A Proofs of Section 1

Proof of Proposition 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 (7) 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 bonus rate $\beta \equiv \log(\bar{z}) - \log(\underline{z})$ is given as a function of the desired effort level ℓ by the incentive constraint (6). 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:

$$\underline{z} = \left(\frac{1-\tau}{1-p}\right)^{-1/(1-p)} e^{\frac{1}{1-p}U} e^{\frac{1}{1-p}[h(\ell)-\ell h'(\ell)]} \tag{33}$$

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 $u(\mathbb{E}z) = \log(\theta\ell)$ and $\mathbb{E}[u(z)] = (1-\ell)\log \underline{z} + \ell \log(e^{\beta}\underline{z}) = \log \underline{z} + \ell\beta$, and using expression (7), we obtain:

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

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 (6). Since the firm takes as given the worker's reservation utility U, we differentiate equation (33) 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 (8). Note that in equilibrium, we can use (7) 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 shows that the optimal effort level ℓ is independent of the worker's productivity θ. ■

Proof of Lemma 1. The first order condition (8) 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 $\frac{\partial \underline{z}}{\partial \ell} = -\ell \underline{z}\frac{h''(\ell)}{1-p}$ 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 Lemma 2. This result follows immediately from equation (6) and Lemma 1. ■

Proof of Lemma 3. Differentiating equation (9) gives

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

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 $\frac{e^{\beta}-1}{1+\ell(e^{\beta}-1)} = \frac{b}{\mathbb{E}z}$ leads to

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

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 Proposition 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 (10). ■

Proof of Theorem 1. 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 (10), we have

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

where $\log(\mathbb{E}z) - \mathbb{E}[\log z] = \log(1+\ell(e^{\beta}-1)) - \beta\ell$. 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

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

where $Z \equiv \int \mathbb{E}[z]dF(\theta)$ is aggregate income, and $C \equiv \int \mathbb{E}[R(z)]dF(\theta)$ is aggregate consumption. Under the assumptions of Theorem 1, we have

$$C = \frac{1-\tau}{1-p}\frac{1+\ell(e^{(1-p)\beta}-1)}{[1+\ell(e^{\beta}-1)]^{1-p}}\int(\theta\ell)^{1-p}dF(\theta), \tag{34}$$

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,

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

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

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

<u>*Behavioral Effect with Exogenous Private Insurance.*</u> By equation (7), in response to a change in progressivity, the income levels change (in percentage terms) by

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

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

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

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[43]

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

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

[43] 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 $\varepsilon_{\ell,1-p}\int \mathbb{E}[T'(z)z]dF(\theta)$. This is the expression we would obtain, for instance, in the full-insurance benchmark.

as:

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

With a CRP tax schedule, we can write

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

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 $\mathbb{E}c = (1-\ell)\underline{c} + \ell e^{(1-p)\beta}\underline{c}$ leads to $\frac{1}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\underline{c}}{\mathbb{E}c}$ and $\frac{e^{(1-p)\beta}}{1+\ell(e^{(1-p)\beta}-1)} = \frac{\bar{c}}{\mathbb{E}c}$. Therefore, $\frac{b}{\mathbb{E}z}$ and $\frac{\gamma}{\mathbb{E}c}$ 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:

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

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

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

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

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

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

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

where the last equality follows from the expression (34) 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

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

The first-order condition for effort (8) implies that $\frac{\partial\ell}{\partial\tau}=0$. Thus, $\frac{\partial Z}{\partial\tau}=0$ and, using expression (34), $\frac{\partial C}{\partial\tau}=-\frac{C}{1-\tau}$. Hence, we obtain

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

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

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

Optimal Rate of Progressivity. The optimal rate of progressivity is the solution to

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

That is, the optimal level of p satisfies

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

Rearranging terms, this formula can be rewritten as

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

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

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

Recall that $\mathbb{V}(\log z) = \beta^2\ell(1-\ell)$, $\frac{\bar{z}}{\mathbb{E}z} - \frac{\bar{c}}{\mathbb{E}c} = (1-\ell)(\frac{b}{\mathbb{E}z} - \frac{\gamma}{\mathbb{E}c})$, and $\frac{Z}{C} = 1 + \frac{G}{Z-G} \equiv 1 + \frac{g}{1-g}$

where g is the ratio of public expenditures G to output $Z = C + G$. We thus get

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

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

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

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

$$\begin{aligned}
\kappa_1\mathbb{V}(\log z) &= \frac{\beta\ell(1-\ell)}{1-p}\frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)} \sim \beta^2\ell(1-\ell) = \mathbb{V}(\log z) \\
\kappa_2\mathbb{V}(\log z) &= \beta\ell(1-\ell)\left[\frac{e^{\beta}-1}{1+\ell(e^{\beta}-1)} - \frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)}\right] \sim p\mathbb{V}(\log z).
\end{aligned}$$

Note that $\kappa_2 > 0$ if and only if $\frac{e^\beta-1}{1+\ell(e^\beta-1)} = \frac{e^{(1-p)\beta}-1}{1+\ell(e^{(1-p)\beta}-1)}$, which easily leads to $p > 0$.

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:

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

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

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

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

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

$$\kappa_2 = \frac{1-p}{\beta p} \left(\frac{\bar{z}_{pp} - \underline{z}_{pp}}{\mathbb{E}z_{pp}} - \frac{\bar{c}_{pp} - \underline{c}_{pp}}{\mathbb{E}c_{pp}} \right)$$

$$\kappa_3 = \frac{1-p}{p} \left(\frac{\bar{z}_{pp} - \underline{z}_{pp}}{\mathbb{E}z_{pp}} - \frac{1}{1-p} \frac{\bar{c}_{pp} - \underline{c}_{pp}}{\mathbb{E}c_{pp}} \right)$$

and the constant κ_4 is given by

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

This concludes the proof. ∎

B Proofs of Section 2

Concavity of the Utility of Earnings v. Our analysis requires that the utility of earnings $z \mapsto v(z) \equiv u(R(z))$ is concave. It is easy to show that this is equivalent to $p_1(z)p_2(z) > -\gamma(z)$ where $\gamma(z) \equiv -\frac{R(z)u''(R(z))}{u'(R(z))}$ is the agent's coefficient of relative risk aversion, and $p_1(z) \equiv \frac{1-T(z)/z}{1-T'(z)}$, $p_2(z) \equiv \frac{zT''(z)}{1-T'(z)}$ are two measures of the local rate of progressivity of the tax schedule: the parameter $p_1(z)$ is the ratio of the average and marginal retention rates, and $p_2(z)$ is (minus) the elasticity of the retention rate with respect to income. If the tax schedule has a constant rate of progressivity p (CRP), these variables are respectively equal to $\frac{1}{1-p}$ and p. When we characterize the optimal tax schedule within the CRP class, we assume that $u(c) = \log c$ which implies that $\gamma(z) = -1$. It is easy to verify that in this case the above restriction is always satisfied regardless of the value of p. ∎

Proof of Lemma 4. Denote the agent's expected utility of effort ℓ by

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

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

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

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. ■

Proof of Proposition 2. The participation constraint reads:

$$(1-\ell)v(\underline{z}, 0) + \ell v(\underline{z}, b) - h(\ell) = U(\theta),$$

and the local incentive constraint reads:

$$v(\underline{z}, b) - v(\underline{z}, 0) = h'(\ell).$$

Solving this linear system of equations for $v(\underline{z}, 0)$ and $v(\underline{z}, b)$ as functions of ℓ and $U(\theta)$ immediately delivers equations (16) and (17). 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 (16) and (17) leads to

$$v_1(\underline{z}, 0)\frac{\partial \underline{z}}{\partial \ell} = -\ell h''(\ell)$$

and

$$v_1(\underline{z}, b)\frac{\partial \underline{z}}{\partial \ell} + v_2(\underline{z}, b)\frac{\partial b}{\partial \ell} = (1-\ell)h''(\ell).$$

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

$$\theta = b - \left[1 - \ell\frac{v_1(\underline{z}, b)}{v_2(\underline{z}, b)}\right]\frac{1}{v_1(\underline{z}, 0)}\ell h''(\ell) + \frac{1}{v_2(\underline{z}, b)}\ell(1-\ell)h''(\ell).$$

Rearranging terms and noting that $\frac{v_1(\underline{z},b)}{v_2(\underline{z},b)} = \frac{R_1(\underline{z},b)}{R_2(\underline{z},b)}$ leads to equation (18). Finally, the zero-profit condition $\underline{z} + \ell b = \ell\theta$ pins down the equilibrium reservation utility $U(\theta)$. ∎

B.1 Incidence and Optimal Taxation on Total Earnings

Proof of Lemma 5. Suppose that the tax system is over total earnings, so that $R(\underline{z}, b) \equiv R(\bar{z}+b)$ for all $b \geq 0$. Equations (16) and (17), which characterize the equilibrium base pay and bonus for a given a recommended effort level ℓ and reservation utility $U(\theta)$, can then be rewritten as:

$$\begin{aligned} u(R(\underline{z})) - h(\ell) &= U(\theta) - \ell h'(\ell) \\ u(R(\bar{z})) - h(\ell) &= U(\theta) + (1-\ell)h'(\ell). \end{aligned}$$

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:

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

Linearizing this system around the initial equilibrium leads to

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

Rearranging terms and noting that $R'u' = v'$ leads to equations (19) and (20). ∎

Proof of Lemma 6. 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 (19) and (20) into this equation and rearranging terms, we obtain

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

But the first-order condition for labor effort (18) when taxes are levied on total earnings can be written as

$$\theta = b + \left[\frac{1}{v'(\bar{z})} - \frac{1}{v'(\underline{z})}\right]\ell(1-\ell)h''(\ell). \tag{35}$$

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 (21). ■

Lemma 8 *Suppose that the initial tax schedule is piecewise linear. The effect of a tax reform $\hat{R}$ on labor effort ℓ is given by:*

$$\frac{\hat{\ell}}{\ell} = \varepsilon_{\ell,R'(\underline{z})}\frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \varepsilon_{\ell,R'(\bar{z})}\frac{\hat{R}'(\bar{z})}{R'(\bar{z})} + \varepsilon_{\ell,R(\underline{z})}\frac{\hat{R}(\underline{z})}{R'(\underline{z})\underline{z}} + \varepsilon_{\ell,R(\bar{z})}\frac{\hat{R}(\bar{z})}{R'(\bar{z})\bar{z}} \tag{36}$$

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

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

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

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

where we denote:

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

In particular, if the utility function u is logarithmic, then we have $E=1$ and $\varepsilon_{\ell,R(\bar{z})}<0$ if and only if $R'(\underline{z}) > \frac{R(\underline{z})}{\underline{z}}$.

Proof of Lemma 8. The first-order condition (35) for labor effort, expressed at the perturbed tax schedule and to a first order as $\delta\to 0$, reads:

$$\begin{aligned}\theta = b+\delta\hat{b}+ \Bigg[&\frac{1}{[R'(\bar{z})+\delta(\hat{R}'(\bar{z})+R''(\bar{z})\hat{\bar{z}})]u'[R(\bar{z})+\delta(\hat{R}(\bar{z})+R'(\bar{z})\hat{\bar{z}})]}-\\ &\frac{1}{[R'(\underline{z})+\delta(\hat{R}'(\underline{z})+R''(\underline{z})\hat{\underline{z}})]u'[R(\underline{z})+\delta(\hat{R}(\underline{z})+R'(\underline{z})\hat{\underline{z}})]}\Bigg]\\ &\times(\ell+\delta\hat{\ell})(1-\ell-\delta\hat{\ell})h''(\ell+\delta\hat{\ell}).\end{aligned}$$

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 (35) 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

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

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:

$$\frac{\partial^2 \Pi(\theta)}{\partial \ell^2} \leq 0. \tag{37}$$

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

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

where we used again equation (35) with $\theta - b = \underline{z}/\ell$. We can therefore rewrite the Gateaux derivative of labor effort as

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

where, by condition (37), 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 (36). 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 (35). This concludes the proof.

<u>*Extension to a locally nonlinear baseline tax schedule.*</u> 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:

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

where we denote

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

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 Proposition 3. Substituting expression (21) into (19) and (20) 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:

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

Defining ε_{out} and ε_{in} as in the text, and noting that, by equation (35),

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

leads to expressions (25) and (26). ■

Proof of Theorem 2. Suppose that there is a top tax bracket $[z^*, \infty)$ with tax rate τ, and that the tax rate below z^* is fixed at t. Let θ_i be the lowest type with high-level pay in the top bracket, and θ_t be the lowest type whose earnings are always in the top bracket. Types $[\theta_i, \theta_t)$ are called "intermediate workers" and types $[\theta_t, \infty)$ are called "top workers". Denote the share of intermediate workers among both the top and intermediate workers by σ^I.

Consider a uniform increase in the marginal tax rate τ in the top bracket $[z^*, \infty)$ by $\delta\hat{\tau} > 0$ with $\delta \to 0$. This perturbation is represented by $\hat{R}(z) = -\hat{\tau}(z - z^*)\mathbb{I}_{\{z \geq z^*\}}$ and $\hat{R}'(z) = -\hat{\tau}\mathbb{I}_{\{z \geq z^*\}}$, so that the tax liability levied after the reform on workers with income $z > z^*$ is $T(z^*) + (\tau + \delta\hat{\tau})(z - z^*)$. Let $\hat{\underline{z}}, \hat{\overline{z}}, \hat{\ell}$ be the changes (Gateaux derivatives) in base pay, high-level pay, and effort in response to this reform.

The Gateaux derivative of a generic variable X of the form $X = \int_{\theta_1}^{\theta_2} x \frac{dF(\theta)}{F(\theta_2)-F(\theta_1)}$ is given by $\hat{X} = \int_{\theta_1}^{\theta_2} \hat{x} \frac{dF(\theta)}{F(\theta_2)-F(\theta_1)}$. We define the elasticity of X with respect to $1-\tau$, keeping the sets of intermediate and top workers fixed, by $e_X = -\frac{1-\tau}{\hat{\tau}} \frac{\hat{X}}{X}$. There are two ways to interpret this variable. First, it is the aggregate elasticity of the variable x keeping the thresholds θ_1, θ_2 fixed, i.e., ignoring the impact of the tax change on composition of the two groups—e.g., the elasticity of mean earnings of top agents, keeping the set of top agents fixed. Alternatively, we can interpret e_X as the average elasticity of x weighted by x/X for a given set of agents—e.g., the elasticity of mean earnings of top agents, weighted by their relative earnings. To see this, note that we can write $e_X = -\int_{\theta_1}^{\theta_2} \frac{x}{X} \cdot \frac{1-\tau}{\hat{\tau}} \frac{\hat{x}}{x} \frac{dF(\theta)}{F(\theta_2)-F(\theta_1)}$.

The impact of the perturbation of the top tax rate $\hat{\tau}$ on the average revenue from top workers is given by

$$dTR^T = \int_{\theta_t}^{\infty} \left\{ \hat{\tau}[\ell\overline{z} + (1-\ell)\underline{z} - z^*] + \tau[\widehat{\ell\overline{z}} + \widehat{(1-\ell)\underline{z}}] \right\} \frac{dF(\theta)}{1-F(\theta_t)}.$$

Denoting average earnings of top workers by $Z^T = \int_{\theta_t}^{\infty} [\ell\overline{z} + (1-\ell)\underline{z}] \frac{dF(\theta)}{1-F(\theta_t)}$, we can rewrite the previous expression as

$$dTR^T = \hat{\tau} \left(Z^T - z^* - \frac{\tau}{1-\tau} Z^T e_{Z^T} \right).$$

Next, the impact of the reform on the tax payments of intermediate workers is given by

$$dTR^I = \int_{\theta_i}^{\theta_t} \left\{ \hat{\tau}(\overline{z} - z^*)\ell + t[(1-\ell)\hat{\underline{z}} + \tau\ell\hat{\overline{z}} + \hat{\ell}(T(\overline{z}) - T(\underline{z}))] \right\} \frac{dF(\theta)}{F(\theta_t) - F(\theta_i)},$$

where $T(\overline{z}) - T(\underline{z}) = \tau(\overline{z} - z^*) - t(z^* - \underline{z})$. Introduce the following notation: the mean effort ℓ of intermediate workers is $L^I = \int_{\theta_i}^{\theta_t} \ell \frac{dF(\theta)}{F(\theta_t)-F(\theta_i)}$, the mean frequency-adjusted high-level pay $\ell\overline{z}$ of intermediate workers is $\overline{Z}^I = \int_{\theta_i}^{\theta_t} \ell\overline{z} \frac{dF(\theta)}{F(\theta_t)-F(\theta_i)}$, and the mean frequency-adjusted base pay $(1-\ell)\underline{z}$ of intermediated workers is $\underline{Z}^I = \int_{\theta_i}^{\theta_t}(1 -$

$\ell)\bar{z}\frac{dF(\theta)}{F(\theta_t)-F(\theta_i)}$. We can rewrite the impact on intermediate workers as

$$
\begin{aligned}
dTR^I &= \int_{\theta_i}^{\theta_t} \left\{ \hat{\tau}(\bar{z}-z^*)\ell + \tau\widehat{\ell\bar{z}} + t\widehat{(1-\ell)\underline{z}} + (t-\tau)z^*\hat{\ell} \right\} \frac{dF(\theta)}{F(\theta_t)-F(\theta_i)} \\
&= \hat{\tau}\left[\bar{Z}^I - L^I z^* - \frac{\tau}{1-\tau}\bar{Z}^I e_{\bar{Z}^I} - \frac{t}{1-\tau}\underline{Z}^I e_{\underline{Z}^I} - \frac{t-\tau}{1-\tau} z^* L^I e_{L^I} \right].
\end{aligned}
$$

The average tax revenue impact on the intermediate and top workers is then

$$
\begin{aligned}
dTR &= (1-\sigma^I)dTR^T + \sigma^I dTR^I \\
&= Z^* - (1-\sigma^I+\sigma^I L^I)z^* - \frac{\tau}{1-\tau}\left[(1-\sigma^I)Z^T e_{Z^T} + \sigma^I \bar{Z}^I e_{\bar{Z}^I}\right] \\
&- \frac{t}{1-\tau}\sigma^I \underline{Z}^I e_{\underline{Z}^I} - \frac{t-\tau}{1-\tau}\sigma^I z^* L^I e_{L^I},
\end{aligned}
$$

where $Z^* = (1-\sigma^I)Z^T + \sigma^I \overline{Z}^I$ denotes the mean frequency-adjusted earnings in the top bracket. It is easy to show that $(1-\sigma^I)Z^T e_{Z^T} + \sigma^I \bar{Z}^I e_{\bar{Z}^I} = Z^* e_{Z^*}$.

Next, denote the average welfare impact of the reform on intermediate and top workers by dW. We have

$$
\begin{aligned}
dW &= -\hat{\tau}\int_{\theta_i}^{\infty} \left\{ \ell\tilde{g}(\bar{z}\mid\theta)(\bar{z}-z^*) + (1-\ell)\tilde{g}(\underline{z}\mid\theta)(\underline{z}-z^*)\mathbb{I}_{\{\underline{z}>z^*\}} \right\} \frac{dF(\theta)}{1-F(\theta_i)} \\
&= -\hat{\tau}[Z^* - (1-\sigma^I+\sigma^I L^I)z^*]\tilde{\mathcal{G}},
\end{aligned}
$$

where $\tilde{\mathcal{G}}$ is the income-weighted average modified marginal social welfare weight in the top bracket. The top tax rate is optimal when

$$
dTR + dW = 0.
$$

Substituting the previous expressions and rearranging, we obtain

$$
\frac{\tau}{1-\tau} = \frac{1-\tilde{\mathcal{G}} - \frac{1}{Z^*-(1-\sigma^I+\sigma^I L^I)z^*}\left[\frac{t}{1-\tau}\sigma^I \underline{Z}^I e_{\underline{Z}^I} + \frac{t-\tau}{1-\tau}\sigma^I z^* L^I e_{L^I}\right]}{\rho\, e_{Z^*}}.
$$

where $\rho = \frac{Z^*}{Z^*-(1-\sigma^I+\sigma^I L^I)z^*}$ is the (observed) Pareto coefficient of the earnings distribution. Now define the earnings share of intermediate workers in the top bracket as $s^I = \frac{\sigma^I(\bar{Z}^I - L^I z^*)}{\sigma^I(\bar{Z}^I-L^I z^*)+(1-\sigma^I)(Z^T-z^*)}$ and the Pareto coefficient for intermediate earners as

$\rho^I = \frac{\bar{Z}^I/L^I}{\bar{Z}^I/L^I - z^*}$. We obtain

$$\frac{\tau}{1-\tau} = \frac{1 - \tilde{\mathcal{G}} - s^I \left(\frac{t}{1-\tau} \frac{\underline{Z}^I}{\bar{Z}^I} \rho^I e_{\underline{Z}_i} + \frac{t-\tau}{1-\tau} (\rho^I - 1) e_{L^I} \right)}{\rho e_{Z^*}}$$

which concludes the proof of equation (27).

Structural expression for $e_{\underline{Z}^I}$. By Proposition 3, the response of base pay to the tax reform $\hat{\tau}$ is given by

$$(1-\ell)\hat{\underline{z}} = -\varepsilon_{out}\ell(\overline{z} - z^*)\frac{\hat{\tau}}{1-\tau} - \overline{z}\varepsilon_{in}\hat{\ell} = -\left[\varepsilon_{out}\ell(\overline{z} - z^*) - \ell\overline{z}\varepsilon_{in}e_\ell\right]\frac{\hat{\tau}}{1-\tau},$$

where $e_\ell = -\frac{1-\tau}{\hat{\tau}}\frac{\hat{\ell}}{\ell}$ is the individual labor effort elasticity with respect to $1-\tau$. Noting that $\widehat{(1-\ell)\underline{z}} = \widehat{(1-\ell)} \cdot \underline{z} - \hat{\ell}\underline{z}$, we can express $e_{\underline{Z}^I}$ as

$$\begin{aligned} e_{\underline{Z}^I} &= -\frac{1-\tau}{\hat{\tau}}\frac{1}{\underline{Z}^I}\int_{\theta_i}^{\theta_t} \widehat{(1-\ell)\underline{z}} \frac{dF(\theta)}{F(\theta_t) - F(\theta_i)} \\ &= \frac{1}{\underline{Z}^I}\int_{\theta_i}^{\theta_t} \left\{\varepsilon_{out}(\bar{z} - z^*)\ell - \ell\bar{z}\varepsilon_{in}e_\ell + (1-\ell)\underline{z}e_{1-\ell}\right\} \frac{dF(\theta)}{F(\theta_t) - F(\theta_i)}, \end{aligned}$$

where $e_{1-\ell}$ is the individual elasticity of $1-\ell$. Define the following averages over intermediate workers: $\overline{\varepsilon_{out}}^I$ is the average crowd-out parameter, $\overline{\varepsilon_{in}e_\ell}^I$ is the average product of the crowd-in parameter and the labor effort elasticity, and $\overline{e_{1-\ell}}^I$ is the average elasticity of one minus labor effort, all appropriately weighted:

$$\begin{aligned} \overline{\varepsilon_{out}}^I &= \int_{\theta_i}^{\theta_t} \varepsilon_{out}\frac{\ell\bar{z} - \ell z^*}{\bar{Z}^I - L^I z^*}\frac{dF(\theta)}{F(\theta_t) - F(\theta_i)}, \\ \overline{\varepsilon_{in}e_\ell}^I &= \int_{\theta_i}^{\theta_t} \varepsilon_{in}e_\ell\frac{\ell\bar{z}}{\bar{Z}^I}\frac{dF(\theta)}{F(\theta_t) - F(\theta_i)}, \\ \overline{e_{1-\ell}}^I &= \int_{\theta_i}^{\theta_t} e_{1-\ell}\frac{(1-\ell)\underline{z}}{\underline{Z}^I}\frac{dF(\theta)}{F(\theta_t) - F(\theta_i)}. \end{aligned}$$

We can then rewrite the elasticity of mean base pay as

$$e_{\underline{Z}_i} = \frac{\bar{Z}^I - L^I z^*}{\underline{Z}^I}\overline{\varepsilon_{out}}^I - \frac{\bar{Z}^I}{\underline{Z}^I}\overline{\varepsilon_{in}e_\ell}^I + \overline{e_{1-\ell}}^I = \kappa_1^{-1}\left(\overline{\varepsilon_{out}}^I - \rho^I\overline{\varepsilon_{in}e_\ell}^I\right) + \overline{e_{1-\ell}}^I.$$

This concludes the proof of equation (28). ■

B.2 Full Optimum Tax Schedule

Theorem 2 provides a formula for the optimal top-bracket tax rate. We now derive and analyze a formula (à la Diamond-Saez) for the full optimal non-linear tax schedule.

Consider an arbitrary tax reform $\hat{T}$ of a given baseline tax schedule T. The effect (Gateaux derivative) of this reform on the social objective is $WE = \frac{1}{\lambda}\int \hat{U}(\theta)dF(\theta)$, where the impact on individual expected utility $\hat{U}(\theta)$ is described by Lemma 6, and λ denotes the marginal value of public funds. The effect of the reform on tax revenue is

$$\begin{aligned} RE = & \int \left[\hat{T}(\overline{z}(\theta))(1-\ell(\theta)) + \hat{T}(\underline{z}(\theta))\ell(\theta)\right] dF(\theta) \\ & + \int \left[T'(\underline{z}(\theta))\hat{\underline{z}}(\theta)(1-\ell(\theta)) + T'(\overline{z}(\theta))\hat{\overline{z}}(\theta)\ell(\theta)\right] dF(\theta) \\ & + \int \left[T(\overline{z}(\theta)) - T(\underline{z}(\theta))\right] \hat{\ell}(\theta)dF(\theta) \end{aligned}$$

where the first integral is the mechanical effect of the reform, the second integral captures the responses of base pay and high-level pay, and the third captures the frequency responses.

We now express these effects in terms of the earnings distribution. We assume throughout that $\underline{z}(\theta)$ and $\bar{z}(\theta)$ are strictly increasing in θ. We can thus change variables from θ to $z = \bar{z}(\theta)$, e.g., $\ell(z)$, $\hat{\underline{z}}(z)$, $\hat{\bar{z}}(z)$, and $\Delta T(z) \equiv T(\bar{z}(\theta)) - T(\underline{z}(\theta))$. We use a different convention whenever a variable is naturally indexed by $\underline{z}$, in which case we use a change variables from θ to $\underline{\theta}$; this is the case below for the elasticities of base earnings with respect to taxes. Denote the c.d.f. of total earnings by F_z, and the (scaled)[44] c.d.f.s of base pay and high-level pay by $F_{\underline{z}}$ and $F_{\bar{z}}$, respectively (the corresponding p.d.f.s are denoted by $f_z, f_{\underline{z}}, f_{\bar{z}}$).

We can then write the welfare effect as $WE = -\int \tilde{g}(z)\hat{T}(z)dF_z(z)$, where $\tilde{g}(z)$ is the average modified marginal social welfare weight conditional on earnings z, defined

[44]The scaling is due to the fact that they would otherwise not converge to 1 as $z \to \infty$, but rather to the average values of $1-\ell$ and ℓ in the economy, respectively.

by:

$$\tilde{g}(z) = \frac{f_{\underline{z}}(z)}{f_z(z)}\tilde{g}(z \mid \theta_{\underline{z}}) + \frac{f_{\bar{z}}(z)}{f_z(z)}\tilde{g}(z \mid \theta_{\bar{z}}),$$

where $\theta_{\underline{z}}$ and $\theta_{\bar{z}}$ follow $\underline{z}(\theta_{\underline{z}}) = \bar{z}(\theta_{\bar{z}}) = z$, and where $\tilde{g}(z \mid \theta)$ are given in Corollary 1.[45]

Next, the revenue effect of the tax reform can be written as

$$RE = \int_0^\infty \hat{T}(z)dF_z(z) + \int_0^\infty T'(z)\hat{\underline{z}}(z)dF_{\underline{z}}(z) + \int_0^\infty T'(z)\hat{\bar{z}}(z) + \Delta T(z)\frac{\hat{\ell}(z)}{\ell(z)}dF_{\bar{z}}(z).$$

Define the elasticities of the variables $x \in \{\underline{z}, \bar{z}, \ell\}$ with respect to the relevant marginal and average tax rates as

$$\frac{\hat{x}(\theta)}{x(\theta)} = \varepsilon_{x,R'(\underline{z})}(\theta)\frac{\hat{R}'(\underline{z}(\theta))}{R'(\underline{z}(\theta))} + \varepsilon_{x,R'(\bar{z})}(\theta)\frac{\hat{R}'(\bar{z}(\theta))}{R'(\bar{z}(\theta))} + \varepsilon_{x,R(\underline{z})}(\theta)\frac{\hat{R}(\underline{z}(\theta))}{R(\underline{z}(\theta))} + \varepsilon_{x,R(\bar{z})}(\theta)\frac{\hat{R}(\bar{z}(\theta))}{R(\bar{z}(\theta))}.$$

These elasticities can be obtained from Lemma 8 (see extension of the result in the proof to allow for a fully non-linear initial tax schedule) and Proposition 3. As explained above we index these elasticities by earnings, e.g., $\varepsilon_{\underline{z},R'(\bar{z})}(z)$ stands for $\varepsilon_{\underline{z},R'(\bar{z})}(\theta)$ for θ s.t. $\underline{z}(\theta) = z$. Furthermore, define the following mappings linking the two levels of pay: $\underline{x} \equiv \underline{z} \circ \bar{z}^{-1}$ and $\overline{x} \equiv \bar{z} \circ \underline{z}^{-1}$, meaning that $\underline{x}(z) = \underline{z}(\theta)$ for θ s.t. $\bar{z}(\theta) = z$, and $\overline{x}(z) = \bar{z}(\theta)$ for θ s.t. $\underline{z}(\theta) = z$.

To derive the optimal tax formula, consider a reduction of the marginal tax rate at earnings level z^* (Saez (2001)). Following Sachs, Tsyvinski, and Werquin (2020), we can express this reform as $\hat{T}'(z) = -\delta(z^* - z)$ and $\hat{T}(z) = -\mathbb{I}_{\{z > z^*\}}$, where δ is the Dirac delta function. This reform generates the following effects:

1. Welfare effect: $WE = \int_{z^*}^\infty g(z)dF_z(z)$, with marginal social welfare weights $g(z)$ defined as in Theorem 2.

2. Mechanical effect: $ME = -(1 - F_z(z^*))$.

3. Behavioral effects due to the reduction of $T'(z^*)$:

[45] The MVPF is defined as the social welfare gain from reducing every agent's tax payment by one unit: $\lambda = \frac{1}{1+\mathcal{I}} \int_0^\infty \tilde{g}(z)u'(R(z))f_z(z)dz$ where $\mathcal{I} \equiv \int_0^\infty T'(z)\frac{\partial z}{\partial R}f_z(z)dz$ is the aggregate income effect that this tax cut induces.

(a) Effect on earnings at z^*

$$BE_{z^*,R'} = \frac{T'(z^*)}{1-T'(z^*)} z^* \left[\varepsilon_{\underline{z},R'(\underline{z})}(z^*) \cdot f_{\underline{z}}(z^*) + \varepsilon_{\bar{z},R'(\bar{z})}(z^*) \cdot f_{\bar{z}}(z^*) \right]$$

Note that we can express the above as $\frac{T'(z^*)}{1-T'(z^*)} \cdot z^* \cdot \bar{\varepsilon}_{z,R'(z)}(z^*) \cdot f_z(z^*)$, where $\bar{\varepsilon}_{z,R'(z)}(z^*)$ is the average elasticity of earnings at z^* with respect to the retention rate at z^*.

(b) Effects on base pay at $\underline{x}(z^*)$ and high-level pay at $\overline{x}(z^*)$

$$\begin{aligned} BE_{z \neq z^*,R'} = & \frac{T'(\overline{x}(z^*))}{1-T'(z^*)} \overline{x}(z^*) \cdot \varepsilon_{\bar{z},R'(\underline{z})}(\overline{x}(z^*)) \cdot f_{\bar{z}}(\overline{x}(z^*)) \\ & + \frac{T'(\underline{x}(z^*))}{1-T'(z^*)} \underline{x}(z^*) \cdot \varepsilon_{\underline{z},R'(\bar{z})}(\underline{x}(z^*)) \cdot f_{\underline{z}}(\underline{x}(z^*)) \end{aligned}$$

(c) Effects on frequencies $\ell(z^*)$ and $\ell(\overline{x}(z^*))$

$$BE_{\ell,R'} = \frac{\Delta T(z^*)}{1-T'(z^*)} \varepsilon_{\ell,R'(\bar{z})}(z^*) \cdot f_{\bar{z}}(z^*) + \frac{\Delta T(\overline{x}(z^*))}{1-T'(z^*)} \varepsilon_{\ell,R'(\underline{z})}(\overline{x}(z^*)) \cdot f_{\bar{z}}(\overline{x}(z^*))$$

4. Behavioral effects due the reduction of $T(z)$ for $z > z^*$:

(a) Effects on base pay

$$BE_{\underline{z},R} = \int_{\underline{x}(z^*)}^{\infty} \frac{T'(z)}{R(\overline{x}(z))} z \cdot \varepsilon_{\underline{z},R(\bar{z})}(z) dF_{\underline{z}}(z) + \int_{z^*}^{\infty} \frac{T'(z)}{R(z)} z \cdot \varepsilon_{\underline{z},R(\underline{z})}(z) dF_{\underline{z}}(z)$$

(b) Effects on high-level pay

$$BE_{\bar{z},R} = \int_{z^*}^{\infty} \frac{T'(z)}{R(z)} z \cdot \varepsilon_{\bar{z},R(\bar{z})}(z) dF_{\bar{z}}(z) + \int_{\overline{x}(z^*)}^{\infty} \frac{T'(z)}{R(\underline{x}(z^*))} z \cdot \varepsilon_{\bar{z},R(\underline{z})}(z) dF_{\bar{z}}(z)$$

(c) Effects on frequency

$$BE_{\ell,R} = \int_{z^*}^{\infty} \frac{\Delta T(z)}{R(z)} \varepsilon_{\ell,R(\bar{z})}(z) dF_{\bar{z}}(z) + \int_{\overline{x}(z^*)}^{\infty} \frac{\Delta T(z)}{R(\underline{x}(z^*))} \varepsilon_{\ell,R(\underline{z})}(z) dF_{\bar{z}}(z)$$

Summing all the effects, equating the sum to zero and rearranging yields the Diamond-

Saez formula with performance pay:

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1-\mathcal{G}-\frac{1}{1-F_z(z^*)}\cdot[BE_{z\neq z^*,R'}+BE_{\ell,R'}+BE_{\underline{z},R}+BE_{\bar{z},R}+BE_{\ell,R}]}{\rho(z^*)\cdot\bar{\varepsilon}_{z,R'(z)}(z^*)}$$

where $\rho(z^*) = \frac{z^* f_z(z^*)}{1-F_z(z^*)}$ is the local Pareto parameter of the earnings distribution and $\mathcal{G} = \frac{\int_{z^*}^{\infty} g(z)dF_z(z)}{1-F_z(z^*)}$ is the average marginal social welfare weight of workers with earnings above z^*.

The presence of performance pay modifies the standard Diamond-Saez formula by adjusting the marginal social welfare weights and adding several new fiscal externalities. To build intuition about how these modifications affect optimal tax rates, suppose that the ability distribution is bounded with strictly positive infimum: $\theta \in [\underline{\theta}, \overline{\theta}]$, with $\underline{\theta} > 0$ and $\overline{\theta} < \infty$. As a result, the support of base pay and high-level pay do not fully overlap: sufficiently low earnings $z < \overline{z}_{lb}$ can only be reached with base pay, while sufficiently high earnings $z > \underline{z}_{ub}$ can only be reached with high-level pay. We describe how tax rates should be chosen at such extremities.[46]

For simplicity, we assume that the baseline tax schedule $T(\cdot)$ is strictly increasing and progressive: $T'(z) > \frac{T(z)}{z}$ for all $z \geq 0$, and that the marginal tax rate is (approximately) constant at the two extremities of the earnings distribution. Thus, the signs of the effort elasticities with respect to $R(\underline{z})$ and $R(\bar{z})$ are as described in Lemma 7. Consider the range of earnings which can be reached only with high-level pay (respectively, base pay). It is straightforward to show that:

1. The marginal social welfare weights are adjusted upwards (downwards) relative to the standard formula, leading to lower (higher) tax rates. Intuitively, for the high extremity, the crowd-out implies that the consumption gains from the lower taxation of high-level pay are partially transferred, via the endogenous earnings adjustment, to the base pay, where the marginal utility of consumption is higher.

2. The fiscal externalities due to frequency responses $BE_{\ell,R'}$ and $BE_{\ell,R}$ are strictly positive (negative), leading to lower (higher) tax rates.

[46] We can characterize the tax rate in the lower extremity using a slightly different (mirror image) tax reform than the Saezian reform described above, namely, a decrease in the marginal tax rate at z^* such that the level of taxes increases uniformly below z^* but is unchanged above.

3. The fiscal externality $BE_{z \neq z^*, R'}$ is negative in both cases, leading to higher tax rates.

4. The signs of the fiscal externalities $BE_{\underline{z},R}$ and $BE_{\bar{z},R}$ depend on the relative strengths of the crowding-out and the crowding-in. If they approximately offset each other, as in Section 1, then these fiscal externalities are approximately equal to zero.

Thus, as long as the crowd-out and crowd-in approximately offset each other and the impact of $BE_{z \neq z^*, R'}$ is approximately uniform across the earnings distribution, performance pay tends to reduce to optimal tax progressivity: It leads to relatively lower tax rates at high earnings levels (that can only be reached with high-level pay) and relatively higher tax rates at low earnings levels (that can only be reached with base pay). This is intuitively consistent with our results on the optimal separation of tax rates on bonuses and base pay. In Proposition 4 we demonstrated that there are efficiency gains from taxing bonuses at lower rate than base pay, since labor effort is much more responsive to the tax on bonuses. The tax rate on earnings in the upper extremity, where no one earns a base pay, is effectively a tax rate on bonuses. Thus, reducing it brings strong efficiency gains due to higher effort. A mirror image of this argument applies to the tax rate in the lower extremity of the earnings distribution, which can be understood as a tax rate on base pay and should be, consequently, set at a higher level.

In the previous argument, we restricted attention to the regions of the earnings distribution where all earnings come either exclusively from base pay or exclusively from high-level pay. We conjecture that this reasoning holds more generally when both forms of pay coincide, in which case the optimal tax rates should be set higher (lower) in the regions of the earnings distribution where base pay (high-level pay) is relatively more prevalent. Given that base pay is necessarily more prevalent than high-level pay at low earnings levels, this implies that performance pay makes the optimal tax schedule more regressive than in the standard setting, consistent with our findings regarding the optimal overall rate of progressivity (Theorem 1).

B.3 Incidence and Optimal Taxation with Separate Tax Schedules

Lemma 9 *Suppose that the tax system is separable between base pay and bonuses, so that $R(\underline{z}, b) \equiv R(\underline{z}) + P(b)$ for all $b \geq 0$. A tax reform $(\hat{R}, \hat{P})$ leads to the following incidence:*

$$\hat{\underline{z}} = -(1 - \varepsilon'_{out}) \frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \varepsilon'_{out} \ell \frac{\hat{P}(b)}{P'(b)} - b\varepsilon'_{in} \hat{\ell} \tag{38}$$

$$\ell \hat{b} = (1 - \varepsilon'_{out}) \frac{\hat{R}(\underline{z})}{R'(\underline{z})} - \varepsilon'_{out} \ell \frac{\hat{P}(b)}{P'(b)} + \left(b\varepsilon'_{in} + \frac{\underline{z}}{\ell}\right) \hat{\ell} \tag{39}$$

and

$$\frac{\hat{U}}{R'(\underline{z}) u'(R(\underline{z}))} = \varepsilon'_{out} \left[\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \ell \frac{\hat{P}(b)}{P'(b)}\right], \tag{40}$$

where the crowd-out and moral-hazard elasticities are defined by

$$\varepsilon'_{out} = \frac{\frac{1}{R'(\underline{z}) u'(R(\underline{z}))}}{\frac{1 - \frac{R'(\underline{z})}{P'(b)} \ell}{R'(\underline{z}) u'(R(\underline{z}))} + \frac{\ell}{P'(b) u'(R(\underline{z}) + P(b))}} \quad \textit{and} \quad \varepsilon'_{in} = \frac{\ell h''(\ell)}{b R'(\underline{z}) u'(R(\underline{z}))}.$$

Suppose for simplicity that the utility of consumption is logarithmic. Then the impact of the tax reform on labor effort is given by

$$\begin{aligned}
&\left(-\frac{1}{h''(\ell)} \frac{\partial^2 \Pi(\theta)}{\partial \ell^2}\right) \frac{\hat{\ell}}{\ell} \qquad (41)\\
&= \left[\left(1 - \frac{R'(\underline{z})}{P'(b)}\right) \frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}} - (1 - \ell) \frac{\frac{P(b)}{P'(b)}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}} - \frac{1}{\ell h''(\ell)} \frac{\frac{P(b)}{P'(b)}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}}\right] \frac{\hat{R}(\underline{z})}{R'(\underline{z})}\\
&+ \left[\ell \left(1 - \frac{R'(\underline{z})}{P'(b)}\right) \frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}} - \ell(1 - \ell) \frac{\frac{P(b)}{P'(b)}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}} + \frac{1}{\ell h''(\ell)} \frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell \frac{P(b)}{P'(b)}}\right] \frac{\hat{P}(b)}{P'(b)}\\
&- \frac{R(\underline{z})}{R'(\underline{z})} \frac{\hat{R}'(\underline{z})}{R'(\underline{z})} + \left[\frac{R(\underline{z})}{P'(b)} + (1 - \ell) \frac{P(b)}{P'(b)}\right] \frac{\hat{P}'(b)}{P'(b)}.
\end{aligned}$$

Proof of Lemma 9. Equations (16) and (17) can be rewritten as follows:

$$\begin{aligned} u(R(\underline{z})) - h(\ell) &= U(\theta) - \ell h'(\ell) \\ u(R(\underline{z}) + P(b)) - h(\ell) &= U(\theta) + (1-\ell)h'(\ell). \end{aligned}$$

After a perturbation $(\delta\hat{R}, \delta\hat{P})$, to a first order as $\delta \to 0$, this system becomes

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

and

$$\begin{aligned} &u[R(\underline{z} + \delta\hat{\underline{z}}) + P(b + \delta\hat{b}) + \delta(\hat{R}(\underline{z}) + \hat{P}(b))] - h(\ell + \delta\hat{\ell}) \\ =&U(\theta) + \delta\hat{U} + (1 - \ell - \delta\hat{\ell})h'(\ell + \delta\hat{\ell}). \end{aligned}$$

Linearizing this system around the initial equilibrium leads to

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

and

$$\begin{aligned} &u'(R(\underline{z}) + P(b))[\hat{R}(\underline{z}) + \hat{P}(b) + R'(\underline{z})\hat{\underline{z}} + P'(b)\hat{b}] - h'(\ell)\hat{\ell} \\ =&\hat{U} + [-h'(\ell) + (1-\ell)h''(\ell)]\hat{\ell}. \end{aligned}$$

Rearranging terms leads to

$$\hat{\underline{z}} = -\frac{\hat{R}(\underline{z})}{R'(\underline{z})} + \frac{\hat{U}}{R'(\underline{z})u'(R(\underline{z}))} - \frac{\ell h''(\ell)}{R'(\underline{z})u'(R(\underline{z}))}\hat{\ell}$$

and

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

The free-entry condition implies

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

Substituting the expressions for $\hat{\underline{z}}$ and $\hat{b}$ into this condition leads to

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

But the first-order condition for labor effort (18) can be rewritten as

$$\theta = b + \left[\frac{1}{P'(b)u'(R(\underline{z})+P(b))} - \frac{\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{1-\ell}}{R'(\underline{z})u'(R(\underline{z}))}\right]\ell(1-\ell)h''(\ell). \tag{42}$$

Thus, the Gateaux derivative of expected utility is given by

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

Substitute this expression into the above equations to get:

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

and

$$\ell\hat{b} = -\left[\frac{\frac{1}{R'(\underline{z})u'(R(\underline{z}))}}{\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{P'(b)u'(R(\underline{z})+P(b))}} - 1\right]\frac{\hat{R}(\underline{z})}{R'(\underline{z})}$$

$$= -\left[\frac{\frac{1}{R'(\underline{z})u'(R(\underline{z}))}}{\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell}{P'(b)u'(R(\underline{z})+P(b))}}\right]\ell\frac{\hat{P}(b)}{P'(b)}$$

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

Using equation (42) to substitute for $\frac{\ell(1-\ell)h''(\ell)}{P'(b)u'(R(\underline{z})+P(b))}$, the above expressions can be rewritten as (38), (39), and (40).

Suppose that the utility function is logarithmic. The crowd-out elasticity is then equal to

$$\varepsilon'_{out} = \frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\underline{z})}{R'(\underline{z})} + \ell\frac{P(b)}{P'(b)}} = \frac{p_{\underline{z}}}{p_{\underline{z}} + \frac{\ell b}{\underline{z}}p_b},$$

where $\frac{\ell b}{\underline{z}}$ is the ratio of expected bonus to base pay, and $p_{\underline{z}}, p_b$ are tax progressivity parameters defined as the ratios of average to marginal tax rates of the base pay and bonus tax schedules:

$$p_{\underline{z}} \equiv \frac{R(\underline{z})}{\underline{z}R'(\underline{z})}, \quad p_b \equiv \frac{P(b)}{bP'(b)}.$$

For more general utility functions, we can approximate the difference in inverse marginal utilities as a function of the difference in consumption levels and the risk aversion (resp., prudence) via first- (resp., second-) order Taylor expansions around $b = 0$. We get

$$\varepsilon'_{out} \approx \frac{\frac{1}{R'(\underline{z})u'(R(\underline{z}))}}{\frac{1}{R'(\underline{z})u'(R(\underline{z}))} + \frac{\ell P(b)}{P'(b)}\left[-\frac{u''(R(\underline{z}))}{(u'(R(\underline{z})))^2}\right]} = \frac{\frac{R(\underline{z})}{R'(\underline{z})}}{\frac{R(\underline{z})}{R'(\underline{z})} + \frac{\ell P(b)}{P'(b)}\left[-\frac{R(\underline{z})u''(R(\underline{z}))}{u'(R(\underline{z}))}\right]} = \frac{p_{\underline{z}}}{p_{\underline{z}} + \frac{\ell b}{\underline{z}}p_b\sigma},$$

where $\sigma \equiv -\frac{R(\underline{z})u''(R(\underline{z}))}{u'(R(\underline{z}))}$ is a risk aversion coefficient.

Next, suppose that $\underline{z}, b$ are located in locally linear portions of the tax schedule, so that $R''(\underline{z}) = P''(b) = 0$. After the perturbation, the first-order condition for effort reads

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

To a first-order as $\delta \to 0$, this implies

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

i.e.,

$$\theta = b + \ell h''(\ell)\left(\frac{1-\ell}{P'(b)u'(R(\underline{z})+P(b))}-\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})u'(R(\underline{z}))}\right)$$
$$+\delta\ell h''(\ell)\left[\left(\frac{1-2\ell}{P'(b)u'(R(\underline{z})+P(b))}-\frac{1-2\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})u'(R(\underline{z}))}\right)+\right.$$
$$\left.\left(\frac{1-\ell}{P'(b)u'(R(\underline{z})+P(b))}-\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})u'(R(\underline{z}))}\right)\frac{\ell h'''(\ell)}{h''(\ell)}\right]\frac{\hat{\ell}}{\ell}$$
$$+\delta\ell h''(\ell)\left(\frac{1-\ell\frac{R'(\underline{z})}{P'(b)}}{R'(\underline{z})}\frac{u''(R(\underline{z}))}{(u'(R(\underline{z})))^2}-\frac{1-\ell}{P'(b)}\frac{u''(R(\underline{z})+P(b))}{(u'(R(\underline{z})+P(b)))^2}\right)\left[\hat{R}(\underline{z})+R'(\underline{z})\hat{\underline{z}}\right]$$
$$-\delta\ell h''(\ell)\left(\frac{1-\ell}{P'(b)}\frac{u''(R(\underline{z})+P(b))}{(u'(R(\underline{z})+P(b)))^2}\right)\left[\hat{P}(b)+P'(b)\hat{b}\right]+\delta\hat{b}$$
$$+\delta\ell h''(\ell)\left[\frac{1}{R'(\underline{z})u'(R(\underline{z}))}\frac{\hat{R}'(\underline{z})}{R'(\underline{z})}-\left(\frac{\ell}{P'(b)u'(R(\underline{z}))}+\frac{1-\ell}{P'(b)u'(R(\underline{z})+P(b))}\right)\frac{\hat{P}'(b)}{P'(b)}\right].$$

Use the first-order condition for effort and assume that the utility is log to get

$$\left[\left(1-\frac{R'(\underline{z})}{P'(b)}\right)\frac{R(\underline{z})}{R'(\underline{z})}+(2\ell-1)\frac{P(b)}{P'(b)}-\left(\left(\frac{R'(\underline{z})}{P'(b)}-1\right)\frac{R(\underline{z})}{R'(\underline{z})}+(1-\ell)\frac{P(b)}{P'(b)}\right)\frac{\ell h'''(\ell)}{h''(\ell)}\right]\frac{\hat{\ell}}{\ell}$$
$$=\left(\frac{R'(\underline{z})}{P'(b)}-1\right)\left[\frac{\hat{R}(\underline{z})}{R'(\underline{z})}+\hat{\underline{z}}\right]+(1-\ell)\frac{\hat{P}(b)}{P'(b)}+\left(1-\ell+\frac{1}{\ell h''(\ell)}\right)\hat{b}$$
$$+\left(\frac{R(\underline{z})}{R'(\underline{z})}\right)\frac{\hat{R}'(\underline{z})}{R'(\underline{z})}-\left(\frac{R(\underline{z})}{P'(b)}+(1-\ell)\frac{P(b)}{P'(b)}\right)\frac{\hat{P}'(b)}{P'(b)}.$$

Next, recall the incidence of the tax reform on $\hat{\underline{z}}$ and $\hat{b}$ when the utility is logarithmic is given by

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

Substitute these expressions into the previous equation to get

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

where

$$D=-\left(1-2\frac{R'(\underline{z})}{P'(b)}\right)\frac{R(\underline{z})}{R'(\underline{z})}-(3\ell-2)\frac{P(b)}{P'(b)}$$
$$-\left[\left(\frac{R'(\underline{z})}{P'(b)}-1\right)\ell\frac{R(\underline{z})}{R'(\underline{z})}-(1-\ell)\left(\frac{R(\underline{z})}{P'(b)}+(1-\ell)\frac{P(b)}{P'(b)}\right)\right]\ell h''(\ell)$$
$$+\left[\left(\frac{R'(\underline{z})}{P'(b)}-1\right)\frac{R(\underline{z})}{R'(\underline{z})}+(1-\ell)\frac{P(b)}{P'(b)}\right]\frac{\ell h'''(\ell)}{h''(\ell)}.$$

Now, differentiate the firm's profit $\Pi(\theta) = \ell\theta - \underline{z} - \ell b$ to get

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

where we used

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

Differentiating once more gives the following expression for $\frac{\partial^2 \Pi(\theta)}{\partial \ell^2}$:

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

i.e., when the utility is logarithmic,

$$\begin{aligned}
\frac{1}{h''(\ell)} \frac{\partial^2 \Pi(\theta)}{\partial \ell^2} &= \left(1 - 2 \frac{R'(\underline{z})}{P'(b)}\right) \frac{R(\underline{z})}{R'(\underline{z})} + (3\ell - 2) \frac{P(b)}{P'(b)} \\
&\quad + \left[(2\ell - 1) \frac{R(\underline{z})}{P'(b)} - \ell \frac{R(\underline{z})}{R'(\underline{z})} - (1-\ell)^2 \frac{P(b)}{P'(b)} \right] \ell h''(\ell) \\
&\quad - \left[\left(\frac{1}{P'(b)} - \frac{1}{R'(\underline{z})} \right) R(\underline{z}) + (1-\ell) \frac{P(b)}{P'(b)} \right] \frac{\ell h'''(\ell)}{h''(\ell)}.
\end{aligned}$$

As a result, the second-order condition of the firm's problem implies

$$D = -\frac{1}{h''(\ell)} \frac{\partial^2 \Pi(\theta)}{\partial \ell^2} \geq 0.$$

Equation (41) follows. ■

Proof of Proposition 4. Suppose that the tax reform satisfies

$$\frac{\hat{R}(\underline{z})}{R'(\underline{z})} = -\ell \frac{\hat{P}(b)}{P'(b)}.$$

Recall that the utility is log and $R'(\underline{z}) = P'(b) = 1 - \tau$ in the baseline tax system. Lemma 9 gives the impact of this perturbation on the worker's base pay:

$$\hat{\underline{z}} = -\frac{\hat{R}(\underline{z})}{1-\tau} - \frac{R(\underline{z})}{1-\tau} \ell h''(\ell)\hat{\ell},$$

on the bonus:

$$\ell \hat{b} = \frac{\hat{R}(\underline{z})}{1-\tau} + \frac{R(\underline{z}) + (1-\ell)P(b)}{1-\tau} \ell h''(\ell)\hat{\ell},$$

and on expected utility:

$$\hat{U}(\theta) = 0.$$

If the tax rates are perturbed for types $[\theta^*, \infty)$, the impact of the reform on government revenue $\hat{\mathcal{R}}$ is given by

$$\begin{aligned}
&\int_{\theta^*}^{\infty} \left[\hat{T}(\underline{z}(\theta)) + \ell(\theta)\hat{T}_B(b(\theta)) + T_B(b(\theta))\hat{\ell}(\theta) + T'(\underline{z}(\theta))\hat{\underline{z}}(\theta) + T_B'(b(\theta))\ell(\theta)\hat{b}(\theta)\right] dF(\theta) \\
&= \int_{\theta^*}^{\infty} \left[T_B(b(\theta)) + \frac{\tau}{1-\tau} P(b(\theta))\ell(\theta)(1-\ell(\theta))h''(\ell(\theta))\right] \hat{\ell}(\theta) dF(\theta)
\end{aligned}$$

where $T(\underline{z}) \equiv \underline{z} - R(\underline{z})$ and $T_B(b) \equiv b - P(b)$, and where the second equality uses the expressions we have derived above for the incidence of the reform around a baseline tax system where $T'(\underline{z}(\theta)) = T_B'(b(\theta)) = \tau$. Since the terms in square brackets are positive, it follows that $\hat{\mathcal{R}} > 0$ if $\hat{\ell}(\theta) > 0$ for all θ.

Now, the incidence of tax reforms on labor effort is given by equation (41). Apply this formula with $\frac{\hat{R}(\underline{z})}{R'(\underline{z})} = -\ell \frac{\hat{P}(b)}{P'(b)}$ to obtain

$$D\frac{\hat{\ell}}{\ell} = \frac{1}{\ell h''(\ell)} \frac{\hat{P}(b)}{P'(b)} + \left[\frac{R(\underline{z})}{P'(b)} + (1-\ell)\frac{P(b)}{P'(b)}\right] \frac{\hat{P}'(b)}{P'(b)} - \frac{R(\underline{z})}{R'(\underline{z})} \frac{\hat{R}'(\underline{z})}{R'(\underline{z})},$$

where $D = -\frac{1}{h''(\ell)} \frac{\partial^2 \Pi(\theta)}{\partial \ell^2} \geq 0$. It follows that, if the tax reform lowers the marginal and total tax rate on bonuses, so that $\hat{P}(b) > 0$ and $\hat{P}'(b) > 0$, labor effort unambiguously increases ($\hat{\ell} > 0$) if the reform also implies $\hat{R}'(\underline{z}) < 0$; that is, if the marginal tax rate on base pay increases. Assuming that the tax schedule is initially linear on

$[0,\infty)$ ensures that this is satisfied, since in that case $\ell(\theta)$ and $b(\theta)/\underline{z}(\theta)$ are constant (by Proposition 1); thus, for a linear downward perturbation of the bonus tax rate $\hat{P}(b) = b$, we have $\hat{R}(\underline{z}(\theta)) = -\ell b(\theta) \propto -\underline{z}(\theta)$, i.e., the base pay tax rate is perturbed upwards linearly. $\blacksquare$

Proof of Theorem 3. Suppose bonuses are taxed with a schedule $T_b(b)$ that has a tax rate τ_b in the top bracket $[b^*,\infty)$. Denote by θ^* the type that earns the bonus b^*. The base pay of top bonus earners is taxed at a fixed rate t. Denote the average bonus tax rate at b^* by $t_b \equiv T_b(b^*)/b^*$. The elasticities of aggregate variables with respect to $1-\tau_b$ used below are constructed by keeping the set of top agents fixed, as in the proof of Theorem 2).

Consider an increase in the top bonus tax rate by $\delta\hat{\tau}_b$ with $\delta \to 0$, i.e., a perturbation $\hat{T}_b(b) = -\hat{\tau}_b(b-b^*)\mathbb{I}_{\{b\geq b^*\}}$ and $\hat{T}_b'(b) = -\hat{\tau}_b\mathbb{I}_{\{b\geq b^*\}}$. At the optimum, the first-order change in social welfare caused by the reform is equal to zero:

$$\int_{\theta^*}^{\infty}\left[\hat{\tau}_b(b-b^*)\ell(1-\tilde{g}(\bar{z}\mid\theta)) + \tau_z\hat{\underline{z}} + \tau_b\ell\hat{b} + T_b(b)\hat{\ell}\right]\frac{dF(\theta)}{1-F(\theta^*)} = 0,$$

where, using the expression for $\hat{U}$ derived in Lemma 9, we define the modified social marginal welfare weights as

$$\tilde{g}(\bar{z}\mid\theta) \equiv \frac{1}{\lambda}\frac{\frac{1}{(1-\tau_b)u'(R(\underline{z},b))}}{\frac{1-\frac{1-t}{1-\tau_b}\ell(\theta)}{(1-t)u'(R(\underline{z},0))} + \frac{\ell(\theta)}{(1-\tau_b)u'(R(\underline{z},b))}}\alpha(\theta)u'(R(\underline{z},b)).$$

Note that, if $\tau_b \geq t$, then $\tilde{g}(\bar{z}\mid\theta) > g(\bar{z}\mid\theta)$. By continuity, there exists $\underline{t} < t$ such that, as long as $\tau_b \geq \underline{t}$, this inequality continues to hold.

Denote the average effort ℓ over top bonus earners by $L = \int_{\theta^*}^{\infty}\ell\frac{dF(\theta)}{1-F(\theta^*)}$, the average frequency-adjusted bonus ℓb by $B = \int_{\theta^*}^{\infty}\ell b\frac{dF(\theta)}{1-F(\theta^*)}$, and the average base pay by $\underline{\mathcal{Z}} = \int_{\theta^*}^{\infty}\underline{z}\frac{dF(\theta)}{1-F(\theta^*)}$. Additionally, denote the bonus-weighted average of the modified marginal social welfare weights in the top bracket by $\tilde{\mathcal{G}} = \int_{\theta^*}^{\infty}\frac{\ell b-\ell b^*}{B-Lb^*}\tilde{g}(\bar{z}\mid\theta)\frac{dF(\theta)}{1-F(\theta^*)}$.

We can now rewrite the previous equation as follows. The mechanical effect is equal to $\hat{\tau}_b(B-Lb^*)$. The welfare effect is $-\hat{\tau}_b(B-Lb^*)\tilde{\mathcal{G}}$. The behavioural effects due to bonus and effort changes are equal to

$$\int_{\theta^*}^{\infty}\left[\tau_b\ell\hat{b}+T_b(b)\hat{\ell}\right]\frac{dF(\theta)}{1-F(\theta^*)}=\int_{\theta^*}^{\infty}\left[\tau_b\widehat{\ell b}+(T_b(b)-\tau_b b)\hat{\ell}\right]\frac{dF(\theta)}{1-F(\theta^*)}$$
$$=-\hat{\tau}_b\left(\frac{\tau_b}{1-\tau_b}Be_B-\frac{t_b-\tau_b}{1-\tau_b}Lb^*e_L\right),$$

where we used $T_b(b)-\tau_b b=\tau_b(b-b^*)+t_b b^*-\tau_b b=(t_b-\tau_b)b^*$. The behavioural effect due to base pay responses is equal to:

$$\int_{\theta^*}^{\infty}\tau_z\hat{\underline{z}}\frac{dF(\theta)}{1-F(\theta^*)}=-\hat{\tau}_b\frac{\tau_z}{1-\tau_b}\underline{Z}e_{\underline{Z}}.$$

Combining the terms, we get

$$\frac{\tau_b}{1-\tau_b}=\frac{1-\tilde{\mathcal{G}}\frac{\tau_z}{1-\tau_b}\frac{\underline{Z}}{B}\rho_b e_{\underline{Z}}-\frac{t_b-\tau_b}{1-\tau_b}(1-\rho_b)e_L}{\rho_b e_B},$$

where $\rho_b=\frac{B/L}{B/L-b^*}$ is the empirical Pareto coefficient of top bonuses.

Structural expression for $e_{\underline{Z}}$. Using the incidence formulas for the case of separate taxation (Lemma 9), we can write

$$\hat{\underline{z}}=-\frac{\hat{\tau}_b}{1-\tau_b}\varepsilon_{out}\ell(b-b^*)-b\varepsilon_{in}\hat{\ell}=-\frac{\hat{\tau}_b}{1-\tau_b}\left[\varepsilon_{out}\ell(b-b^*)-\ell b\varepsilon_{in}e_\ell\right].$$

Plugging this expression into the definition of $e_{\underline{Z}}=-\frac{1-\tau_b}{\hat{\tau}_b}\frac{1}{\underline{Z}}\int_{\theta^*}^{\infty}\hat{\underline{z}}\frac{dF(\theta)}{1-F(\theta^*)}$, we get

$$e_{\underline{Z}}=\frac{(B-Lb^*)}{\underline{Z}}\underbrace{\int_{\theta^*}^{\infty}\frac{\ell(b-b^*)}{B-Lb^*}\varepsilon_{out}\frac{dF(\theta)}{1-F(\theta^*)}}_{=\overline{\varepsilon_{out}}}-\frac{B}{\underline{Z}}\underbrace{\int_{\theta^*}^{\infty}\frac{\ell b}{B}\varepsilon_{in}e_\ell\frac{dF(\theta)}{1-F(\theta^*)}}_{=\overline{\varepsilon_{in}e_\ell}}$$

This expression easily leads to equation (31). ∎

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 $\theta\ell - \mathbb{E}z$ subject to the incentive constraint

$$\ell = \arg\max_{l\geq 0} \mathbb{E}[U(c,l)] \tag{43}$$

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

We show below that the incentive compatibility constraint (43) 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 D 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

$$h'(\ell) = \frac{(1-\tau)\theta}{1+\gamma h''(\ell)\sigma_\eta^2}. \tag{44}$$

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 (44), 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.

Corollary 2 *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 1, 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 3 *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. We relax this assumption in section C.4. 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:

$$\ell(\theta) \in \arg\max_{\hat{\ell}} u(R(z(\theta, \eta + \hat{\ell} - \ell(\theta)))) - h(\hat{\ell}), \quad \forall \eta. \tag{45}$$

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 1, 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 3. Finally, the free-entry condition (4) holds.

Assumption 4 *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.

Proposition 5 *The earnings schedule is log-linear and given by:*

$$\log z(\theta, \eta) = \log(\theta \ell) + \beta \eta - \frac{1}{2}\beta^2 \sigma_\eta^2 \quad \textit{with} \quad \beta = \frac{h'(\ell)}{1-p}. \tag{46}$$

Effort ℓ is independent of θ and satisfies:

$$\ell = [(1-p)(1-\varepsilon_{\beta,\ell}\beta^2\sigma_\eta^2)]^{\varepsilon_\ell^F/(1+\varepsilon_\ell^F)}, \tag{47}$$

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

$$U(\theta) = \log(R(\theta\ell)) - h(\ell) - \frac{1}{2}(1-p)\beta^2\sigma_\eta^2. \tag{48}$$

Proposition 5 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 (6)), namely $\beta = h'(\ell)/(1-p)$. As in Section 1, this property follows immediately by taking the first-order condition in the incentive compatibility constraint (45). This implies in turn that the crowd-out and crowd-in elasticities are given by $\varepsilon_{\beta,1-p} = -1$ and $\varepsilon_{\beta,\ell} = 1/\varepsilon_\ell^F$. Lemma 2 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.

Theorem 4 *Suppose that the social welfare objective is utilitarian. The optimal rate* *of progressivity satisfies*

$$\frac{p}{(1-p)^2} = \frac{\sigma_\theta^2 + (1+\varepsilon_{\beta,1-p})\beta^2\sigma_\eta^2}{\left[1+\frac{g}{(1-g)p}\right]\varepsilon_{\ell,1-p} + (1-p)\varepsilon_{\beta,\ell}\varepsilon_{\ell,1-p}\beta^2\sigma_\eta^2}. \tag{49}$$

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 (49) up to a second order as $\beta \to 0$. We derive further theoretical and quantitative results in Doligalski et al. (2020).

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 et al. (2012).[47] Workers are indexed by their constant productivity θ. They live for $S \geq 2$ periods and discount the 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

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

[47] Our results of Section 1 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).

subject to the incentive constraint:

$$\begin{aligned} & \mathbb{E}_1\left[\sum_{t=1}^{S} \beta^{t-1}(u(R_t(z_t(\theta,\eta^t))) - h(\tilde{\ell}_t(\eta^t)))\right] \qquad (50) \\ \leq \quad & \mathbb{E}_1\left[\sum_{t=1}^{S} \beta^{t-1}(u(R_t(z_t(\theta,\eta^t))) - h(\ell_t(\theta)))\right], \quad \forall \{\tilde{\ell}_t(\eta^t)\}_{1\leq t\leq S} \end{aligned}$$

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.

Proposition 6 *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*

$$\beta_t = \delta_t \frac{h'(\ell_t)}{1-p}. \qquad (51)$$

The earnings schedule satisfies

$$\log(z_t(\theta,\eta^t)) = \log(z_{t-1}(\theta,\eta^{t-1})) + \beta_t\eta_t - \frac{1}{2}\beta_t^2\sigma_\eta^2, \qquad (52)$$

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 (52) 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.

Theorem 5 *Suppose that the planner is utilitarian. The optimal rate of progressivity is given by*

$$\frac{p}{(1-p)^2} = \frac{\sigma_\theta^2}{\varepsilon_{L,1-p} + (1-p)\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\frac{\delta_1}{\delta_s}\varepsilon_{\beta_s,\ell_s}\varepsilon_{\ell_s,1-p}\beta_s^2\sigma_\eta^2} \tag{53}$$

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 (53) is similar to its static counterpart (49). 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 No Commitment of Workers, Competitive Screening and Adverse Selection

In this section, we consider a modified version of the Edmans and Gabaix (2011) model in which workers cannot commit to stay with their employer once they privately observe their idiosyncratic performance shock η. Effectively, the performance shock becomes a hidden type, and the framework becomes one of adverse selection: workers with heterogenous performance shocks are screened by competitive firms that offer a menu of contracts for workers to select from.

Our results are as follows. First, the equilibrium in this model, taking the income tax schedule as given, always exists and is unique. Second, the equilibrium is very simple: labor effort satisfies the standard first-order condition with the tax rate as the only distortion, workers' earnings are equal to their realized output and there is no private insurance within firms. Below we describe the setup of the model, derive our results, and discuss them in the context of other models of adverse selection.

Setup. The setting is as in Edmans and Gabaix (2011), described in Section C.2. Workers are characterized by a fixed ability θ and a performance shock η. They choose labor effort $\ell \geq 0$ and produce output $y = \theta(\ell + \eta)$. The utility over consumption c and effort ℓ is given by $u(c) - h(\ell)$, where u is strictly increasing and concave, and h is strictly increasing and strictly convex, both continuously differentiable. Competitive firms observe workers' ability θ. Thus, when describing the equilibrium we will focus on a market for a particular realization of ability θ. Furthermore, firms observe the individual's output, but do not observe the performance shock nor labor effort. This means that the performance shock realization is a hidden type. We define a labor contract (y, z) as a pair of output y and earnings z. Firms are competitive: free entry ensures that they cannot make positive profits in equilibrium. To be precise, we use the equilibrium notion of Miyazaki (1977)-Wilson (1977)-Spence (1978), defined formally below. Defining an equilibrium as in M. Rothschild and Stiglitz (1976) would make no difference: the two equilibrium notions always coincide in this setting.

Definition 1 ***(Miyazaki-Wilson-Spence equilibrium)*** *A set of contracts is an equilibrium if (i) firms make zero profits on their overall portfolio of contracts offered, and (ii) there is no other potential contract which would make positive profits, if offered, after all contracts rendered unprofitable by its introduction have been withdrawn.*

We make a number of simplifying assumptions for ease of exposition. These assumptions can be relaxed without affecting the main results.

1. The performance shock η takes two values: η_H with probability $\lambda \in (0,1)$ and η_L with the remaining probability, where $\eta_H > \eta_L \geq 0$.

2. The utility function is quasilinear: $u(c) = c$.

3. The income tax schedule is affine with tax rate t and lump-sum transfer T.

Since we consider a labor market with asymmetric information, the equilibrium set of contracts needs to satisfy the incentive-compatibility (IC) constraints. Denote the contract designed for type η_i, $i \in \{L, H\}$, by (y_i, z_i), with the associated effort level $\ell_i = \frac{y_i}{\theta} - \eta_i$. The IC constraints then ensure that both types of workers have incentives to select the contract that is designed for them:

$$(1-t)z_i - h(\ell_i) \geq (1-t)z_{-i} - h(\ell_{-i} + \eta_{-i} - \eta_i) \quad \forall i \in \{L, H\}. \tag{54}$$

Consider the following candidate for the equilibrium—it is the one we would expect to see in the absence of information frictions. Namely, firms break even on each type of worker and labor effort follows the standard first-order condition. Lemma 10 below shows that in such a candidate equilibrium, the IC constraints are always slack. Proposition 7, which follows, demonstrates further that the candidate is the unique equilibrium and always exists.

Lemma 10 *Consider a candidate equilibrium in which a worker with performance shock η_i, $i \in \{L, H\}$, exerts effort ℓ^* which satisfies $(1-t)\theta = h'(\ell^*)$ and receives pre-tax earnings $z_i = \theta(\ell^* + \eta_i)$. The incentive-compatibility constraints are slack.*

Proof. Rewrite the IC constraint of type L as $h(\ell^* + \eta_H - \eta_L) - h(\ell^*) \geq \theta(1-t)(\eta_H - \eta_L)$. We can bound the left-hand side from below:

$$\begin{aligned} h(\ell^* + \eta_H - \eta_L) - h(\ell^*) &= \int_0^{\eta_H - \eta_L} h'(\ell^* + x)dx \\ &> \int_0^{\eta_H - \eta_L} h'(\ell^*)dx = h'(\ell^*)(\eta_H - \eta_L) = \theta(1-t)(\eta_H - \eta_L), \end{aligned} \tag{55}$$

where the inequality follows from the strict convexity of $h(\cdot)$ and the last equality follows from the first-order condition for effort. Thus, the IC constraint of the L type is always slack. It is straightforward to show the analogous result for type H. ■

Proposition 7 *There exists a unique equilibrium, which is given by the candidate equilibrium from Lemma 10.*

Proof. First we will prove the properties of the equilibrium conditional on existence, then we will prove the existence and uniqueness. The proposition effectively postulates that in equilibrium *(i)* firms break even on each contract and *(ii)* labor effort satisfies the standard first-order condition.

First, suppose that condition *(i)* is not satisfied. Without loss of generality, suppose that the incumbent firm makes positive profits on worker type i by offering a contract (y_i, z_i) with $y_i > z_i$ and is incurring losses on the other type. A profitable deviation for the competitor is to offer a single contract $(\tilde{y}, \tilde{z}) = (y_i, \frac{y_i+z_i}{2})$ to both types. Type i is always better-off accepting the contract, while type $-i$ may or may not be better-off. The competitor makes positive profits on each worker who accepts the contract regardless of their type, since output is higher than wage payments: $\tilde{y} > \tilde{z}$.

Second, suppose that condition *(ii)* is not satisfied for type i. The profitable deviation is then to offer a contract $(\tilde{y}, \tilde{z}) = (\theta(\ell^* + \eta_i)), \theta(\ell^* + \eta_i) - \epsilon)$ for sufficiently small $\epsilon > 0$. It is straightforward to show that for sufficiently small $\epsilon > 0$ such contracts attract type i. Since the contracted output is strictly higher than the wage payment, the competitor makes positive profits on any worker who accepts the new contract. Hence, it does not matter whether type $-i$ is attracted by the contract or not.

The candidate equilibrium of Lemma 10 exists by construction. To prove that it is indeed an equilibrium, we will show that there exists no profitable deviation starting from it. At the candidate equilibrium each worker receives a utility-maximizing contract subject to the employer making no losses on that worker. Consequently, there exists no contract which yields higher utility to a worker and generates positive profits. Thus, there exists no profitable deviation. The uniqueness follows from the strict concavity of h, implying a unique solution ℓ^* to the first-order condition $\theta(1-t) = h'(\ell)$. ■

We find that the equilibrium in the model without workers' commitment and asymmetric information about the performance shock coincides with the allocation that we would expect to see in the absence of information frictions. Thus, the labor market operates exactly as in the standard Mirrlees model with two-dimensional heterogeneity (θ, η). In equilibrium workers are paid exactly what they produce, which can be interpreted as workers having performance-based contracts, e.g., a 100 percent commission or piece rate. However, the exact slope of the contract is not uniquely pinned down. The performance shock is fully passed through to earnings, which means that there is no private insurance within the firm.

It is useful to compare these results to other settings with competitive labor market screening. In particular, Miyazaki (1977) finds rat-race effects (upward effort distortion of high types) and cross-subsidization from high to low types, while Stantcheva (2014), who studies tax policy in Miyazaki's environment, demonstrates that the government can redistribute more than in the standard Mirrlees model. Using the Nash equilibrium notion as in M. Rothschild and Stiglitz (1976) could even lead to a non-existence of equilibrium.

Our model has none of these features. The fundamental difference is in the information available to the firms: in Miyazaki (1977) and Stantcheva (2014), firms observe the worker's effort (hours worked), while in our setting firms observe the worker's output. Screening workers via effort is costly (IC constraints are binding) and complex: it involves additional effort distortions and cross-subsidization between types. By contrast, screening workers via output is easy and costless—IC constraints are slack. In equilibrium with observable output the firm can simply promise each worker a pay equal to the realized output and let workers choose their effort level.

D Proofs of Section C

D.1 Proofs of Section C.1

The incentive constraint reads

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

Taking the first-order condition implies

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

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

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

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

D.2 Proofs of Section C.2

Proof of Proposition 5. 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 (45) implies that $\frac{\partial z(\theta,\eta)}{\partial \eta} = \frac{\partial z(\theta,\eta)}{\partial \ell}$ so that the first-order condition reads

$$
v'(z(\theta,\eta))\frac{\partial z(\theta,\eta)}{\partial \eta} = h'(\ell(\theta)). \tag{56}
$$

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

$$v(z(\theta,\eta)) = h'(\ell(\theta))\eta + k, \tag{57}$$

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

$$k = U(\theta) + h(\ell(\theta)). \tag{58}$$

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:

$$v(z(\theta,\eta)) = h'(\ell(\theta))\eta + [U(\theta) + h(\ell(\theta))]. \tag{59}$$

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 (57) 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:

$$\log(z(\theta,\eta)) = \frac{\ell^{\frac{1}{\varepsilon}}}{1-p}\eta - \frac{1}{1-p}\log\frac{1-\tau}{1-p} + \frac{k}{1-p}, \tag{60}$$

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 (58) implies:

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

and hence

$$\log(z(\theta,\eta)) = \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\eta + \frac{1}{1-p}\frac{1}{1+\frac{1}{\varepsilon_\ell^F}}\ell^{1+\frac{1}{\varepsilon_\ell^F}} - \frac{1}{1-p}\log\frac{1-\tau}{1-p} + \frac{U(\theta)}{1-p}. \tag{61}$$

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

$$\log(z(\theta,\eta)) = \log(\theta\ell) + \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\eta - \frac{1}{2}\left(\frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}\right)^2\sigma_\eta^2. \tag{62}$$

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

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

We have

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

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.

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

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 $\mathbb{E}[z(\theta,\eta)]$ to get

$$e^{\frac{1}{1-p}\frac{1}{1+1/\varepsilon_\ell^F}\ell^{1+1/\varepsilon_\ell^F}+\frac{1}{2}\frac{\ell^{2/\varepsilon_\ell^F}}{(1-p)^2}\sigma_\eta^2-\frac{1}{1-p}\log\frac{1-\tau}{1-p}+\frac{U(\theta)}{1-p}} = \theta\ell. \tag{63}$$

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

$$\ell^{1+1/\varepsilon_\ell^F} + \frac{1}{\varepsilon_\ell^F}\frac{\ell^{2/\varepsilon_\ell^F}}{1-p}\sigma_\eta^2 = 1-p. \tag{64}$$

Using the definition $\beta \equiv \frac{\ell^{\frac{1}{\varepsilon_\ell^F}}}{1-p}$ for the pass-through easily leads to (47). Note that if $\varepsilon_\ell^F = 1$, we obtain optimal effort in closed form:

$$\ell = \left(\frac{1}{1-p} + \frac{\sigma_\eta^2}{(1-p)^2}\right)^{-1/2}. \tag{65}$$

Taking logs in equation (63) easily leads to (48).

Differentiating equation (64) 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 Theorem 4. 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 $\mathbb{E}[z(\theta,\eta)] = \theta\ell$ and $\mathbb{E}[(z(\theta,\eta))^{1-p}] = (\theta\ell)^{1-p}e^{-\frac{p\ell^{2/\varepsilon_\ell^F}\sigma_\eta^2}{2(1-p)}}$, so that government revenue is equal to

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

Budget balance thus requires

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

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

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

with

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

We therefore obtain

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

Using the first-order condition for effort leads to

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

Rearranging this equation leads to the result. ■

D.3 Proofs of Section C.3

Lemma 11 *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,*

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

Proof of Lemma 11. Starting from an incentive compatible allocation, 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 leads to the result. ■

Proof of Proposition 6. We provide a heuristic proof of this proposition; the formal argument follows the same steps as in Edmans et al. (2012). 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

$$\mathbb{E}_{t-1}\left[\frac{\partial U}{\partial z_t}\right] = -\frac{1}{\theta}\frac{\partial U}{\partial \ell_t} \tag{66}$$

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

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

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:

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

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

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

The incentive constraint for effort (66) 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 (52). The expressions for optimal effort and utility are derived in the next proof. ∎

Proof of Theorem 5. Recall that the earnings schedule is given by

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

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

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

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

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

The first-order condition for optimal effort reads

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

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

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

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

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

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. ■