frac{\partial \bar{z}}{\partial \ell}\right]\ell(1-\ell)h''(\ell)\\ &-\frac{1}{\ell}\left[\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}\right]\left[\frac{1-2\ell}{1-\ell} + \frac{\ell h'''(\ell)}{h''(\ell)}\right]\ell(1-\ell)h''(\ell).\end{aligned}$$

But recall that

$$\frac{\partial \underline{z}}{\partial \ell} = -\frac{\ell h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))} \quad \text{and} \quad \frac{\partial b}{\partial \ell} = \left[\frac{1-\ell}{R'(\bar{z})u'(R(\bar{z}))} + \frac{\ell}{R'(\underline{z})u'(R(\underline{z}))}\right]h''(\ell).$$

Hence, we obtain

$$-\frac{\ell^2}{\underline{z}}\frac{\partial^2 \Pi(\theta)}{\partial \ell^2} = \frac{1-2\ell}{1-\ell} + \frac{\ell h'''(\ell)}{h''(\ell)} + \ell h''(\ell)\times$$
$$\left(\frac{\ell}{\underline{z}}\frac{1-\ell}{R'(\bar{z})u'(R(\bar{z}))} + \frac{\ell}{\underline{z}}\frac{\ell}{R'(\underline{z})u'(R(\underline{z}))} - \frac{\ell\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} + (1-\ell)\frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}}{\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}}\right).$$

where we used again equation (34) with $\theta - b = \underline{z}/\ell$. We can therefore rewrite the

Gateaux derivative of labor effort as

$$\left(-\ell\frac{\partial^2\Pi(\theta)}{\partial\ell^2}\right)\frac{\hat{\ell}}{\ell} = -\left[\frac{\ell(1-\ell)h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))}\right]\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \left[\frac{\ell(1-\ell)h''(\ell)}{R'(\bar{z})u'(R(\bar{z}))}\right]\frac{\hat{R}'(\bar{z})}{R'(\bar{z})}$$
$$-\left[\frac{\frac{\frac{1}{1-\ell}}{R'(\bar{z})u'(R(\bar{z}))} + \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right](1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})}$$
$$+\left[\frac{\frac{\frac{1}{\ell}}{R'(\underline{z})u'(R(\underline{z}))} - \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right]\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})}$$

where, by condition (35), the term multiplying $\hat{\ell}/\ell$ in the left hand side is positive. Using the definition of ε_{out}, and noting that

$$\frac{\ell b}{\underline{z}}\varepsilon_{in} = \frac{\ell}{\underline{z}}\frac{\ell(1-\ell)h''(\ell)}{v'(\underline{z})} \quad \text{and} \quad \frac{\ell b}{\underline{z}}\varepsilon_{in} + 1 = \frac{\ell}{\underline{z}}\frac{\ell(1-\ell)h''(\ell)}{v'(\bar{z})},$$

leads to equation (31). Note finally that if the utility function is logarithmic, this expression simplifies to:

$$\left(-\frac{\ell^2}{\underline{z}}\frac{\partial^2\Pi(\theta)}{\partial\ell^2}\right)\frac{\hat{\ell}}{\ell} = -\left[\frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\bar{z})}{R'(\bar{z})} - \frac{R(\underline{z})}{R'(\underline{z})}}\right]\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \left[\frac{\frac{R(\bar{z})}{R'(\bar{z})}}{\frac{R(\bar{z})}{R'(\bar{z})} - \frac{R(\underline{z})}{R'(\underline{z})}}\right]\frac{\hat{R}'(\bar{z})}{R'(\bar{z})}$$
$$-\left[\frac{\frac{\ell}{1-\ell}\frac{R(\bar{z})}{\underline{z}R'(\bar{z})} + 1}{(1-\ell)\frac{R(\underline{z})}{R'(\underline{z})} + \ell\frac{R(\bar{z})}{R'(\bar{z})}}\right](1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \left[\frac{1 - \frac{R(\underline{z})}{\underline{z}R'(\underline{z})}}{(1-\ell)\frac{R(\underline{z})}{R'(\underline{z})} + \ell\frac{R(\bar{z})}{R'(\bar{z})}}\right]\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})}.$$

where we used again equation (34). This concludes the proof.

Extension to a locally nonlinear baseline tax schedule. Accounting for the terms involving R'' in the above Taylor expansion leads to the following more general expression for the response of labor effort to tax reforms:

$$
\begin{aligned}
D'\frac{\hat{\ell}}{\ell} =& -\frac{\ell}{\underline{z}}\left[\frac{\ell(1-\ell)h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))}\right]\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \frac{\ell}{\underline{z}}\left[\frac{\ell(1-\ell)h''(\ell)}{R'(\bar{z})u'(R(\bar{z}))}\right]\frac{\hat{R}'(\bar{z})}{R'(\bar{z})} \\
& -\frac{\ell}{\underline{z}}\left[\frac{A\cdot\frac{\frac{1}{1-\ell}}{R'(\bar{z})u'(R(\bar{z}))} + \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right](1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} \\
& +\frac{\ell}{\underline{z}}\left[\frac{A\cdot\frac{\frac{1}{\ell}}{R'(\underline{z})u'(R(\underline{z}))} - \left(\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} - \frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}\right)\ell(1-\ell)h''(\ell)}{\frac{1-\ell}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{R'(\bar{z})u'(R(\bar{z}))}}\right]\ell\frac{\hat{R}(\bar{z})}{R'(\bar{z})}
\end{aligned}
$$

where we denote

$$
\begin{aligned}
D' =& \frac{1-2\ell}{1-\ell} + \frac{\ell h'''(\ell)}{h''(\ell)} - \frac{\frac{\frac{\ell}{1-\ell}R''(\underline{z})}{(R'(\underline{z}))^3(u'(R(\underline{z})))^2} + \frac{R''(\bar{z})}{(R'(\bar{z}))^3(u'(R(\bar{z})))^2}}{\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}} + \ell h''(\ell)\times \\
& + \left(\frac{\ell}{\underline{z}}\frac{1-\ell}{R'(\bar{z})u'(R(\bar{z}))} + \frac{\ell}{\underline{z}}\frac{\ell}{R'(\underline{z})u'(R(\underline{z}))} - \frac{\ell\frac{u''(R(\underline{z}))}{R'(\underline{z})(u'(R(\underline{z})))^3} + (1-\ell)\frac{u''(R(\bar{z}))}{R'(\bar{z})(u'(R(\bar{z})))^3}}{\frac{1}{R'(\bar{z})u'(R(\bar{z}))} - \frac{1}{R'(\underline{z})u'(R(\underline{z}))}}\right),
\end{aligned}
$$

and

$$
A = 1 - \left(\frac{\ell R''(\underline{z})}{(R'(\underline{z}))^2 u'(R(\underline{z}))} + \frac{(1-\ell)R''(\bar{z})}{(R'(\bar{z}))^2 u'(R(\bar{z}))}\right).
$$

Note that this term appears in the second and third line of the right hand side of the expression for $\hat{\ell}/\ell$, which is otherwise identical to the formula derived above. ∎

Proof of Lemma 10. Substituting expression (30) into (28) and (29) implies

$$
\begin{aligned}
\hat{\underline{z}} &= \left[\frac{\frac{1-\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}} - 1\right]\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \frac{\frac{\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\bar{z})}{R'(\bar{z})} - \frac{\ell h''(\ell)}{v'(\underline{z})}\hat{\ell} \\
\hat{\bar{z}} &= \frac{\frac{1-\ell}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \left[\frac{\frac{\ell}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}} - 1\right]\frac{\hat{R}(\bar{z})}{R'(\bar{z})} + \frac{(1-\ell)h''(\ell)}{v'(\bar{z})}\hat{\ell}.
\end{aligned}
$$

This system can be rewritten as follows:

$$\hat{\underline{z}} = -\left[1 - \frac{\frac{1-\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\right] \frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \frac{1}{1-\ell} \frac{\frac{1-\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}} \ell \frac{\hat{R}(\bar{z})}{R'(\bar{z})} - \frac{1}{1-\ell} \frac{\ell(1-\ell)h''(\ell)}{v'(\underline{z})} \hat{\ell}$$

$$\hat{\bar{z}} = \frac{1}{\ell}\left[\frac{\frac{\ell}{v'(\bar{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\right](1-\ell)\frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \left[\frac{\frac{1-\ell}{v'(\underline{z})}}{\frac{1-\ell}{v'(\underline{z})} + \frac{\ell}{v'(\bar{z})}}\right] \frac{\hat{R}(\bar{z})}{R'(\bar{z})} + \frac{1}{\ell}\left[\frac{\ell(1-\ell)h''(\ell)}{v'(\bar{z})}\right]\hat{\ell}.$$

Defining ε_{out} and ε_{in} as in the text, and noting that, by equation (34),

$$\frac{\ell(1-\ell)h''(\ell)}{v'(\bar{z})} = \frac{\ell(1-\ell)h''(\ell)}{v'(\underline{z})} + \theta - b$$

leads to expressions (32) and (33). ■

C Alternative Models of Performance Pay

C.1 Linear Contracts: Piece Rates and Commissions

Preferences are represented by the utility function $U(c,\ell) = -\frac{1}{\gamma}\exp(-\gamma(c-h(\ell)))$, where h is convex. The income tax schedule is affine: $c = T_0 + (1-\tau)z$. Providing effort ℓ yields output $\theta(\ell+\eta)$, where $\eta \sim \mathcal{N}(0,\sigma_\eta^2)$. The firm observes the worker's output but not her effort nor performance shock. Following Holmstrom and Milgrom (1987), we can restrict attention to linear contracts, i.e., pre-tax earnings are given as a function of observed output by $z = z_0 + \beta\theta(\ell+\eta)$, for some $(z_0,\beta) \in \mathbb{R}^2$. The firm maximizes expected profits $\theta\ell - \mathbb{E}[z|\theta]$ subject to the incentive constraint

$$\ell = \arg\max_{l \geq 0} \mathbb{E}[U(c,l)|\theta] \tag{36}$$

and the participation constraint $\mathbb{E}[U(c,\ell)|\theta] \geq U(\theta)$. The free-entry condition holds and determines the equilibrium reservation values.

We show below that the incentive compatibility constraint (36) implies $h'(\ell) = (1-\tau)\beta\theta$. In other words, if the firm wants to elicit an effort level ℓ from the worker, it must design a contract such that the sensitivity of pay to performance is equal to

$$\beta = \frac{1}{\theta}\frac{h'(\ell)}{1-\tau}.$$

This equation shows the worker's exposure to output risk, measured by the slope of the equilibrium contract, has a similar expression as in our baseline model, and identical crowd-out and crowd-in elasticities $\varepsilon_{\beta,1-p} = -1$ and $\varepsilon_{\beta,\ell} = 1/\varepsilon_\ell^F$. In Section C.4 in the Appendix we derive expressions for the demogrant z_0 and the equilibrium expected utility $U(\theta)$.

The optimal effort level is chosen to maximize the firm's profit. We find that ℓ satisfies

$$h'(\ell) = \frac{(1-\tau)\theta}{1+\gamma h''(\ell)\sigma_\eta^2}. \tag{37}$$

Suppose in particular that $h(\ell) = \frac{\ell^2}{2}$. We then get $\beta = \frac{1}{\theta}\frac{\ell}{1-\tau}$ and $\ell = \frac{(1-\tau)\theta}{1+\gamma\sigma_\eta^2}$. Thus, $\beta = \frac{1}{1+\gamma\sigma_\eta^2}$ is independent of the tax rate. More generally, the net effect of the tax rate on the pass-through is given by

$$\frac{d\ln\beta}{d\ln(1-\tau)} = -1 + \frac{\varepsilon_{\ell,1-\tau}}{\varepsilon_\ell^F},$$

and the elasticity of labor effort with respect to the retention rate $1-\tau$ is given by

$$\varepsilon_{\ell,1-\tau} = \frac{\partial\ln\ell}{\partial\ln(1-\tau)} = \frac{\varepsilon_\ell^F}{1+(1-\beta)\frac{h'(\ell)h'''(\ell)}{h''(\ell)^2}},$$

where $\varepsilon_\ell^F = \frac{h'(\ell)}{\ell h''(\ell)}$ is Frisch elasticity. These expressions imply that an increase in the tax rate leads to an increase in the pass through β, so that the crowd-out dominates the crowd-in, if and only if the labor effort elasticity $\varepsilon_{\ell,1-\tau}$ is smaller than the Frisch elasticity ε_ℓ^F, or equivalently whenever $h'''(\ell) > 0$. If the disutility of effort is isoelastic, this is the case iff $\varepsilon_\ell^F < 1$.

The framework of Holmstrom and Milgrom (1987) allows us to verify that our main prediction—the offsetting of the crowd-out and crowd-in effects—is robust to the degree of risk aversion of workers. Suppose that the Frisch elasticity is constant, in which case the effort elasticity becomes

$$\varepsilon_{\ell,1-\tau} = \frac{\varepsilon_\ell^F}{1+(1-\beta)(1-\varepsilon_\ell^F)}.$$

This elasticity depends on $\beta = \frac{1}{\theta}\frac{h'(\ell)}{1-\tau}$, which is increasing in the level of effort. Note further that by the first-order condition (37), effort is strictly decreasing in the coefficient of absolute risk aversion γ. Intuitively, motivating effort requires exposing workers to earnings risk, and more risk-averse workers require higher compensation for this risk—a higher z_0—which is costly to the firm. Thus, the firm optimally chooses a lower level of effort when γ is higher. This comparative statics allows us to sharply characterise how risk aversion affects both the labor effort elasticity and the degree to which the crowd-in offsets the crowd-out.

Proposition 4 *The effort elasticity $\varepsilon_{\ell,1-\tau}$ is a monotonic function of the coefficient of absolute risk aversion γ, and takes values between ε_ℓ^F when $\gamma = 0$ and $\frac{\varepsilon_\ell^F}{2-\varepsilon_\ell^F}$ when $\gamma \to \infty$. Thus, $\frac{d\ln\beta}{d\ln(1-\tau)}$ takes values between 0 when $\gamma = 0$ and $-\frac{1-\varepsilon_\ell^F}{2-\varepsilon_\ell^F}$ when $\gamma \to \infty$.*

Suppose that, in line with the existing evidence, Frisch elasticity is equal $\varepsilon_\ell^F = 0.5$. We know that, regardless of the degree of risk aversion, the crowd-in will offset at least two-thirds of the crowd-out: $\frac{d\ln\beta}{d\ln(1-\tau)} > -1/3$. Furthermore, the lower is coeff. of absolute risk aversion γ, the higher is this offset rate, reaching 100 percent when the risk aversion vanishes.

C.2 Convex Contracts: Stock-Options

We now build on the model of performance pay proposed by Edmans and Gabaix (2011). This framework gives rise to convex optimal contracts and has been used to describe forms of executive compensation such as stock options. Here, we focus on a simple version of the model, and we refer to our earlier Working Paper (Doligalski, Ndiaye, and Werquin 2020) for a thorough analysis of taxation a general environment that allows for arbitrary utility function, distribution of performance shocks, and tax schedule.

The setup is similar to our baseline model of Section 2, except that agents can now draw continuous performance shocks. A worker with ability θ who provides effort ℓ produces output $\theta(\ell + \eta)$, where $\eta \in \mathbb{R}$ is a random variable with mean 0. As in Edmans and Gabaix (2011), we impose the following assumption.

Assumption 1 *The agent chooses effort ℓ* after *observing the realization of her performance shock η. The firm recommends the same effort level $\ell(\theta)$ for all agents with the same ability θ.*

Importantly, we assume here that the worker is committed to stay with an employer regardless of the realisation of the performance shock. Since the design of the contract ensures that effort is incentive compatible, the firm is able to infer the underlying type η from the worker's output. We thus denote the earnings schedule by $z(\theta, \eta)$. The firm's problem is to maximize expected profit (1) subject to the participation constraint (3) and the incentive compatibility constraint, which reads:

$$\ell(\theta) \in \arg\max_{\hat{\ell}} u(R(z(\theta, \eta + \hat{\ell} - \ell(\theta)))) - h(\hat{\ell}), \quad \forall \eta. \tag{38}$$

That is, when the worker exerts effort $\hat{\ell}$, the employer assumes that she has exerted the recommend effort $\ell(\theta)$ and deduces that η is $\eta + \hat{\ell} - \ell(\theta)$ and pays her according to that calculation. Incentive compatibility then implies that $\hat{\ell} = \ell(\theta)$ is optimal. Notice that, in contrast to our baseline framework of Section 2, the effort level $\ell(\theta)$ must maximize utility state-by-state (i.e., for each performance shock realization η) rather than in expectation. This is a consequence of the timing Assumption 1. Finally, the free-entry condition (4) holds.

Assumption 2 *The utility of consumption is logarithmic, $u(c) = \log c$. The Frisch elasticity of labor supply $\varepsilon_\ell^F \equiv h'(\ell)/\ell h''(\ell)$ is constant. The performance shocks are normally distributed, $\eta \sim \mathcal{N}(0, \sigma_\eta^2)$. The tax schedule has a constant rate of progressivity (CRP), $T(z) = z - \frac{1-\tau}{1-p} z^{1-p}$.*

We denote by $\beta \equiv \partial \log z(\theta, \eta)/\partial \eta$ the pass-through of performance shocks to log-earnings. The following proposition characterizes the equilibrium labor contract.

Lemma 11 *The earnings schedule is log-linear and given by:*

$$\log z(\theta, \eta) = \log(\theta \ell) + \beta \eta - \frac{1}{2}\beta^2 \sigma_\eta^2 \quad \text{with} \quad \beta = \frac{h'(\ell)}{1-p}. \tag{39}$$

Effort ℓ is independent of θ and satisfies:

$$\ell = [(1-p)(1 - \varepsilon_{\beta,\ell}\beta^2 \sigma_\eta^2)]^{\varepsilon_\ell^F/(1+\varepsilon_\ell^F)}, \tag{40}$$

where $\varepsilon_{\beta,\ell} \equiv \frac{\partial \log \beta}{\partial \log \ell} = 1/\varepsilon_\ell^F$. Expected utility is given by

$$U(\theta) = \log(R(\theta \ell)) - h(\ell) - \frac{1}{2}(1-p)\beta^2 \sigma_\eta^2. \tag{41}$$

Lemma 11 shows that earnings risk, measured by the pass-through parameter β, is constant and has the exact same expression as in our discrete model (equation (7)), namely $\beta = h'(\ell)/(1-p)$. As in Section 2, this property follows immediately by taking the first-order condition in the incentive compatibility constraint (38). This implies in turn that the crowd-out and crowd-in elasticities are given by $\varepsilon_{\beta,1-p} = -1$ and $\varepsilon_{\beta,\ell} = 1/\varepsilon_\ell^F$. Proposition 1 and the subsequent discussion on the relative magnitude of these two forces thus applies identically to this framework. Only the expression for the labor effort elasticity is different, namely,

$$\varepsilon_{\ell,1-p} = \frac{\varepsilon_\ell^F}{1+\varepsilon_\ell^F} \cdot \frac{1+\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2}{1+\frac{1-\varepsilon_\ell^F}{1+\varepsilon_\ell^F}\varepsilon_{\beta,\ell}\,\beta^2\sigma_\eta^2}.$$

This expression shows that the labor effort elasticity is strictly larger in the presence of moral hazard ($\varepsilon_{\beta,\ell} > 0$) than in the benchmark model with exogenous risk ($\varepsilon_{\beta,\ell} = 0$), due to the marginal cost of incentives (MCI) in the first-order condition for effort.

We can now derive the optimal rate of progressivity in this framework. We obtain the following result.

Proposition 5 *Suppose that the social welfare objective is utilitarian. The optimal rate of progressivity satisfies*

$$\frac{p}{(1-p)^2} = \frac{\sigma_\theta^2 + (1+\varepsilon_{\beta,1-p})\beta^2\sigma_\eta^2}{\left[1+\frac{s}{(1-s)p}\right]\varepsilon_{\ell,1-p} + (1-p)\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}\beta^2\sigma_\eta^2}. \tag{42}$$

Thus, the optimal rate of progressivity is strictly smaller in the model with endogenous private insurance than in the benchmark environment with exogenous risk where $\varepsilon_{\beta,1-p} = \varepsilon_{\beta,\ell} = 0$.

Interestingly, the optimal tax progressivity in our baseline setting (equation (13)) coincides with formula (42) up to a second order as $\beta \to 0$.

C.3 Dynamic Contracts: Career Incentives

We now extend our results to a dynamic model of the labor market based on the model of Edmans, Gabaix, Sadzik, and Sannikov (2012).[30] Workers are indexed by their constant productivity θ. They live for $S \geq 2$ periods and discount the

[30] Our results of Section 2 also extend to the dynamic framework of Sannikov (2008), in which the one-shot deviation principle implies that the sensitivity of utility to output shocks is, again, given by the marginal disutility of effort $h'(\ell)$ (see equation (4) on p. 962).

future at rate r. Preferences are separable, logarithmic in consumption and isoelastic in effort. Productivity θ is lognormally distributed with mean μ_θ and variance σ_θ^2. The government levies a CRP income tax given by $R_t(z) = \frac{1-\tau_t}{1-p} z^{1-p}$. The rate of progressivity p is time-independent while the intercept τ_t ensures that the budget is balanced in each period. Private savings are ruled out, so that $c_t = R_t(z_t)$.

We denote the history of a random variable x up to time $t \leq S$ by x^t. Flow output at time t is given by $y_t = \theta(\ell_t + \eta_t)$ where $\{\eta_t\}_{1\leq t\leq S}$ are i.i.d. random variables. We assume that η_t are normally distributed with mean 0 and variance σ_η^2. As in Section C.2, we assume that the agent chooses period-t effort ℓ_t after observing the realization of the history of performance shocks up to and including time t, η^t. Firms discount future profits at rate r. In each period they observe the agent's productivity and history of output realizations. A labor contract specifies for each t a recommended effort level $\ell_t(\theta)$ and an earnings function $z_t(\theta, \eta^t)$. The firm maximizes its expected profit

$$\Pi(\theta) = \max_{\{\ell_t(\theta), z_t(\theta,\eta^t)\}_{1\leq t\leq S}} \sum_{t=1}^{S} \left(\frac{1}{1+r}\right)^{t-1} \left\{\theta \ell_t - \mathbb{E}_0\left[z_t(\theta, \eta^t)\right]\right\}$$

subject to the incentive constraint:

$$\begin{aligned} & \mathbb{E}_1\left[\sum_{t=1}^{S} \beta^{t-1}(u(R_t(z_t(\theta,\eta^t))) - h(\tilde{\ell}_t(\eta^t)))\right] \\ \leq\ & \mathbb{E}_1\left[\sum_{t=1}^{S} \beta^{t-1}(u(R_t(z_t(\theta,\eta^t))) - h(\ell_t(\theta)))\right], \quad \forall \{\tilde{\ell}_t(\eta^t)\}_{1\leq t\leq S} \end{aligned} \tag{43}$$

and the participation constraint:

$$\mathbb{E}_0\left[\sum_{t=1}^{S} \beta^{t-1}(u(R(z_t(\theta,\eta^t))) - h(\ell_t(\theta)))\right] \geq U(\theta).$$

The free-entry condition (4) holds.

Lemma 12 *Let $\sum_{s=0}^{S-t}\left(\frac{1}{1+r}\right)^s \equiv 1/\delta_t$, and denote the present value of effort by $L \equiv \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\ell_s$. Define the sequence of pass-through parameters $\{\beta_t\}_{1\leq t\leq S}$ by*

$$\beta_t = \delta_t \frac{h'(\ell_t)}{1-p}. \tag{44}$$

The earnings schedule satisfies

$$\log(z_t(\theta, \eta^t)) = \log(z_{t-1}(\theta, \eta^{t-1})) + \beta_t \eta_t - \frac{1}{2}\beta_t^2 \sigma_\eta^2, \tag{45}$$

where initial earnings are given by $z_0 \equiv \delta_1 \theta L$. *Period-t effort level* ℓ_t *is independent of* θ *and satisfies*

$$\ell_t = \left[(1-p)\left(\frac{\ell_t}{\delta_1 L} - \frac{1}{\delta_t}\varepsilon_{\beta_t,\ell_t}\beta_t^2\sigma_\eta^2\right)\right]^{\varepsilon_\ell^F/(1+\varepsilon_\ell^F)}$$

where $\varepsilon_{\beta_t,\ell_t} = 1/\varepsilon_\ell^F$ *is the elasticity of the pass-through parameter* β_t *with respect to effort* ℓ_t. *Expected utility is given by*

$$U(\theta) = \sum_{t=1}^{S}\left(\frac{1}{1+r}\right)^{t-1}\left[u(R(\delta_1\theta L)) - h(\ell_t) - \frac{1}{2\delta_t}\beta_t^2\sigma_\eta^2\right].$$

Equation (45) shows that, as in the static setting of Section C.2, earnings in each period t are a log-linear function of the performance shock η_t in that period. Note that $\delta_S = 1$ in the last period, so that β_S is exactly the same as in the static model. In earlier periods we have $\delta_t < 1$ for all $t \leq S-1$, so that the pass-through of output risk is smaller than in the static environment. This is because an increase in output realization in a given period, either due to effort or to random shocks, boosts log-earnings in the current and all future periods equally. Indeed, since the agent is risk-averse it is efficient to spread the rewards over her entire horizon. In other words, a given increase in lifetime utility necessary to elicit higher effort requires a higher increase in flow utility if there are fewer remaining periods over which to smooth these benefits. As a result, the sequence $\{\delta_t\}_{1\leq t\leq S}$ is strictly increasing and the degree of performance pay gets stronger over time. Nevertheless, the pass-through of performance shocks to log-earnings β_t keeps the same expression as in the static model. Thus, our insight that tax progressivity affects the private contract via offsetting crowd-out and crowd-in forces carries over to this dynamic environment.

Proposition 6 *Suppose that the planner is utilitarian. The optimal rate of progressivity is given by*

$$\frac{p}{(1-p)^2} = \frac{\sigma_\theta^2}{\varepsilon_{L,1-p} + (1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{\delta_1}{\delta_s}\varepsilon_{\beta_s,\ell_s}\varepsilon_{\ell_s,1-p}\beta_s^2\sigma_\eta^2} \tag{46}$$

where $\varepsilon_{L,1-p}$ is the elasticity of the present discounted value of effort with respect to progressivity, and $\varepsilon_{\beta_s,\ell_s} = 1/\varepsilon_\ell^F$.

Equation (46) is similar to its static counterpart (42). Assuming first that private insurance is exogenous ($\varepsilon_{\beta_s,\ell_s} = \varepsilon_{\beta_s,1-p} = 0$ for all $s \geq 1$), note that the relevant labor effort elasticity is that of the present-value of effort, $\varepsilon_{L,1-p}$. With endogenous earnings risk, the optimal rate of progressivity accounts for the negative fiscal externality due to the crowding-in of private insurance (second term in the denominator). The only difference with the static expression is that the relevant discount factor is not $(1/(1+r))^{s-1}$ but $(1/(1+r))^{s-1}\delta_1/\delta_s$. Since δ_s is increasing over time, this implies that the fiscal externalities caused by the future performance-pay effects are discounted at a higher rate than the standard deadweight losses from distorting effort.

C.4 Proofs of Section C

Proofs of Section C.1. The incentive constraint reads

$$\ell = \arg\max_l -\frac{1}{\gamma}\mathbb{E}\left[e^{-\gamma[T_0+(1-\tau)(z_0+\beta\theta(l+\eta))-h(l)]}|\theta\right]$$

Taking the first-order condition implies

$$\mathbb{E}\left[\{(1-\tau)\beta\theta - h'(\ell)\}e^{-\gamma[T_0+(1-\tau)(z_0+\beta\theta(\ell+\eta)-h(\ell)]}|\theta\right] = 0$$

and hence

$$(1-\tau)\beta\theta = h'(\ell).$$

The slope of the optimal contract is thus given by $\beta = \frac{1}{\theta}\frac{h'(\ell)}{1-\tau}$. Expected utility is given by

$$\begin{aligned}
\mathbb{E}[U(c,\ell)|\theta] =& \mathbb{E}\left[-\frac{1}{\gamma}e^{-\gamma[T_0+(1-\tau)z_0+(1-\tau)\beta\theta(\ell+\eta)-h(\ell)]}|\theta\right] \\
=& -\frac{1}{\gamma}e^{-\gamma T_0}e^{-\gamma(1-\tau)z_0}e^{-\gamma(\ell h'(\ell)-h(\ell))}\mathbb{E}\left[e^{-\gamma h'(\ell)\eta}|\theta\right] \\
=& -\frac{1}{\gamma}e^{-\gamma T_0}e^{-\gamma(1-\tau)z_0}e^{-\gamma(\ell h'(\ell)-h(\ell))}e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2}.
\end{aligned}$$

The participation constraint then implies

$$z_0 = -\frac{\log(-\gamma U(\theta))}{\gamma(1-\tau)} - \frac{T_0}{1-\tau} - \frac{\ell h'(\ell) - h(\ell)}{1-\tau} + \frac{1}{2}\frac{\gamma}{1-\tau}(h'(\ell))^2\sigma_\eta^2$$

Free entry implies $0 = (1-\beta)\theta\ell - z_0$, and hence

$$z_0 = (1-\beta)\theta\ell.$$

Thus expected utility is equal to

$$\begin{aligned} U(\theta) =& -\frac{1}{\gamma}e^{-\gamma T_0}e^{-\gamma(1-\tau)z_0}e^{-\gamma(\ell h'(\ell)-h(\ell))}e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2} \\ =& -\frac{1}{\gamma}e^{-\gamma(1-\tau)(1-\beta)\theta\ell}e^{-\gamma T_0}e^{-\gamma(\ell h'(\ell)-h(\ell))}e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2} \\ =& -\frac{1}{\gamma}e^{-\gamma(1-\tau)\theta\ell}e^{\gamma\ell h'(\ell)}e^{-\gamma T_0}e^{-\gamma(\ell h'(\ell)-h(\ell))}e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2} \\ =& -\frac{1}{\gamma}e^{-\gamma[T_0+(1-\tau)\theta\ell-h(\ell)]}e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2} = U(R(\theta\ell),\ell)e^{\frac{1}{2}\gamma^2(h'(\ell))^2\sigma_\eta^2}. \end{aligned}$$

Firm profits are given by

$$\begin{aligned} \Pi =&(1-\beta)\theta\ell - z_0 \\ =& \left(1 - \frac{1}{\theta}\frac{h'(\ell)}{1-\tau}\right)\theta\ell + \frac{\ell h'(\ell)-h(\ell)}{1-\tau} - \frac{1}{2}\frac{\gamma}{1-\tau}(h'(\ell))^2\sigma_\eta^2 + \frac{\log(-\gamma U(\theta))}{\gamma(1-\tau)} + \frac{T_0}{1-\tau}. \end{aligned}$$

The optimal choice of effort maximizes the firm's profits:

$$0 = \theta\left(1 - \frac{1}{\theta}\frac{h'(\ell)}{1-\tau}\right) - \frac{\ell h''(\ell)}{1-\tau} + \frac{\ell h''(\ell)}{1-\tau} - \frac{\gamma}{1-\tau}h'(\ell)h''(\ell)\sigma_\eta^2$$

so that

$$h'(\ell) = \frac{(1-\tau)\theta}{1+\gamma h''(\ell)\sigma_\eta^2}.$$

This first-order condition implies that

$$\frac{\partial \ell}{\partial 1-\tau} = \frac{\theta}{h''(\ell) + \gamma\sigma_\eta^2(h''(\ell)^2 + h'(\ell)h'''(\ell))}$$

which leads to the following effort elasticity

$$\varepsilon_{\ell,1-\tau} = \frac{\partial \ln \ell}{\partial \ln(1-\tau)} = \frac{\varepsilon_\ell^F}{1+(1-\beta)\frac{h'(\ell)h'''(\ell)}{h''(\ell)^2}},$$

where $\varepsilon_\ell^F = \frac{h'(\ell)}{\ell h''(\ell)}$ is the Frisch elasticity. Assuming that the Frisch elasticity is constant, the effort elasticity becomes $\varepsilon_{\ell,1-\tau} = \frac{\varepsilon_\ell^F}{1+(1-\beta)(1-\varepsilon_\ell^F)}$. ∎

Proof of Lemma 11. Consider first the general case of a concave utility function u and a nonlinear retention function R. Given the earnings contract $\{z(\theta,\eta) : \eta \in \mathbb{R}\}$, an agent with ability θ and performance shock η chooses effort $\ell(\theta)$ to maximize utility $v(z(\theta,\eta)) - h(\ell(\theta))$ with $v = u \circ R$. Equation (38) implies that $\frac{\partial z(\theta,\eta)}{\partial \eta} = \frac{\partial z(\theta,\eta)}{\partial \ell}$ so that the first-order condition reads

$$v'(z(\theta,\eta))\frac{\partial z(\theta,\eta)}{\partial \eta} = h'(\ell(\theta)). \tag{47}$$

This equation pins down the slope of the earnings schedule that the firm must implement in order to induce the effort level $\ell(\theta)$. Integrating this incentive constraint over η given $\ell(\theta)$ leads to

$$v(z(\theta,\eta)) = h'(\ell(\theta))\eta + k, \tag{48}$$

for some constant $k \in \mathbb{R}$. Since in equilibrium the participation constraint (3) must hold with equality, the agent's expected utility must be equal to his reservation value $U(\theta)$. Therefore, the value of k must be chosen by the firm such that the agent's participation constraint holds with equality. Imposing the participation constraint with $\mathbb{E}\eta = 0$ implies

$$k = U(\theta) + h(\ell(\theta)). \tag{49}$$

The previous two equations fully characterize the wage contract given the desired effort level $\ell(\theta)$ and the reservation value $U(\theta)$. They imply that, for a given pair $(a(\theta), U(\theta))$, the wage given performance shock η satisfies:

$$v(z(\theta,\eta)) = h'(\ell(\theta))\eta + [U(\theta) + h(\ell(\theta))]. \tag{50}$$

The first-order condition for effort is obtained by taking the first-order condition with respect to $\ell(\theta)$ in the firm's problem, taking as given the earnings contract required to satisfy the workers' incentive and participation constraints.

Suppose now that the tax schedule is CRP, so that $R(z) = \frac{1-\tau}{1-p} z^{1-p}$. Equation (48) then implies that in order to induce agents with ability θ to choose the same effort ℓ regardless of their noise realization η, the earnings contract must satisfy:

$$\log(z(\theta,\eta)) = \frac{\ell^{\frac{1}{\varepsilon}}}{1-p}\eta - \frac{1}{1-p}\log\frac{1-\tau}{1-p} + \frac{k}{1-p}, \tag{51}$$

for some $k \in \mathbb{R}$. Thus, log-earnings are linear in the performance shock $\eta = \frac{z}{\theta} - \ell$ that the firm infers upon observing realized output z. Imposing that the agent's participation constraint holds with equality pins down the value of k as a function of $U(\theta)$. Namely, equation (49) implies:

$$k = U(\theta) + \frac{1}{1+\frac{1}{\varepsilon_\ell^F}}\ell^{1+\frac{1}{\varepsilon_\ell^F}}$$

and hence

$$\log(z(\theta,\eta)) = \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\eta + \frac{1}{1-p}\frac{1}{1+\frac{1}{\varepsilon_\ell^F}}\ell^{1+\frac{1}{\varepsilon_\ell^F}} - \frac{1}{1-p}\log\frac{1-\tau}{1-p} + \frac{U(\theta)}{1-p}. \tag{52}$$

Below we derive the equilibrium value of the reservation utility $U(\theta)$ and obtain the equilibrium wage given (ℓ,η):

$$\log(z(\theta,\eta)) = \log(\theta\ell) + \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\eta - \frac{1}{2}\left(\frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\right)^2\sigma_\eta^2. \tag{53}$$

Define the sensitivity of the before-tax and after-tax wages to output in the optimal contract by the semi-elasticities $\beta(\theta,\eta) \equiv \frac{1}{z(\theta,\eta)}\frac{\partial z(\theta,\eta)}{\partial\eta}$ and $\beta^c(\theta,\eta) \equiv \frac{1}{R(z(\theta,\eta))}\frac{\partial R(z(\theta,\eta))}{\partial\eta}$, respectively. We have $\beta(\theta,\eta) = \frac{\ell^{1/\varepsilon_\ell^F}}{1-p}$ and $\beta^c(\theta,\eta) = \ell^{1/\varepsilon_\ell^F}$. Both $\beta(\theta,\eta)$ and $\beta^c(\theta,\eta)$ depend on the tax schedule through its effect on optimal effort, and there is an additional crowding-out effect on the before-tax sensitivity.

Next, since $v'(z) = \frac{R'(z)}{R(z)} = \frac{1-p}{z}$ the firm's first-order condition reads

$$\begin{aligned}\theta &= \mathbb{E}\left[\frac{h'(\ell)}{v'(z(\theta,\eta))} + \frac{h''(\ell)}{v'(z(\theta,\eta))}\eta|\theta\right] \\ &= \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\mathbb{E}[z(\theta,\eta)|\theta] + \frac{1}{\varepsilon_\ell^F}\frac{\ell^{\frac{1}{\varepsilon_\ell^F}-1}}{1-p}\mathbb{E}[z(\theta,\eta)\eta|\theta].\end{aligned}$$

We have

$$\begin{aligned}\mathbb{E}[z(\theta,\eta)|\theta] &= \mathbb{E}\big[e^{\frac{\ell^{1/\varepsilon_\ell^F}}{1-p}\eta}|\theta\big]e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}} \\ &= e^{\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2}e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}}.\end{aligned}$$

where we used the fact that that η is normally distributed with mean 0 and variance σ_η^2 so that $\mathbb{E}[e^{x\eta}] = e^{\frac{1}{2}x^2\sigma_\eta^2}$ for any x. Moreover, we have $\mathbb{E}[\eta e^{x\eta}] = x\sigma^2 e^{\frac{1}{2}x^2\sigma_\eta^2}$ for any x. Indeed, let φ the (normal) pdf of η. We have $\varphi'(\eta) = -\frac{\eta}{\sigma_\eta^2}\varphi(\eta)$, so that $\mathbb{E}[\eta e^{x\eta}] = \int \eta e^{x\eta}\varphi(\eta)d\eta = -\sigma_\eta^2\int e^{x\eta}\varphi'(\eta)d\eta = x\sigma_\eta^2\int e^{x\eta}\varphi(\eta)d\eta = x\sigma_\eta^2 e^{\frac{1}{2}x^2\sigma_\eta^2}$, where the third equality follows from an integration by parts.

$$\begin{aligned}\mathbb{E}[z(\theta,\eta)\eta|\theta] &= \mathbb{E}\big[\eta e^{\frac{\ell^{1/\varepsilon_\ell^F}}{1-p}\eta}|\theta\big]e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}} \\ &= \frac{\ell^{1/\varepsilon_\ell^F}}{1-p}\sigma_\eta^2 e^{\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2}e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}}.\end{aligned}$$

Plugging these expressions into the firm's first order condition leads to

$$\theta\ell = \left[\frac{\ell^{1+1/\varepsilon_\ell^F}}{1-p} + \frac{1}{\varepsilon_\ell^F}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2\right]e^{\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2}e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}}$$

and hence

$$\frac{\ell^{1+\frac{1}{\varepsilon_\ell^F}}}{1-p} + \frac{1}{\varepsilon_\ell^F}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2 = \theta\ell e^{-\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}-\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2+\frac{1}{1-p}\log\frac{1-\tau}{1-p}-\frac{U(\theta)}{1-p}}.$$

Now use the free-entry condition and the expression derived above for $\mathbb{E}[z(\theta,\eta)|\theta]$ to get

$$e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}+\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}} = \theta\ell. \tag{54}$$

Combining this equation with the first-order condition for optimal effort therefore leads to:

$$\ell^{1+1/\varepsilon_\ell^F} + \frac{1}{\varepsilon_\ell^F}\frac{\ell^{2/\varepsilon_\ell^F}}{1-p}\sigma_\eta^2 = 1-p. \tag{55}$$

Using the definition $\beta \equiv \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}$ for the pass-through easily leads to (40). Note that if $\varepsilon_\ell^F = 1$, we obtain optimal effort in closed form:

$$\ell = \left(\frac{1}{1-p} + \frac{\sigma_\eta^2}{(1-p)^2}\right)^{-1/2}. \tag{56}$$

Taking logs in equation (54) easily leads to (41).

Differentiating equation (55) with respect to $(1-p)$ leads to

$$\left[\left(1+\frac{1}{\varepsilon_\ell^F}\right)\ell^{1/\varepsilon_\ell^F} + \frac{2\sigma_\eta^2}{(1-p)(\varepsilon_\ell^F)^2}\ell^{2/\varepsilon_\ell^F-1}\right]\frac{\partial \ell}{\partial(1-p)} - \frac{\sigma_\eta^2}{(1-p)^2\varepsilon_\ell^F}\ell^{2/\varepsilon_\ell^F} = 1,$$

and hence

$$\left[\left(1+\frac{1}{\varepsilon_\ell^F}\right)\ell^{1/\varepsilon_\ell^F+1} + \frac{2\sigma_\eta^2}{(1-p)(\varepsilon_\ell^F)^2}\ell^{2/\varepsilon_\ell^F}\right]\varepsilon_{\ell,1-p} - \frac{\sigma_\eta^2}{(1-p)\varepsilon_\ell^F}\ell^{2/\varepsilon_\ell^F} = 1-p.$$

Using the first-order condition again to substitute for $1-p$ leads to

$$\varepsilon_{\ell,1-p} = \frac{\ell^{1/\varepsilon_\ell^F+1} + \frac{2\sigma_\eta^2}{(1-p)\varepsilon_\ell^F}\ell^{2/\varepsilon_\ell^F}}{\left(1+\frac{1}{\varepsilon_\ell^F}\right)\ell^{1/\varepsilon_\ell^F+1} + \frac{2\sigma_\eta^2}{(1-p)(\varepsilon_\ell^F)^2}\ell^{2/\varepsilon_\ell^F}}.$$

We finally express this elasticity in terms of the pass-through elasticities. We have $\beta = \frac{\ell^{1/\varepsilon_\ell^F}}{1-p}$ and $\varepsilon_{\beta,\ell} = 1/\varepsilon_\ell^F$. We can thus write

$$\varepsilon_{\ell,1-p} = \frac{\ell^{1/\varepsilon_\ell^F+1} + 2(1-p)\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2}{\left(1+\frac{1}{\varepsilon_\ell^F}\right)\ell^{1/\varepsilon_\ell^F+1} + \frac{2}{\varepsilon_\ell^F}(1-p)\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2}.$$

But the first-order condition for labor effort reads

$$\ell^{1+1/\varepsilon_\ell^F} = (1-p)(1-\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2).$$

Substituting into the previous equation and rearranging terms leads to

$$\varepsilon_{\ell,1-p} = \frac{1+\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2}{\left(1+\frac{1}{\varepsilon_\ell^F}\right)+\left(\frac{1}{\varepsilon_\ell^F}-1\right)\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2}.$$

This easily yields the expression given in the text. ■

Proof of Proposition 5. Recall that the earnings schedule of agents with ability θ can be written as

$$\log(z(\theta,\eta)) = \log(\theta\ell) + \beta\eta - \frac{1}{2}(\beta\sigma_\eta)^2$$

and their expected utility as

$$U(\theta) = \log\frac{1-\tau}{1-p} + (1-p)\log(\theta\ell) - \frac{1}{2}(1-p)(\beta\sigma_\eta)^2 - h(\ell).$$

Utilitarian social welfare is therefore equal to

$$\int_\Theta U(\theta)\mathrm{d}F(\theta) = (1-p)\mu_\theta + (1-p)\log\ell - (1-p)\frac{\beta^2\sigma_\eta^2}{2} - h(\ell) + \log\frac{1-\tau}{1-p}.$$

The first-order condition for effort, taking taxes as given, reads

$$0 = \frac{\partial U(\theta)}{\partial\ell} = (1-p)\frac{1}{\ell} - (1-p)\beta\sigma_\eta^2\frac{\partial\beta}{\partial\ell} - h'(\ell).$$

Now recall that expected pre-tax and post-tax earnings are respectively given by $\mathbb{E}[z(\theta,\eta)|\theta] = \theta\ell$ and $\mathbb{E}[(z(\theta,\eta))^{1-p}|\theta] = (\theta\ell)^{1-p}e^{-\frac{p\ell^{2/\varepsilon_\ell^F}\sigma_\eta^2}{2(1-p)}}$, so that government revenue is equal to

$$\int_\Theta \mathbb{E}[R(z(\theta,\eta))|\theta]f(\theta)d\theta = \ell e^{\mu_\theta+\frac{\sigma_\theta^2}{2}} - \frac{1-\tau}{1-p}e^{-\frac{p\ell^{2/\varepsilon_\ell^F}\sigma_\eta^2}{2(1-p)}}\ell^{1-p}e^{(1-p)\mu_\theta+(1-p)^2\frac{\sigma_\theta^2}{2}}.$$

Budget balance thus requires

$$\frac{1-\tau}{1-p} = \frac{\ell e^{\mu_\theta+\frac{\sigma_\theta^2}{2}} - G}{e^{-\frac{p\ell^{2/\varepsilon_\ell^F}\sigma_\eta^2}{2(1-p)}}\ell^{1-p}e^{(1-p)\mu_\theta+(1-p)^2\frac{\sigma_\theta^2}{2}}} = \frac{(1-s)\ell e^{\mu_\theta+\frac{\sigma_\theta^2}{2}}}{e^{-\frac{p\ell^{2/\varepsilon_\ell^F}\sigma_\eta^2}{2(1-p)}}\ell^{1-p}e^{(1-p)\mu_\theta+(1-p)^2\frac{\sigma_\theta^2}{2}}}.$$

As a result, maximizing with respect to $1-p$ leads to:

$$0 = \mu_\theta + \log \ell + (1-p)\frac{1}{\ell}\frac{\partial \ell}{\partial (1-p)} - h'(\ell)\frac{\partial \ell}{\partial(1-p)} - \frac{\beta^2\sigma_\eta^2}{2} - (1-p)\beta\sigma_\eta^2\left[\frac{\partial\beta}{\partial(1-p)} + \frac{\partial\beta}{\partial\ell}\frac{\partial\ell}{\partial(1-p)}\right] + \frac{\partial \log\frac{1-\tau}{1-p}}{\partial(1-p)},$$

with

$$\frac{\partial \log\frac{1-\tau}{1-p}}{\partial(1-p)} = \frac{s}{1-s}\frac{\partial\log\ell}{\partial(1-p)} - \mu_\theta - (1-p)\sigma_\theta^2 - \log\ell + p\frac{1}{\ell}\frac{\partial\ell}{\partial(1-p)} - \left(\frac{1}{2}-p\right)\beta^2\sigma_\eta^2 + p(1-p)\beta\sigma_\eta^2\left[\frac{\partial\beta}{\partial\,(1-p)} + \frac{\partial\beta}{\partial\ell}\frac{\partial\ell}{\partial(1-p)}\right].$$

We therefore obtain

$$0 = \left[(1-p)\frac{1}{\ell} - h'(\ell) - (1-p)\beta\sigma_\eta^2\frac{\partial\beta}{\partial\ell}\right]\frac{\partial\ell}{\partial(1-p)} + p\frac{1}{\ell}\frac{\partial\ell}{\partial(1-p)} + \frac{s}{1-s}\frac{\partial\log\ell}{\partial(1-p)} - (1-p)\sigma_\theta^2 - (1-p)\beta^2\sigma_\eta^2 - (1-p)^2\beta\sigma_\eta^2\frac{\partial\beta}{\partial(1-p)} + p(1-p)\beta\sigma_\eta^2\frac{\partial\beta}{\partial\ell}\frac{\partial\ell}{\partial(1-p)}.$$

Using the first-order condition for effort leads to

$$0 = \frac{1}{1-p}\left[p + \frac{s}{1-s}\right]\varepsilon_{\ell,1-p} + p\beta^2\sigma_\eta^2\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p} - (1-p)[\sigma_\theta^2 + \psi^2\sigma_\eta^2] - (1-p)\beta^2\sigma_\eta^2\varepsilon_{\beta,1-p}.$$

Rearranging this equation leads to the result. ■

Proof of Lemma 12. We provide a heuristic proof of this proposition; the formal argument follows the same steps as in Edmans et al. (2012).

We start by showing that the earnings process $z_t(\theta, \eta^t)$ is a martingale. That is, expected period-t earnings are equal to realized period-$(t-1)$ earnings,

$$\mathbb{E}_{t-1}[z_t(\theta, \eta^{t-1}, \eta_t)] = z_{t-1}(\theta, \eta^{t-1}).$$

To see this, start from an incentive compatible allocation and consider the following variations in retained earnings and utility:

$$\hat{u}_{t-1} = v(z_{t-1}(\theta, \eta^{t-1})) - \frac{1}{1+r}\Delta$$

and

$$\hat{u}_t = v(z_t(\theta, \eta^{t-1}, \eta_t)) + \Delta$$

and $\hat{u}_s = v(z_s(\theta, \eta^s))$ for all $s \notin \{t-1, t\}$. These perturbations preserve utility and incentive compatibility since for all ℓ_{t-1},

$$\hat{u}_{t-1} - h(\ell_{t-1}) + \frac{1}{1+r}\mathbb{E}_{t-1}[\hat{u}_t] = v(z_{t-1}(\theta, \eta^{t-1})) - h(\ell_{t-1}) + \frac{1}{1+r}\mathbb{E}_{t-1}[v(z_t(\theta, \eta^{t-1}, \eta_t))].$$

The optimal allocation must be unaffected by such deviations, so that

$$0 = \arg\min_{\Delta} \mathbb{E}\left[\sum_{s=1}^{S}(1+r)^{-t}(z_s - v^{-1}(\hat{u}_s))\right].$$

The associated first-order condition evaluated at $\Delta = 0$ reads

$$\mathbb{E}\left[\frac{1}{v'(z_t(\theta, \eta^{t-1}, \eta_t))} \mid z^{t-1}\right] = \frac{1}{v'(z_{t-1}(\theta, \eta^{t-1}))}.$$

The inverse Euler equation (see Golosov et al. (2003)) holds in our setting. With log utility and a CRP tax schedule, this equation can be rewritten as

$$(1-p)\mathbb{E}[z_t(\theta, \eta^{t-1}, \eta_t) \mid z^{t-1}] = (1-p)z_{t-1}(\theta, \eta^{t-1}),$$

which yields the martingale property.

Now, assume that a unique level of effort is implemented at each time t, that this effort level is independent of past output noise, and that local incentive constraints are sufficient conditions. Consider a local deviation in effort ℓ_t after history (η^{t-1}, η_t). By incentive compatibility the effect of such a deviation on the worker's lifetime utility U should be zero,

$$\mathbb{E}_{t-1}\left[\frac{\partial U}{\partial z_t}\frac{\partial z_t}{\partial \ell_t} + \frac{\partial U}{\partial \ell_t}\right] = 0.$$

Since $\frac{\partial z_t}{\partial \ell_t} = \theta$, we obtain

$$\mathbb{E}_{t-1}\left[\frac{\partial U}{\partial z_t}\right] = -\frac{1}{\theta}\frac{\partial U}{\partial \ell_t} \tag{57}$$

Applying incentive compatibility for effort in the final period we obtain:

$$v'(z_S(\theta,\eta^S))\frac{\partial z(\theta,\eta^{S-1},\eta_S)}{\partial \eta_S} = h'(\ell_S(\theta)).$$

Fixing η^{S-1} and integrating this incentive constraint over η_S leads to

$$v(z_S(\theta,\eta^S)) = h'(\ell_S(\theta))\eta_S + g^{S-1}(\eta^{S-1})$$

for some function of past output $g^{S-1}(\eta^{S-1})$. This implies in particular that

$$\frac{\partial v(z_S(\theta,\eta^S))}{\partial \eta_{S-1}} = \frac{\partial g^{S-1}(\eta^{S-1})}{\partial \eta_{S-1}}.$$

Analogously, the incentive constraint for effort in the second to last period reads

$$v'(z_{S-1}(\theta,\eta^{S-1}))\frac{\partial z_{S-1}(\theta,\eta^{S-1})}{\partial \eta_{S-1}} + \frac{1}{1+r}v'(z_S(\theta,\eta^S))\frac{\partial z_S(\theta,\eta^S)}{\partial \eta_{S-1}} = h'(\ell_{S-1}(\theta)).$$

Integrating over η_{S-1} and using the previous equation implies

$$v(z_{S-1}(\theta,\eta^{S-1})) + \frac{1}{1+r}g^{S-1}(\eta^{S-1}) = h'(\ell_{S-1}(\theta))\eta_{S-1} + g^{S-2}(\eta^{S-2}).$$

We now want to show that $g^{S-1}(\eta^{S-1})$ is a linear function of η_{S-1}. Since the utility function is logarithmic and the tax schedule is CRP, we obtain

$$(1-p)\log(z_S(\theta,\eta^S)) = h'(\ell_S(\theta))\eta_S + g^{S-1}(\eta^{S-1}) - \log\frac{1-\tau_S}{1-p}$$

and

$$\begin{aligned}&(1-p)\log(z_{S-1}(\theta,\eta^{S-1}))\\ &= h'(\ell_{S-1}(\theta))\eta_{S-1} - \frac{1}{1+r}g^{S-1}(\eta^{S-1}) + g^{S-2}(\eta^{S-2}) - \log\frac{1-\tau_{S-1}}{1-p}.\end{aligned}$$

Now recall that the inverse Euler equation reads

$$\mathbb{E}_{S-1}[z_S(\theta,\eta^S)] = z_{S-1}(\theta,\eta^{S-1}).$$

Using the previous expressions, this equality can be rewritten as

$$\begin{aligned}
&\mathbb{E}_{S-1}\left[e^{\frac{1}{1-p}h'(\ell_S(\theta))\eta_S}\right] e^{\frac{1}{1-p}g^{S-1}(\eta^{S-1})} \\
&= \left(\frac{1-\tau_S}{1-\tau_{S-1}}\right)^{\frac{1}{1-p}} e^{\frac{1}{1-p}h'(\ell_{S-1}(\theta))\eta_{S-1}} e^{-\frac{1}{1+r}\frac{1}{1-p}g^{S-1}(\eta^{S-1})+\frac{1}{1-p}g^{S-2}(\eta^{S-2})}.
\end{aligned}$$

This in turn implies

$$\begin{aligned}
&\left(1+\frac{1}{1+r}\right) g^{S-1}(\eta^{S-1}) \\
&= h'(\ell_{S-1}(\theta))\eta_{S-1} + g^{S-2}(\eta^{S-2}) - \frac{1}{2}\frac{(h'(\ell_S(\theta)))^2}{1-p}\sigma_\eta^2 + \frac{1}{1-p}\log\frac{1-\tau_S}{1-\tau_{S-1}}.
\end{aligned}$$

Therefore, $g^{S-1}(\eta^{S-1})$, and in turn $v(z_{S-1}(\theta, \eta^{S-1}))$, is linear in η_{S-1}. Moreover, the last-period utility is linear in both η_S and η_{S-1}. By induction, we can show that the utility in each period is a linear function of the performance shock in every past period. Now suppose for simplicity that $S = 2$, $r = 0$, $\theta = 1$, so that $\delta_1 = \frac{1}{2}$ and $\delta_2 = 1$. From the arguments above we guess a log-linear specification for earnings:

$$\begin{aligned}
\log z_1 &= \beta_1\eta_1 + k_1 \\
\log z_2 &= \beta_{21}\eta_1 + \beta_2\eta_2 + k_1 + k_2.
\end{aligned}$$

The martingale property derived above requires $z_1 = \mathbb{E}_1[z_2]$, so that for all η_1, $e^{\beta_1\eta_1+k_1} = e^{\beta_{21}\eta_1+k_1}\mathbb{E}[e^{\beta_2\eta_2+k_2} \mid \eta_1]$. This requires $\beta_1 = \beta_{21}$ and $e^{-k_2} = \mathbb{E}[e^{\beta_2\eta_2} \mid \eta_1]$. Now, the total utility of the agent is given by

$$\begin{aligned}
U = (1-p)&[2\beta_1\eta_1 + \beta_2\eta_2 + 2k_1 + k_2] \\
&- h(\ell_1) - h(\ell_2) + \log\frac{1-\tau_1}{1-p} + \log\frac{1-\tau_2}{1-p}.
\end{aligned}$$

The incentive constraint for effort (57) implies

$$\beta_1 = \frac{h'(\ell_1)}{2(1-p)}, \quad \text{and} \quad \beta_2 = \frac{h'(\ell_2)}{1-p}$$

and therefore

$$k_2 = -\frac{h'(\ell_2)}{1-p} - \frac{\sigma_\eta^2}{2}\left(\frac{h'(\ell_2)}{1-p}\right)^2.$$

Replacing in the expression for log earnings leads to

$$\log z_1 = k_1' + \frac{h'(\ell_1)}{2(1-p)}\eta_1 - \frac{\sigma_\eta^2}{2}\left(\frac{h'(\ell_1)}{2(1-p)}\right)^2$$

and

$$\log z_2 = k_1' + \frac{h'(\ell_1)}{2(1-p)}\eta_1 - \frac{\sigma_\eta^2}{2}\left(\frac{h'(\ell_1)}{2(1-p)}\right)^2 + \left(\frac{h'(\ell_2)}{1-p}\right)\eta_2 - \frac{\sigma_\eta^2}{2}\left(\frac{h'(\ell_2)}{1-p}\right)^2,$$

where $k_1' \equiv k_1 + \beta_1\ell_1 - \frac{\sigma_\eta^2}{2}\beta_1^2$. This constant is pinned down by the zero profit condition $\mathbb{E}[z_1 + z_2] = \ell_1 + \ell_2$, that is, $2e^{k_1'} = \ell_1 + \ell_2$. This implies

$$k_1' = \log\frac{\ell_1 + \ell_2}{2},$$

which concludes the proof of equation (45). The expressions for optimal effort and utility are derived in the next proof. ■

Proof of Proposition 6. Recall that the earnings schedule is given by

$$\log z_1 = \log(\delta_1\theta L) + \beta_1\eta_1 - \frac{\beta_1^2\sigma_\eta^2}{2},$$
$$\log z_t = \log z_{t-1} + \beta_t\eta_t - \frac{\beta_t^2\sigma_\eta^2}{2}.$$

The expected utility of workers with productivity θ is therefore equal to

$$U(\theta) = (1-p)\left[\frac{1}{\delta_1}\log(\delta_1\theta L) - \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\frac{\beta_s^2\sigma_\eta^2}{2}\right]$$
$$- \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}h(\ell_s) + \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\log\frac{1-\tau_s}{1-p},$$

from which the expression given in the text easily follows. Thus, utilitarian social welfare is

$$\int_\Theta U(\theta)\mathrm{d}F(\theta) = (1-p)\left[\frac{1}{\delta_1}\log(\delta_1 L) + \frac{1}{\delta_1}\mu_\theta - \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\frac{\beta_s^2\sigma_\eta^2}{2}\right]$$
$$- \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1} h(\ell_s) + \sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\log\frac{1-\tau_s}{1-p}.$$

The first-order condition for optimal effort reads

$$0 = \frac{\partial U(\theta)}{\partial \ell_t} = (1-p)\left[\frac{1}{\delta_1}\frac{1}{L}\frac{\partial L}{\partial \ell_t} - \left(\frac{1}{1+r}\right)^{t-1}\frac{1}{\delta_t}\beta_t\sigma_\eta^2\frac{\partial \beta_t}{\partial \ell_t}\right] - \left(\frac{1}{1+r}\right)^{t-1} h'(\ell_t)$$
$$= (1-p)\left[\frac{1}{\delta_1}\frac{\left(\frac{1}{1+r}\right)^{t-1}\ell_t}{L} - \left(\frac{1}{1+r}\right)^{t-1}\frac{1}{\delta_t}\beta_t^2\sigma_\eta^2\varepsilon_{\beta_t,\ell_t}\right]\frac{1}{\ell_t} - \left(\frac{1}{1+r}\right)^{t-1} h'(\ell_t),$$

which easily leads to the equation given in the text. Now, the expected present value of pre-tax and post-tax earnings in period t are given by $\mathbb{E}[z_t] = \delta_1\theta L$ and

$$\mathbb{E}[z_t^{1-p}] = (\delta_1\theta L)^{1-p}\mathbb{E}\left[e^{\sum_{s=1}^{t}(1-p)\beta_s\eta_s}\right]e^{-\sum_{s=1}^{t}(1-p)\frac{\beta_s^2\sigma_\eta^2}{2}} = (\delta_1\theta L)^{1-p}e^{-p(1-p)\sum_{s=1}^{t}\frac{\beta_s^2\sigma_\eta^2}{2}}$$

respectively, so that expected government revenue in period t is equal to

$$\int_\Theta \mathbb{E}[T(z_t)]\mathrm{d}F(\theta)$$
$$= \delta_1 L e^{\mu_\theta + \frac{\sigma_\theta^2}{2}} - \frac{1-\tau_t}{1-p}(\delta_1 L)^{1-p}e^{-p(1-p)\sum_{s=1}^{t}\frac{\beta_s^2\sigma_\eta^2}{2}}e^{(1-p)\mu_\theta + (1-p)^2\frac{\sigma_\theta^2}{2}}.$$

Imposing period-by-period budget balance therefore requires

$$\frac{1-\tau_t}{1-p} = \frac{(\delta_1 L)^p e^{\mu_\theta + \frac{\sigma_\theta^2}{2}}}{e^{-p(1-p)(\sum_{s=1}^{t}\beta_s^2)\frac{\sigma_\eta^2}{2}}e^{(1-p)\mu_\theta + (1-p)^2\frac{\sigma_\theta^2}{2}}}.$$

Substituting this expression into the social welfare function $\int_\Theta U(\theta)\mathrm{d}F(\theta)$ implies that social welfare is equal to

$$\frac{1}{\delta_1}\left[\log(\delta_1 L)+\mu_\theta+(1-(1-p)^2)\frac{\sigma_\theta^2}{2}\right]-\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}h(\ell_s)$$
$$+p(1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\sum_{i=1}^{s}\frac{\beta_i^2\sigma_\eta^2}{2}-(1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\frac{\beta_s^2\sigma_\eta^2}{2}$$
$$=\frac{1}{\delta_1}\left[\log(\delta_1 L)+\mu_\theta+(1-(1-p)^2)\frac{\sigma_\theta^2}{2}\right]-\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}h(\ell_s)$$
$$-(1-p)^2\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\frac{\beta_s^2\sigma_\eta^2}{2}.$$

We can now maximize this expression with respect to $1-p$ to get

$$\sum_{s=1}^{S}\left[\frac{1}{\delta_1}\frac{1}{L}\left(\frac{1}{1+r}\right)^{s-1}-\left(\frac{1}{1+r}\right)^{s-1}h'(\ell_s)\right]\frac{\partial\ell_s}{\partial(1-p)}$$
$$-(1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\varepsilon_{\beta_s,\ell_s}\varepsilon_{\ell_s,1-p}\beta_s^2\sigma_\eta^2$$
$$=(1-p)\left[\frac{1}{\delta_1}\sigma_\theta^2+\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\beta_s^2\sigma_\eta^2\right]+(1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\varepsilon_{\beta_s,1-p}\beta_s^2\sigma_\eta^2.$$

Using the first-order condition for effort derived above to simplify the left hand side of this expression implies

$$\frac{p}{1-p}\frac{1}{\delta_1 L}\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\ell_s\varepsilon_{\ell_s,1-p}+p\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\varepsilon_{\beta_s,\ell_s}\varepsilon_{\ell_s,1-p}\beta_s^2\sigma_\eta^2$$
$$=(1-p)\left[\frac{1}{\delta_1}\sigma_\theta^2+\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}(1+\varepsilon_{\beta_s,1-p})\beta_s^2\sigma_\eta^2\right].$$

But the elasticity of the present discounted value of effort is equal to

$$\varepsilon_{L,1-p}\equiv\frac{1-p}{L}\frac{\partial\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\ell_s}{\partial(1-p)}=\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{\ell_s}{L}\varepsilon_{\ell_s,1-p}.$$

Moreover, we have $1+\varepsilon_{\beta_s,1-p}=0$. Substituting these two expressions into the previous equation and rearranging terms leads to

$$\frac{p}{(1-p)^2}\left[\frac{1}{\delta_1}\varepsilon_{L,1-p}+(1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{1}{\delta_s}\varepsilon_{\beta_s,\ell_s}\varepsilon_{\ell_s,1-p}\beta_s^2\sigma_\eta^2\right]=\frac{1}{\delta_1}\sigma_\theta^2.$$

This concludes the proof. ■