

Taxing Top CEO Incomes: ONLINE APPENDIX

By LAURENCE ALES AND CHRISTOPHER SLEET *

I. Proofs from Section I and Further Analysis of the CEO Assignment Economy

We now give three results that characterize a tax equilibrium and re-express it in a form suitable for optimal tax analysis. The first result confirms that assortative matching occurs in equilibrium.

LEMMA I.1: *If (μ, z, w) is an equilibrium at (T, \bar{U}) , then either (a) $\mu = u$, no firm produces and all candidate CEOs take their outside option or (b) there is a $\bar{v} \in I$ such that (i) for all $v \in (\bar{v}, 1]$, $\mu(v) = u$ and (ii) for all $v \in (0, \bar{v})$, $\mu(v) = v$.*

PROOF OF LEMMA I.1:

If $v' > v$, $\mu(v) = u$ and $\mu(v') \in I$, then $V(S(v), z(v')) - w(v') > V(S(v'), z(v')) - w(v') \geq 0$ and firm v is made strictly better off matching with CEO $\mu(v')$ at income $w(v')$ and effective labor supply $z(v')$. Since CEO $\mu(v')$ is obviously no worse off, this cannot be an equilibrium outcome and, hence, if $\mu(v') \in I$, then $\mu(v) \in I$. It follows that either (i) $\mu = u$ or (ii) there is some $\bar{v} \in I$ such that for $v \in (0, \bar{v})$, $\mu(v) \in I$ and for $v \in (\bar{v}, 1]$, $\mu(v) = u$.

We next argue that μ is increasing on $(0, \bar{v})$. Suppose not, then there exists a pair $v < v' \leq \bar{v}$ such that $\mu(v') < \mu(v)$. We argue that by exchanging partners v and $\mu(v')$ can both improve their payoffs contradicting the fact that (μ, z, w) is an equilibrium. If (μ, z, w) is an equilibrium then, letting $c(v) := w(v) - T[w(v)]$ and $c(v') := w(v') - T[w(v')]$,

$$U\left(c(v), \frac{z(v)}{h(\mu(v))}\right) \geq U\left(c(v'), \frac{z(v')}{h(\mu(v))}\right)$$

and:

$$U\left(c(v'), \frac{z(v')}{h(\mu(v'))}\right) \geq U\left(c(v), \frac{z(v)}{h(\mu(v'))}\right).$$

If the first of these conditions did not hold, then CEO $\mu(v)$ could make herself and firm v' slightly better off by offering to supply $z(v')$ to firm v' for an income very slightly below $w(v')$. Similarly, if the second did not hold, then CEO $\mu(v')$ could make herself and firm v slightly better off by offering to supply $z(v)$ to firm v for an income very slightly below $w(v)$. The Spence-Mirrlees property of U then implies that $z(v') \geq z(v)$ and $c(v') \geq c(v)$ (and, hence, since no firm would pay a higher income to obtain a lower after-tax income and consumption

* Ales: Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA. email: ales@cmu.edu. Sleet: Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA. email: csleet@andrew.cmu.edu.

for its CEO, $w(v') \geq w(v)$. Thus, $\mu(v')$ supplies more effective labor to v' than $\mu(v)$ supplies to v . If $\mu(v')$ instead works for v , supplies the same effective labor $z(v')$ and accepts the same income as before, then the payoff of firm v is changed by: $V(S(v), z(v')) - V(S(v), z(v)) - [w(v') - w(v)]$. If this change is positive, then a contradiction is obtained since firm v is made better off by the partner swap, while $\mu(v')$ is no worse off and so (μ, z, w) cannot be an equilibrium. If this change is non-positive and $z(v') > z(v)$, then:

$$\begin{aligned} 0 &\geq V(S(v), z(v')) - V(S(v), z(v)) - [w(v') - w(v)] \\ &> V(S(v'), z(v')) - V(S(v'), z(v)) - [w(v') - w(v)], \end{aligned}$$

where the second equality uses the strict super-modularity of V and so:

$$V(S(v'), z(v)) - w(v) > V(S(v'), z(v')) - w(v').$$

Thus, firm v' is made strictly better off by swapping partners with firm v , which again contradicts the requirement that (μ, z, w) is an equilibrium. Finally, consider the case in which $V(S(v), z(v')) - V(S(v), z(v)) - [w(v') - w(v)] = 0$, $z(v') = z(v)$ and $w(v) = w(v')$. If firms v and v' swap partners and continue to pay the same incomes and require the same effective labors from their CEOs, then no firm or CEO is made worse off. Denote the common effective labor amount by \hat{z} and the common income by \hat{w} and, to simplify the exposition suppose that the tax function T is differentiable at \hat{w} with derivative $T_w[\hat{w}]$. It cannot be that: $V_z(S(v), \hat{z})h(\mu(v')) = -\frac{U_e(\hat{w}-T[\hat{w}], \hat{z}/h(\mu(v')))}{U_c(\hat{w}-T[\hat{w}], \hat{z}/h(\mu(v')))}(1 - T_w[\hat{w}])$ and $V_z(S(v'), \hat{z})h(\mu(v)) = -\frac{U_e(\hat{w}-T[\hat{w}], \hat{z}/h(\mu(v)))}{U_c(\hat{w}-T[\hat{w}], \hat{z}/h(\mu(v)))}(1 - T_w[\hat{w}])$ since if so the following contradiction emerges:

$$\begin{aligned} V_z(S(v), \hat{z})h(\mu(v')) &= -\frac{U_e(\hat{w} - T[\hat{w}], \hat{z}/h(\mu(v')))}{U_c(\hat{w} - T[\hat{w}], \hat{z}/h(\mu(v')))}(1 - T_w[\hat{w}]) \\ &< -\frac{U_e(\hat{w} - T[\hat{w}], \hat{z}/h(\mu(v)))}{U_c(\hat{w} - T[\hat{w}], \hat{z}/h(\mu(v)))}(1 - T_w[\hat{w}]) \\ &= V_z(S(v'), \hat{z})h(\mu(v)) < V_z(S(v), \hat{z})h(\mu(v')), \end{aligned}$$

where the first inequality follows from the fact that $h(\mu(v')) > h(\mu(v))$ and the Spence-Mirrlees property of U and the second inequality follows from $S(v) > S(v')$, the strict super-modularity of V and $h(\mu(v')) > h(\mu(v))$. Thus, after re-matching at least one pair $(v, \mu(v'))$ or $(v', \mu(v))$ is not at a Pareto optimum. It is then possible for this pair to adjust CEO effort and income to make both firm and CEO better off. Again, this contradicts the assumption that (μ, z, w) is an equilibrium. We conclude that μ is increasing on $(0, \bar{v}]$.

Finally, we show that for $\bar{v} > 0$, μ is the identity map on $(0, \bar{v})$. Since μ is measure-preserving and increasing, it is sufficient to show that there are no

discontinuities in μ and that $\lim_{v \downarrow 0} \mu(v) = 0$. Suppose that μ has a discontinuity at some \hat{v} , but (without loss of generality) is continuous from the right. Then $\mu(\hat{v}) > \mu(\hat{v}_-) := \lim_{v \uparrow \hat{v}} \mu(v)$ and CEOs between $(\mu(\hat{v}_-), \mu(\hat{v}))$ are unmatched. But for $\tilde{m} \in (\mu(\hat{v}_-), \mu(\hat{v}))$, $U(w(\hat{v}) - T[w(\hat{v})], z(\hat{v})/h(\tilde{m})) > U(w(\hat{v}) - T[w(\hat{v})], z(\hat{v})/h(\mu(\hat{v}))) \geq \bar{U}$, contradicting the definition of equilibrium. Thus, μ is continuous. By a very similar argument if $\lim_{v \downarrow 0} \mu(v) > 0$, then there are unmatched CEOs in $(0, \lim_{v \downarrow 0} \mu(v)]$. These CEOs would be made strictly better off by matching with a firm and accepting the terms the firm is giving to her current CEO. Again this is inconsistent with an equilibrium.

Finally, we characterize μ at \tilde{v} . Suppose $\tilde{v} \in (0, 1)$ and let $\underline{v}_n \uparrow \tilde{v}$ and $\bar{v}_n \downarrow \tilde{v}$ (with each $0 < \underline{v}_n < \tilde{v} < \bar{v}_n < 1$). We have:

$$\begin{aligned} W_n &:= U(w(\underline{v}_n) - T[w(\underline{v}_n)], z(\underline{v}_n)/h(\underline{v}_n)) \geq \bar{U} \\ &\geq U(w(\underline{v}_n) - T[w(\underline{v}_n)], z(\underline{v}_n)/h(\bar{v}_n)). \end{aligned}$$

As observed previously $w_n = w(\underline{v}_n)$, $c_n = w(\underline{v}_n) - T[w(\underline{v}_n)]$ and $z_n = z(\underline{v}_n)$ are bounded, decreasing sequences. Denote the limits of these sequences by $(w_\infty, c_\infty, z_\infty)$. Since $\lim h(\underline{v}_n) - h(\bar{v}_n) \downarrow 0$ and U is continuous, it follows that $U(c_n, z_n/h(\underline{v}_n)) - U(c_n, z_n/h(\bar{v}_n))$ converges to 0. Hence, $W_n \downarrow \bar{U}$. In addition, by a similar argument, $V(S(\underline{v}_n), z(\underline{v}_n)) - w(\underline{v}_n) \downarrow 0$. It follows that if T is continuous, then the \tilde{v} firm and CEO are indifferent about matching at the effective labor-income (z_∞, w_∞) .

Without loss of generality we select equilibria in which if $\tilde{v} > 0$, then $\mu(\tilde{v}) = \tilde{v}$. The next proposition simplifies the equilibrium conditions in Definition 1 in the paper in a way that is convenient for tax analysis. It shows that given (T, \bar{U}) , if a pair of effective labor and income functions (z, w) on a domain $(0, \tilde{v}]$ are such that no CEO $v \in I \cap [0, \tilde{v}]$ is made strictly better off exchanging places with CEO $v' \in (0, \tilde{v}]$ and accepting the terms v' receives and similarly for firms, then no firm-CEO pair $(v, v') \in (0, \tilde{v}]^2$ can benefit (both weakly and at least one side strictly) from re-matching and selecting an arbitrary effective labor and income in the codomain of w . Furthermore, if the \tilde{v} -ranked CEO and firm receive the outside options \bar{U} and 0 respectively and if T is such that $T(w) = w$ at all w outside of the co-domain of w , then (\tilde{v}, w, z) is an equilibrium at (T, \bar{U}) . Thus, the stability conditions on firms and CEOs in Definition 1 are decoupled and re-expressed as separate firm and CEO incentive conditions. In addition, Proposition I.1 supplies a converse result: associated with any (non-trivial) equilibrium is a (\tilde{v}, z, w) satisfying the conditions described above.

PROPOSITION I.1: *Let $T : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a tax function, \tilde{v} be a number in I and $z : (0, \tilde{v}] \rightarrow \mathbb{R}_+$ and $w : (0, \tilde{v}] \rightarrow \mathbb{R}_+$ be a pair of effective labor and income functions satisfying the **participation conditions**, for all $v \in (0, \tilde{v}]$,*

$$(I.1) \quad U\left(w(v) - T[w(v)], \frac{z(v)}{h(v)}\right) \geq \bar{U} \quad \text{and} \quad V(S(v), z(v)) - w(v) \geq 0,$$

and the *incentive conditions*, for all $v, v' \in (0, \bar{v}]$,

$$(I.2) \quad U\left(w(v) - T[w(v)], \frac{z(v)}{h(v)}\right) \geq U\left(w(v') - T[w(v')], \frac{z(v')}{h(v')}\right),$$

and

$$(I.3) \quad V(S(v), z(v)) - w(v) \geq V(S(v), z(v')) - w(v').$$

Then there is no tuple (v, v', z', w') with $(v, v') \in (0, \bar{v}]$ and $w' \in w((0, \bar{v}])$ such that:

$$U\left(w' - T[w'], \frac{z'}{h(v')}\right) \geq U\left(w(v') - T[w(v')], \frac{z(v')}{h(v')}\right),$$

and

$$V(S(v), z') - w' \geq V(S(v), z(v)) - w(v),$$

with at least one of these inequalities strict. In addition, if (i) $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) = \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) \geq 0$ and (ii) for all $w' \notin w((0, \bar{v}])$, $T(w') = w'$, then (w, z) defines an equilibrium at (T, \bar{U}) . Conversely, if (\bar{v}, w, z) is an equilibrium at (T, \bar{U}) , then (\bar{v}, w, z) satisfies (I.1) to (I.3), $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) \geq \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) \geq 0$ with equality if $\bar{v} \in (0, 1)$.

PROOF OF PROPOSITION I.1:

Suppose the first claim in the proposition is false and that there is a tuple (v, v', z', w') with $v, v' \in (0, \bar{v}]$ and $w' = w(\hat{v}) \in w((0, \bar{v}])$ such that:

$$U\left(w' - T[w'], \frac{z'}{h(v')}\right) \geq U\left(w(v') - T[w(v')], \frac{z(v')}{h(v')}\right),$$

and

$$V(S(v), z') - w' \geq V(S(v), z(v)) - w(v),$$

with at least one of the previous inequalities strict. If $z' \geq z(\hat{v})$, then:

$$\begin{aligned} U\left(w(\hat{v}) - T[w(\hat{v})], \frac{z(\hat{v})}{h(v')}\right) &\geq U\left(w' - T[w'], \frac{z'}{h(v')}\right) \\ &\geq U\left(w(v') - T[w(v')], \frac{z(v')}{h(v')}\right) \geq U\left(w(\hat{v}) - T[w(\hat{v})], \frac{z(\hat{v})}{h(v')}\right), \end{aligned}$$

where the first inequality follows from $z' \geq z(\hat{v})$, the strict monotonicity of U in e and $w' = w(\hat{v})$, the second inequality is by assumption and the third follows from (I.2). If $z' > z(\hat{v})$, the first of the preceding inequalities is strict implying a contradiction. Thus, $z' \leq z(\hat{v})$ and if $z' = z(\hat{v})$, the v' -ranked CEO is no better

off working for the v -ranked firm at (z', w') . If $z' < z(\hat{v})$, then

$$V(S(v), z(v)) - w(v) \geq V(S(v), z(\hat{v})) - w(\hat{v}) > V(S(v), z') - w(\hat{v}),$$

where the first inequality is by (I.3) and the second is from the strict monotonicity of V in z , and the v -ranked firm is worse off matching with the v' -ranked CEO at (z', w') . For this firm to be strictly better off, $z' > z(\hat{v})$. We conclude that v -ranked firm cannot be made strictly better off without making the v' -ranked CEO strictly worse off and vice versa. A contradiction is attained.

It follows that if (z, w) satisfies the conditions in the proposition over the domain $(0, \bar{v}]$, then each firm (resp. CEO) $v \in (0, \bar{v}]$ is better off matched with CEO (resp. firm) v , than matched with an alternative partner $v' \in (0, \bar{v}]$ at an income $w' \in w((0, \bar{v}])$ and an effective labor supply that improves their alternative partner's payoff relative to $(z(v'), w(v'))$. Moreover, if, $T(w') = w'$ for all $w' \in \mathbb{R}_+ \setminus w((0, \bar{v}])$, then there is 100 per cent taxation of any price outside of the range of w on $(0, \bar{v}]$. Clearly, no firm or CEO would wish to choose such an income and, hence, no firm or CEO in $(0, \bar{v}]$ can benefit from rematching with another partner v' in this set and choosing *any* effective labor supply and income that improves their alternative partner's payoff relative to $(z(v'), w(v'))$. In addition, if $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) = \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) = 0$, then no firm or CEO $v \in (0, \bar{v}]$ is better off unmatched than matched at $(w(v), z(v))$, i.e.

$$\begin{aligned} U\left(w(v) - T[w(v)], \frac{z(v)}{h(v)}\right) &\geq U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) \\ &\geq U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) \geq \bar{U}, \end{aligned}$$

and similarly for firms. Finally, if $\bar{v} \in (0, 1]$ and $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) = \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) = 0$, then by similar logic to that given above, it is readily verified that all firms and CEOs $v \in (\bar{v}, 1]$ are better off unmatched than matched with a partner in I at a (z', w') that gives the partner as much as it could obtain from remaining its current match or remaining unmatched.

For the converse, if (\bar{v}, z, w) is an equilibrium at (T, \bar{U}) , then it is immediate that it satisfies (I.1) to (I.3) and $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) \geq \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) \geq 0$. If $\bar{v} \in (0, 1)$, then $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) = \bar{U}$ and $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) = 0$ since, otherwise, there is an interval (\bar{v}, v') such that for each $v \in (\bar{v}, v')$ either $U\left(w(\bar{v}) - T[w(\bar{v})], \frac{z(\bar{v})}{h(\bar{v})}\right) > \bar{U}$ or $V(S(v), z(\bar{v})) - w(\bar{v}) > 0$. This contradicts the equilibrium definition.

PROOF OF PROPOSITION 1

The proof of Proposition 1 in the main paper is a now a direct consequence of Lemma I.1 and Proposition I.1. □

Lemma I.2 provides monotonicity results for various equilibrium functions. It also proves the existence of a function ω relating equilibrium income to effective labor supply.

LEMMA I.2: *If (\bar{v}, z, w) is an equilibrium threshold and a pair of equilibrium effective labor and income functions at (T, \bar{U}) , then (z, w, c) , with $c(v) = w(v) - T[w(v)]$, $v \in (0, \bar{v}]$ are non-increasing and there exists a function $\omega : z((0, \bar{v})) \rightarrow \mathbb{R}_+$ satisfying for each $v \in (0, \bar{v})$, $\omega(z(v)) = w(v)$. In addition, equilibrium CEO utility Φ , $\Phi(v) = U(c(v), z(v)/h(v))$ for $v \in (0, \bar{v})$, and firm profits π , $\pi(v) = V(S(v), z(v)) - w(v)$ for $v \in (0, \bar{v})$, are decreasing.*

PROOF OF LEMMA I.2:

Monotonicity of z and c follow from (I.2), the Spence-Mirrlees property of U and standard arguments. Hence, if $v > v'$, then $c(v) \leq c(v')$ and so if $w(v) > w(v')$, then $T[w(v)] - T[w(v')] > w(v) - w(v')$. But clearly no firm would choose to buy from a CEO at income $w(v)$ (they could strictly reduce the income they pay and weakly raise their CEO's consumption). Hence, w must be non-decreasing as well. Moreover, w is $\sigma(z)$ -measurable (where $\sigma(z)$ denotes the sigma-algebra induced by z). Hence, there exists a function ω , with $\omega(z(v)) = w(v)$ (see, for example, Klenke (2008), Corollary 1.97, p. 41). Finally, if $v < v'$, then $\Phi(v) = U(c(v), z(v)/h(v)) \geq U(c(v'), z(v)/h(v)) > U(c(v'), z(v')/h(v')) = \Phi(v')$ and similarly for π .

II. Derivation of Elasticities for Section II

In this appendix, we derive expressions for elasticities used in Section II of the paper under the assumption of a multiplicative firm surplus function DSz . We first assume general preferences and then specialize to the quasilinear/constant effort elasticity case. The elasticities given in the main text are "aggregate elasticities" that summarize the CEO income and firm profit responses of populations of CEOs and firms to a tax rate change. Below we build these elasticities up from individual level responses and equilibrium conditions.

INDIVIDUAL INCOME ELASTICITIES FOR CEOS

As in the main text let $\omega(z, 1 - \tau)$ give CEO income as a function of effective labor supply and the tax rate. In this appendix, we also make the dependence of a CEO's effective labor on the retention rate explicit in the notation and let $z(v, 1 - \tau)$ denote the equilibrium effective labor supply of CEO v given retention rate $1 - \tau$. Suppressing dependence of functions on their arguments in the

notation, the CEOs first order conditions are:

$$(II.1) \quad (1 - \tau) \frac{\partial \omega}{\partial z} h U_c \left(c, \frac{z}{h} \right) + U_e \left(c, \frac{z}{h} \right) = 0.$$

where: $c(v, 1 - \tau) = \omega(z(v, 1 - \tau), 1 - \tau) - \tau(\omega(z(v, 1 - \tau), 1 - \tau) - w_0) - T[w_0]$.
The corresponding firm's first order condition is:

$$(II.2) \quad \frac{\partial \omega}{\partial z} = DS.$$

Since $z(v, 1 - \tau) = h(v)e(v, 1 - \tau)$, we have $z_v = z \left\{ \frac{h_v}{h} + \frac{e_v}{e} \right\}$. Substituting (II.2) into (II.1), totally differentiating and using the definitions of the uncompensated and compensated CEO effort elasticities \mathcal{E}^u and \mathcal{E}^c gives:

$$\frac{e_v}{e} = \mathcal{E}^u \frac{h_v}{h} + \mathcal{E}^c \frac{e_v}{e}.$$

And so:

$$(II.3) \quad \frac{z_v}{z} = \left\{ (1 + \mathcal{E}^u) \frac{h_v}{h} + \mathcal{E}^c \frac{S_v}{S} \right\}.$$

Differentiating (II.2) with respect to v and combining with (II.3) gives:

$$\frac{\partial^2 \omega}{\partial z^2} = \frac{DS_v}{z \left\{ (1 + \mathcal{E}^u) \frac{h_v}{h} + \mathcal{E}^c \frac{S_v}{S} \right\}}.$$

Totally differentiating the CEO's first order condition (II.1) with respect to $1 - \tau$ gives:

$$\frac{\partial z}{\partial(1 - \tau)} = \frac{\frac{\partial \omega}{\partial z} h U_c + \left\{ w - w_0 + (1 - \tau) \frac{\partial \omega}{\partial(1 - \tau)} \right\} \left\{ (1 - \tau) \frac{\partial \omega}{\partial z} h U_{cc} + U_{ec} \right\} + (1 - \tau) \frac{\partial^2 \omega}{\partial z \partial(1 - \tau)} h U_c}{(1 - \tau) \frac{\partial^2 \omega}{\partial z^2} h U_c + \left\{ (1 - \tau) \frac{\partial \omega}{\partial z} \right\}^2 h U_{cc} + 2(1 - \tau) \frac{\partial \omega}{\partial z} U_{ce} + U_{ee} / h}.$$

Totally differentiating the firm's first order condition (II.2) gives:

$$\frac{\partial^2 \omega}{\partial z^2} \frac{\partial z}{\partial(1 - \tau)} + \frac{\partial^2 \omega}{\partial z \partial(1 - \tau)} = 0.$$

Combining the preceding conditions:

$$(II.4) \quad \frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} = - \left(\frac{1-\tau}{z} \right) \frac{DShU_c + \left[\frac{w-w_0}{w} + \frac{1-\tau}{\omega} \frac{\partial \omega}{\partial(1-\tau)} \right] w \{ (1-\tau) DShU_{cc} + U_{ec} \}}{\{ (1-\tau) DS \}^2 h U_{cc} + 2(1-\tau) DS U_{ce} + U_{ee}/h}.$$

Note if CEO utility is quasilinear in consumption and there are no income effects, then $U_{cc} = U_{ce} = 0$ and the elasticity in (II.4) reduces to the usual behavioral elasticity $\mathcal{E} = \frac{U_e}{U_{ee}}$. If income effects on effort supply are negative and $\frac{1-\tau}{\omega} \frac{\partial \omega}{\partial(1-\tau)} \leq 0$, then

$$\begin{aligned} \frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} &\geq \left(\frac{U_e}{U_{ce}} \right) \frac{U_c + \frac{w}{DSh} \{ -\frac{U_e}{U_c} U_{cc} + U_{ec} \} e}{\{ \frac{U_e}{U_c} \}^2 U_{cc} - 2 \frac{U_e}{U_c} U_{ce} + U_{ee}} \\ &> \left(\frac{U_e}{U_{ce}} \right) \frac{U_c + \{ -\frac{U_e}{U_c} U_{cc} + U_{ec} \} e}{\{ \frac{U_e}{U_c} \}^2 U_{cc} - 2 \frac{U_e}{U_c} U_{ce} + U_{ee}}, \end{aligned}$$

where the final right hand side is the usual behavioral uncompensated effort elasticity. In particular, in this case, if the latter elasticity is non-negative, then so too is $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$. In addition, at v_0 , $w = w_0$ and if $\frac{\partial \omega(z(v_0), 1-\tau)}{\partial(1-\tau)} = 0$ (as will be shown below), then $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$ equals the usual compensated effort elasticity.

We now verify that at v_0 neither the firm's nor the CEO's payoff changes in response to a small retention rate change. Let Φ_0 denote the utility of the v_0 -ranked CEO prior to the tax change. Let z_0 denote this CEO's effective labor, w_0 her income, h_0 her talent, π_0 firm v_0 's profit and S_0 its asset size (all prior to the tax change). Then:

$$(II.5) \quad U \left(w_0 - T[w_0], \frac{z_0}{h_0} \right) = \Phi_0.$$

A small change to $1-\tau$ perturbs CEO v_0 's utility by: $U_c(1-\tau) \frac{\partial \omega}{\partial(1-\tau)}(z_0, 1-\tau)$. If this CEO continues to receive Φ_0 after the change, then $\frac{\partial \omega}{\partial(1-\tau)}(z_0, 1-\tau) = 0$. Furthermore,

$$\begin{aligned} \frac{\partial \pi}{\partial(1-\tau)}(v_0, 1-\tau) &= \left\{ DS_0 - \frac{\partial \omega}{\partial z}(z_0, 1-\tau) \right\} \frac{\partial z}{\partial(1-\tau)}(v_0, 1-\tau) \\ &\quad - \frac{\partial \omega}{\partial(1-\tau)}(z_0, 1-\tau) = 0. \end{aligned}$$

Consider three cases. First $v_0 = \bar{v} = 1$ and all firms are matched prior to the tax

change with $\Phi_0 = \bar{U}$. The v_0 -firm cannot reduce the utility of its CEO (which must remain above \bar{U}), has no desire to raise the utility of its CEO and by our equilibrium assumption no need to. Thus, the CEO's payoff remains at \bar{U} and the firm's profit remains at π_0 . Second, $v_0 = \bar{v} < 1$. In this case not all firms are matched prior to the tax change, but $\Phi_0 = \bar{U}$. Again, the v_0 -firm cannot reduce the utility of its CEO and has no desire to raise this utility. If it continues to offer a utility of \bar{U} to its CEO, then its payoff is unchanged (and equal to 0). Firms $v \in (v_0, 1)$ (continue to) make strictly smaller and, hence, negative profits if they enter and match with the v_0 ranked CEO or with their correspondingly ranked (candidate) CEO. Hence, these firms do not enter the assignment market and firm v_0 does not need to offer CEO v_0 a utility above \bar{U} . Similar logic ensures that in the third case $v_0 < \bar{v} \leq 1$, firm v_0 continues to offer the CEO the same utility as was offered prior to the small retention rate variation and continues to receive the same payoff.

Since $\frac{\partial \omega}{\partial(1-\tau)}(z_0, 1-\tau) = 0$ there is no adjustment to the equilibrium CEO income schedule at w_0 . Hence, at $v = v_0$ using (II.2) and (II.4), the elasticity of the CEO's income with respect to taxes is:

$$\begin{aligned} \mathcal{E}_w(v_0) &:= \frac{1-\tau}{w} \frac{dw}{d(1-\tau)} \Big|_{v_0} = \frac{1-\tau}{w_0} \frac{\partial \omega}{\partial z} \frac{\partial z}{d(1-\tau)} + \frac{1-\tau}{w_0} \frac{\partial \omega}{\partial(1-\tau)} \Big|_{v_0} \\ &= \left(\frac{z}{w_0} \frac{\partial \omega}{\partial z} \right) \left(\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} \right) \Big|_{v_0} = \frac{DS_0 z_0}{w_0} \mathcal{E}_0^c, \end{aligned}$$

where \mathcal{E}_0^c is the compensated effort elasticity of CEO v_0 in equilibrium. Define $v(z', 1-\tau)$, $z' \in z([0, 1], 1-\tau)$, to be the rank of the CEO exerting effective labor z' when the retention rate is $1-\tau$, i.e.

$$(II.6) \quad v(z(v, 1-\tau), 1-\tau) = v.$$

At all points of differentiability, (II.6) implies:

$$(II.7) \quad \frac{\partial v}{\partial z}(z, 1-\tau) \frac{\partial z}{\partial(1-\tau)}(v(z, 1-\tau), 1-\tau) + \frac{\partial v}{\partial(1-\tau)}(z, 1-\tau) = 0.$$

It follows from (II.7) that:

$$(II.8) \quad \frac{\partial v}{\partial(1-\tau)}(z, 1-\tau) > 0$$

if the corresponding effective labor elasticity $\frac{1-\tau}{z} \frac{dz}{d(1-\tau)}$ (at $(v(z, 1-\tau), 1-\tau)$) is positive (and if there is no bunching at $v(z, 1-\tau)$ so that $z_v(v(z, 1-\tau)) > 0$).

Using (II.2), we have:

$$(II.9) \quad \omega(z, 1 - \tau) = \omega(z_0, 1 - \tau) + \int_{z_0}^z DS(v(z', 1 - \tau)) dz'.$$

Differentiating (II.9) with respect to $1 - \tau$ (and using $\frac{\partial \omega}{\partial(1-\tau)}(v_0) = 0$) gives:

$$(II.10) \quad \frac{\partial \omega}{\partial(1-\tau)}(z, 1 - \tau) = \int_{z_0}^z \left\{ DS_v(v(z', 1 - \tau)) \frac{\partial v}{\partial(1-\tau)}(z', 1 - \tau) \right\} dz'.$$

This equation combined with (II.7) and the elasticity formula (II.4) gives an explicit (but complicated) formula for $\frac{\partial \omega}{\partial(1-\tau)}$. It follows from (II.10) and our earlier discussion of $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$ that if the uncompensated behavioral elasticity for effort is positive over an interval of effective labors and there is no bunching, then $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$ and $\frac{\partial v}{\partial(1-\tau)}$ are both positive over the interval and $\frac{\partial \omega}{\partial(1-\tau)}(z, 1 - \tau)$ is negative. The logic is straightforward. A rise in $1 - \tau$ induces all CEO's to work harder. Consequently, increments in effective labor between z_0 and z are now associated with lower ranked (i.e. higher v) firms and CEOs. Lower ranked firms pay less for the last unit of effective labor they hire. (Specifically, firm v 's optimality condition may be expressed as: $DS(v) = \frac{\partial \omega}{\partial z}(z(v, 1 - \tau), 1 - \tau)$. Thus, firm v pays $DS(v)$ (only) for the last unit of effective labor it buys. It must do so to secure this marginal unit in the face of competition from slightly less productive firms.) Hence, the additional income paid for each given incremental unit of effective labor hired (by now lower ranked firms) is reduced after a rise in $1 - \tau$ and the income paid to a CEO supplying a given z is reduced.

This does not mean that the income earned by CEO v falls after a rise in $1 - \tau$ since that CEO will supply more effective labor in equilibrium. Specifically, the overall income elasticity of CEO v with respect to the retention rate $1 - \tau$ is:

$$(II.11) \quad \begin{aligned} \mathcal{E}_w(v) &= \left(\frac{z}{w} \frac{\partial \omega}{\partial z} \right) \left(\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} \right) + \left(\frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)} \right) \\ &= \left(\frac{DS(v)z(v)}{w} \right) \left(\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} \right) + \left(\frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)} \right), \end{aligned}$$

where $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$ is defined as in (II.4) and the first term is positive if $\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)}$ is and, using (II.10),

$$(II.12) \quad \frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)}(z, 1 - \tau) = \frac{1-\tau}{w} \int_{z_0}^z \left\{ DS_v(v(z', 1 - \tau)) \frac{\partial v}{\partial(1-\tau)}(z', 1 - \tau) \right\} dz'.$$

The preceding formulas are simplified in the absence of income effects. Then

$\frac{1-\tau}{z} \frac{\partial z}{\partial(1-\tau)} = \mathcal{E}^c = \mathcal{E}^u = \mathcal{E} := \frac{U_e}{U_{ee}}$ and, after a change of variables and an integration by parts,

$$(II.13) \quad -\frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)}(z(v)) = \frac{DSz\mathcal{E}(v) - DS_0z_0\mathcal{E}(v_0)}{w} + \frac{1}{w} \int_v^{v_0} DS(v')z(v')\mathcal{E}(v')\mathcal{K}(v')dv',$$

with:

$$\begin{aligned} \mathcal{K}(v') := & \left(1 + \mathcal{E}(v') + \frac{1-\tau}{\mathcal{E}} \frac{\partial \mathcal{E}}{\partial(1-\tau)}(v')\right) \frac{h_v}{h}(v') \\ & + \left(\mathcal{E}(v') + \frac{1-\tau}{\mathcal{E}} \frac{\partial \mathcal{E}}{\partial(1-\tau)}(v')\right) \frac{S_v}{S}(v') \end{aligned}$$

In the constant elasticity case, (II.13) is further reduced to:

$$\frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)} = -\left(\frac{\pi - \pi_0}{w}\right) \mathcal{E} < 0.$$

Absent income effects, the total elasticity of CEO income with respect to the retention rate $1 - \tau$ is:

$$\mathcal{E}_w(v) = \frac{DSz}{w} \mathcal{E}(v) + \frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)}(v),$$

with $\frac{1-\tau}{w} \frac{\partial \omega}{\partial(1-\tau)}$ defined as in (II.13). In the constant effort elasticity case, this is further reduced to:

$$(II.14) \quad \mathcal{E}_w(v) = \left\{ \frac{DS_0z_0 - w_0}{w(v)} + 1 \right\} \mathcal{E} \geq \mathcal{E} \geq 0.$$

Note that in this case as $v \downarrow 0$, $\mathcal{E}_w(v)$ converges to \mathcal{E} .

AGGREGATE INCOME ELASTICITIES FOR CEOs

Defining aggregate CEO income above w_0 to be $W := \int_0^{v_0} w(v)dv$, the elasticity of W with respect to $1 - \tau$ is thus:

$$\mathcal{E}_W = \frac{1}{W} \int_0^{v_0} w(v) \mathcal{E}_w(v) dv,$$

with $\mathcal{E}_w(v)$ defined as in (II.11). In the quasi-linear/constant effort elasticity case, from (II.14), this is simply:

$$(II.15) \quad \mathcal{E}_W = \mathcal{E} \int_0^{v_0} \frac{w(v)}{W} \left\{ \frac{DS_0 z_0 - w_0}{w(v)} + 1 \right\} dv = \left(\frac{R}{W} - \frac{\Delta\Pi}{W} \right) \mathcal{E} \geq 0,$$

where $R := \int_0^{v_0} DS(v)z(v, 1 - \tau)dv$ and $\Delta\Pi := \int_0^{v_0} \{\pi(v) - \pi_0\}dv$.

INDIVIDUAL AND AGGREGATE FIRM PROFIT ELASTICITIES

Next consider the elasticity of firm profit $\pi = Sz - w$ with respect to $1 - \tau$. Using the firm's first order condition (II.2), this is:

$$(II.16) \quad \frac{1 - \tau}{\pi} \frac{d\pi}{d(1 - \tau)} = - \frac{w}{\pi} \frac{1 - \tau}{w} \frac{\partial \omega}{\partial(1 - \tau)},'$$

with $\frac{1 - \tau}{w} \frac{\partial \omega}{\partial(1 - \tau)}$ defined as in (II.10) or (II.13). Note that since $\frac{1 - \tau}{w} \frac{\partial \omega}{\partial(1 - \tau)} < 0$ if the uncompensated behavioral elasticity is positive, it flows that $\frac{1 - \tau}{\pi} \frac{d\pi}{d(1 - \tau)} > 0$. In the quasilinear/constant elasticity case, (II.16) reduces to:

$$\frac{1 - \tau}{\pi} \frac{d\pi}{d(1 - \tau)} = \left(\frac{\pi - \pi_0}{\pi} \right) \mathcal{E} \geq 0,$$

where the inequality uses the monotonicity of π in v . Defining total firm profit above $\pi_0 = \pi(v_0)$ to be $\Pi := \int_0^{v_0} \pi(v)dv$, the elasticity of Π with respect to $1 - \tau$ is:

$$\mathcal{E}_\Pi = \int_0^{v_0} \frac{\pi}{\Pi} \frac{1 - \tau}{\pi} \frac{d\pi}{d(1 - \tau)} dv.$$

Again this is positive if the underlying uncompensated behavioral effort elasticity. In the quasilinear/constant elasticity case, this formula reduces to:

$$\mathcal{E}_\Pi = \frac{\Delta\Pi}{\Pi} \mathcal{E} \geq 0.$$

OPTIMAL TAX FORMULAS

The preceding "aggregate" elasticities can be substituted into the tax formula (30) given in the main text to give optimal taxes in terms of individual elasticities, the equilibrium CEO income schedule and the various distributions. In the remainder of this appendix we extend (30) to allow for societal concern for top CEOs. We then consider the case $\chi = 1$ in detail and derive (34) in the main text.

SOCIETAL CONCERN FOR TOP CEOs

Let ψ denote the welfare weight on top CEOs. Then the policymaker's first order condition is modified as:

$$\begin{aligned}
& - \int_0^{v_0} \{\omega(z^*(v); 1 - \tau^*) - w_0\} dv \\
& + \int_0^{v_0} \psi(v) U_c(v) \left\{ \omega(z^*(v); 1 - \tau^*) - w_0 + (1 - \tau^*) \frac{\partial \omega}{\partial (1 - \tau)}(z^*(v); 1 - \tau^*) \right\} dv \\
& + \tau^* \int_0^{v_0} \left\{ \frac{\partial \omega}{\partial z}(z^*(v); 1 - \tau^*) \frac{\partial z}{\partial (1 - \tau)}(v, 1 - \tau^*) + \frac{\partial \omega}{\partial (1 - \tau)}(z^*(v); 1 - \tau^*) \right\} dv \\
\text{(II.17)} \quad & - \chi \int_0^{v_0} \frac{\partial \omega}{\partial (1 - \tau)}(z^*(v); 1 - \tau^*) dv = 0,
\end{aligned}$$

where $z^*(v) := z(v, 1 - \tau^*)$. Specializing to the quasilinear/constant elasticity case, rearranging and expressing in terms of aggregate elasticities and attributes of the CEO income and firm profit distributions evaluated at the optimum (and denoted by stars) gives:

$$\text{(II.18)} \quad \tau^* = \frac{1}{1 + A_W^* \frac{\mathcal{E}_W^* + \chi \frac{\Pi^*}{W^*} \mathcal{E}_\Pi^*}{1 - \phi - (\chi - \phi) A_W^* \frac{\Pi^*}{W^*} \mathcal{E}_\Pi^*}}.$$

THE $\chi = 1$ CASE

In the main body of the paper we show that the optimal linear tax rate when $\chi = 1$ is:

$$\text{(II.19)} \quad \tau^* = \frac{1}{1 + A_W^* \frac{\mathcal{E}_W^* + \frac{\Pi^*}{W^*} \mathcal{E}_\Pi^*}{1 - A_W^* \frac{\Pi^*}{W^*} \mathcal{E}_\Pi^*}},$$

where the aggregate elasticities in (II.19) are evaluated at the optimum. We now show that under the assumptions of quasi-linear/constant elasticity preferences and a constant talent Pareto coefficient (II.19) reduces to (34) in the main text. After simple calculations:

$$\text{(II.20)} \quad \frac{\Pi^*}{W^*} \mathcal{E}_\Pi^* = \frac{R^*}{W^*} \mathcal{E}_R^* - \mathcal{E}_W^*,$$

where $R^* = \int_0^{v_0} DS(v)z(v, 1 - \tau^*)dv$ and:

$$(II.21) \quad \mathcal{E}_R^* := \frac{1 - \tau}{R} \frac{dR}{d(1 - \tau)} \Big|_{1 - \tau^*} = \int_0^{v_0} \frac{r^*(v)}{R^*} \frac{1 - \tau^*}{z(v, 1 - \tau^*)} \frac{dz(v, 1 - \tau^*)}{d(1 - \tau)} dv,$$

with $r^*(v) = DS(v)z(v, 1 - \tau^*)$. It then follows from (II.19) and (II.20) that:

$$(II.22) \quad \tau^* = \frac{1}{1 + \frac{\mathcal{E}_R^*}{\frac{W^*}{R^*} \frac{1}{A_W^*} - \frac{\Pi^*}{R^*} \mathcal{E}_\Pi^*}}.$$

From the definition of A_W^* ,

$$\frac{W^*}{A_W^*} = \Delta W^* = \int_0^{v_0} \{w^*(v) - w_0\} dv.$$

Next using the definition of the local Pareto coefficient $\alpha_w^*(v) := \frac{m(w^*(v))w^*(v)}{1 - M(w^*(v))}$ and the fact that $v = 1 - M(w^*(v))$,

$$\int_0^{v_0} \frac{w^*(v)}{\alpha_w^*(v)} dv = \int_0^{v_0} \frac{1 - M(w^*(v))}{m(w^*(v))} dv = \int_0^{v_0} w_v^*(v) v dv = \int_0^{v_0} \{w^*(v) - w_0\} dv,$$

where the last equality is via an integration by parts. Hence,

$$(II.23) \quad \frac{W^*}{A_W^*} = \int_0^{v_0} \frac{w^*(v)}{\alpha_w^*(v)} dv.$$

Denoting the local Pareto coefficients for CEO talent and firm profit (at the optimum) by α_h and α_π^* respectively and using (21) in the main body of the paper gives:

$$(II.24) \quad \frac{w}{\alpha_w^*} = DSz(1 + \mathcal{E}^u) \frac{1}{\alpha_h} + \pi \mathcal{E}^c \frac{1}{\alpha_\pi^*}.$$

So, combining (II.23) and (II.24) and evaluating at the optimum:

$$\frac{\Delta W^*}{R^*} = \frac{1}{R^*} \frac{W^*}{A_W^*} = \int_0^{v_0} \frac{r^*}{R^*} (1 + \mathcal{E}^{u*}) \frac{1}{\alpha_h} dv + \int_0^{v_0} \frac{\pi^*}{R^*} \mathcal{E}^{c*} \frac{1}{\alpha_\pi^*} dv.$$

Consequently,

(II.25)

$$\begin{aligned} \frac{1}{R^*} \frac{W^*}{A_W^*} - \frac{\Pi^*}{R^*} \mathcal{E}_{\Pi}^* &= \int_0^{v_0} \frac{r^*}{R^*} (1 + \mathcal{E}^{u^*}) \frac{1}{\alpha_h} dv + \frac{1}{R^*} \int_0^{v_0} \pi^* \mathcal{E}^{c^*} \frac{1}{\alpha_{\pi}^*} dv \\ &\quad + \frac{1 - \tau^*}{R^*} \int_0^{v_0} \int_{z_0}^{z^*(v)} \left\{ DS_v(v(z', 1 - \tau)) \frac{\partial v}{\partial(1 - \tau)}(z', 1 - \tau^*) \right\} dz' dv. \end{aligned}$$

Together (II.21), (II.22) and (II.25) imply:

$$\tau^* = \frac{1}{1 + \frac{\int_0^{v_0} \frac{r^*}{R^*} \frac{1 - \tau}{z} \frac{dz}{d(1 - \tau)} dv}{\int_0^{v_0} \frac{r^*}{R^*} (1 + \mathcal{E}^{u^*}) \frac{1}{\alpha_h} dv + \frac{1}{R^*} \int_0^{v_0} \mathcal{E}^{c^*} \frac{\pi^*}{\alpha_{\pi}^*} dv + \frac{1 - \tau^*}{R^*} \int_0^{v_0} \int_{z_0}^{z^*(v)} \left\{ DS_v(v(z', 1 - \tau)) \frac{\partial v}{\partial(1 - \tau)}(z', 1 - \tau^*) \right\} dz' dv}}.$$

This formula is greatly simplified in the quasilinear/constant effort elasticity case. Then $\frac{1 - \tau}{z} \frac{dz}{d(1 - \tau)}$, \mathcal{E}^c and \mathcal{E}^u take the common value \mathcal{E} which can be pulled through integrals. In this case, using (II.13) gives:

$$\begin{aligned} \frac{1}{R^*} \frac{W^*}{A_W^*} - \frac{\Pi^*}{R^*} \mathcal{E}_{\Pi}^* &= (1 + \mathcal{E}) \int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv + \frac{\mathcal{E}}{R^*} \int_0^{v_0} \frac{\pi^*}{\alpha_{\pi}^*} dv \\ &\quad - \frac{\mathcal{E}}{R^*} \int_0^{v_0} \left\{ DS(v) z^*(v) - DS_0 z_0 + \int_v^{v_0} w_v^*(v') dv' \right\} dv \\ &= (1 + \mathcal{E}) \int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv + \frac{\mathcal{E}}{R^*} \int_0^{v_0} \frac{\pi^*}{\alpha_{\pi}^*} dv \\ &\quad - \frac{\mathcal{E}}{R^*} \int_0^{v_0} \{ DS(v) z^*(v) - w^*(v) - \{ DS_0 z_0 - w_0 \} \} dv. \end{aligned}$$

Then using $\pi^*(v) = DS(v) z^*(v) - w^*(v)$, $\pi_0 = DS_0 z_0 - w_0$ and $\int_0^{v_0} \frac{\pi^*}{\alpha_{\pi}^*} dv = \int_0^{v_0} \{ \pi^*(v) - \pi_0 \} dv$, the previous equation reduces to:

$$\frac{1}{R^*} \frac{W^*}{A_W^*} - \frac{\Pi^*}{R^*} \mathcal{E}_{\Pi}^* = (1 + \mathcal{E}) \int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv.$$

The right hand side of this expression is a weighted integral of (reciprocals of) talent local Pareto coefficients. Furthermore, $\mathcal{E}_R^* = \mathcal{E}$. Hence,

$$\tau^* = \frac{1}{1 + \frac{\mathcal{E}}{1 + \mathcal{E}} \int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv}}.$$

If the local talent Pareto coefficient is constant, then the previous formula reduces to:

$$\tau^* = \frac{1}{1 + \frac{\varepsilon}{1+\varepsilon} \frac{1}{\alpha_h}}.$$

INDEPENDENCE OF $\mathcal{E}_W A_W$ FROM THE MARGINAL TAX RATE

We conclude this appendix by showing that the product $\mathcal{E}_W A_W$ is independent of the marginal tax rate under the assumptions of quasilinear/constant elasticity CEO preferences, a multiplicative firm objective and a tax system that is linear above a threshold. This assumption allows us to relate estimates of the product $\mathcal{E}_W A_W$ in US data to the optimal value and, hence, the optimal tax rate in the Diamond-Saez formula. Note that from (II.15) under the assumptions just described:

$$A_W \mathcal{E}_W = \frac{W}{\Delta W} \frac{\varepsilon}{W} \int_0^{v_0} \{DS_0 z_0 - w_0 + w(v)\} dv = \varepsilon \left\{ 1 + \frac{DS_0 z_0}{\Delta W} (1 - v_0) \right\}$$

and

$$\Delta W = \int_0^{v_0} \{w(v) - w_0\} dv = \int_0^{v_0} w_v(v) v dv = \int_0^{v_0} \mathcal{H}(v) z(v) dv,$$

with $\mathcal{H}(v)$ independent of marginal taxes. Hence,

$$\frac{\partial \Delta W}{\partial (1 - \tau)} = \frac{\varepsilon}{1 - \tau} \Delta W.$$

Similarly,

$$\frac{\partial DS_0 z_0}{\partial (1 - \tau)} = \frac{\varepsilon}{1 - \tau} DS_0 z_0.$$

Thus, the ratio $\frac{DS_0 z_0}{\Delta W}$ is unaffected by the marginal tax rate and neither is the product $A_W \mathcal{E}_W$.

III. Elasticity of Firm Profits to Top Income Tax Rate

In this section we take a closer look at the empirical evidence on the elasticity of firm profits to the top income tax rate, i.e. on the magnitude of \mathcal{E}_Π . The literature gives little guidance on the size of this elasticity. However, Figure III.1 provides some suggestive evidence. It displays the time series for net corporate dividends of US domestic industries and the top marginal tax rate on income and corporate profits over the period 1919 to 2014.¹ We begin looking at the

¹Top marginal rates are taken from the *Tax Foundation*. Data on dividends from 1940 to 2014 is from the BEA Table 6.20A (series: A3302C0) data on GDP is from Table 1.1.5. Data on dividends from 1919 to 1939 is

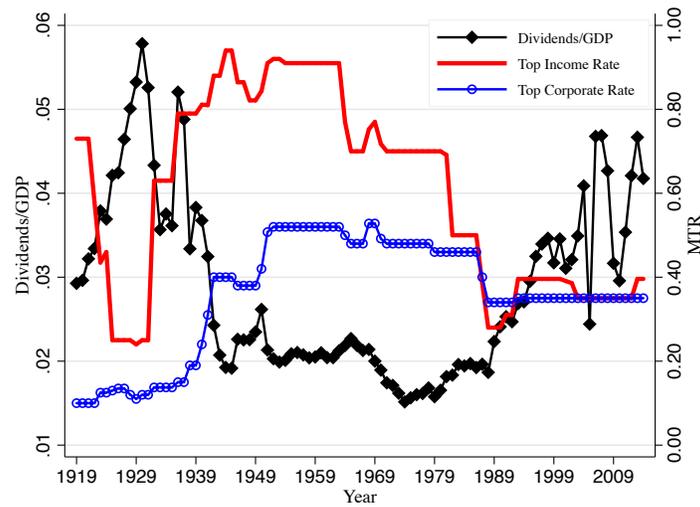


Figure III.1. : The relationship between net corporate dividend payments (left axis) and marginal top income and corporate tax rates (right axis).

relationship between dividends and top marginal income tax rate. A negative correlation between these variables is apparent. Along the lines of Piketty, Saez, and Stantcheva (2014) (who focus on the relationship between top incomes and the retention rate), we estimate the following log linear relationship:

$$(III.1) \quad \log(\text{Dividends}/\text{GDP}_t) = \beta_0 + \beta_1 \log(1 - \bar{\tau}_t) + \varepsilon_t,$$

where $\bar{\tau}_t$ is the top marginal tax rate on income at time t . Estimates for β_1 from 1919 to 2014 provide a positive and statistically significant elasticity of 0.232 (0.042). Of course, this does not establish a casual relationship between the time series as an omitted third factor might be responsible for both time profiles. A candidate for this third factor is the top marginal rate on corporate profits, also displayed in Figure III.1. The two top marginal tax rates (after 1940) display very similar profiles. We complement the relationship (III.1) with the factor $\beta_2 \log(1 - \bar{\tau}_t^c)$ where $\bar{\tau}_t^c$ is the top marginal corporate tax rate at time t . The estimate of β_1 in this case is equal to 0.0574 (0.0333). In this case we can reject a value of $\beta_1 = 0$ at the 10 per cent level. Overall the estimate for the elasticity is close to the value implied for \mathcal{E}_{Π} by the structural model in the body of the paper. A formal connection between the two, however, cannot be fully established

from the NBER Macroeconomy Database (series: a08185). The NBER and BEA dividend series overlap between 1929 to 1939, differences are small but systematic. We use the average difference observed between 1929 to 1939 to adjust the NBER data.

as in the assignment model profit is defined to be the rents accruing to owners of the asset S . These rents exclude payments to adjustable capital and may not be realized contemporaneously with the application of CEO effective labor.

IV. Derivation of optimal nonlinear tax formulas

When firm claimants have welfare weight χ , CEOs have welfare weights ψ , the policymaker's (relaxed) mechanism design problem reduces to:

$$(IV.1) \quad \sup_{\Phi, z, w, \bar{v}} \int_0^{\bar{v}} \{ \chi V(S(v), z(v)) + (1 - \chi)w(v) - \mathcal{C}[\Phi(v), z(v)/h(v)] + \psi(v)\Phi(v) \} dv \\ + \int_{\bar{v}}^1 \{ \psi(v)\bar{U} + T^0 \} dv,$$

subject to $\bar{v} \in I = (0, 1]$, $\Phi(\bar{v}) = \bar{U}$, $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) \geq 0$, and for almost all $v \in I$,

$$\Phi_v(v) = -U_e \left(\mathcal{C} \left[\Phi(v), \frac{z(v)}{h(v)} \right], \frac{z(v)}{h(v)} \right) \frac{z(v)}{h(v)} \frac{h_v(v)}{h(v)} \\ w_v(v) = V_z(S, z)z_v(v).$$

In the relaxed problem (IV.1) the incentive constraints on firms and CEO's are replaced with the firms' first order and the CEOs' envelope conditions respectively. Monotonicity of optimal z and w (and, hence, π) for the relaxed problem are checked ex post in all of our numerical calculations. Problem (IV.1) may be formulated as an optimal control problem in which Φ , z and w are the state variables, z_v is the control variable and \bar{v} is a choice variable. The Hamiltonian for this optimal control problem is:

$$\mathcal{H}(v) = -p^\Phi(v)U_e \left(\mathcal{C} \left[\Phi(v), \frac{z(v)}{h(v)} \right], \frac{z(v)}{h(v)} \right) \frac{z(v)}{h(v)} \frac{h_v(v)}{h(v)} \\ + p^z(v)z_v(v) + p^w(v)V_z(S(v), z(v))z_v(v) \\ + \chi V(S(v), z(v)) + (1 - \chi)w(v) - \mathcal{C} \left[\Phi(v), \frac{z(v)}{h(v)} \right] + \psi(v)\Phi(v),$$

with co-states p^Φ , p^z and p^w . Let q^V denote the multiplier on $V(S(\bar{v}), z(\bar{v})) - w(\bar{v}) \geq 0$ and q^U the multiplier on $\Phi(\bar{v}) - \bar{U} = 0$. The first order condition for z_v implies that:

$$(IV.2) \quad p^z + p^w V_z(S, z) = 0.$$

Differentiating (IV.2) with respect to v gives:

$$(IV.3) \quad p_v^z + p_v^w V_z(S, z) + p^w [V_{zS}(S, z)S_v + V_{zz}(S, z)z_v] = 0.$$

The optimal co-state equations are:

$$(IV.4) \quad p_v^\Phi = \frac{1}{U_c} + p^\Phi \frac{U_{ec} z h_v}{U_c h} - \psi$$

$$(IV.5) \quad p_v^w = -(1 - \chi)$$

$$(IV.6) \quad p_v^z = p^\Phi \left\{ \left[U_{ec} \left(-\frac{U_e}{U_c} \right) + U_{ee} \right] \frac{z}{h} + U_e \right\} \frac{h_v}{h} \frac{1}{h} - p^w V_{zz} z_v - \frac{U_e}{U_c} \frac{1}{h} - \chi V_z.$$

The first order condition for \bar{v} is:

$$(IV.7) \quad \mathcal{H}(\bar{v}) - \psi(\bar{v})\bar{U} - T^0 + q^V V_S(S(\bar{v}), z(\bar{v})) S_v(\bar{v}) \geq 0,$$

with equality if $\bar{v} \in (0, 1)$. The transversalities at $v = \bar{v}$ are:

$$(IV.8) \quad p^w(\bar{v}) = -q^V \quad p^z(\bar{v}) = q^V V_z(S(\bar{v}), z(\bar{v})) \quad p^\Phi(\bar{v}) = q^U.$$

There are also transversalities at $v = 0$. This is because (unlike typical optimal control problems), there are no initial conditions for w and Φ . We have:

$$\lim_{v \downarrow 0} p^w(v) = \lim_{v \downarrow 0} p^z(v) = \lim_{v \downarrow 0} p^\Phi(v) = 0.$$

Integrating (IV.4) gives:

$$(IV.9) \quad p^\Phi(v) = \int_0^v \left[\left(\frac{1}{U_c(u)} - \psi(u) \right) \exp \left\{ - \int_v^u \frac{U_{ec}(u')}{U_c(u')} \left(\frac{e(u')}{h(u')} \right) h_v(u') du' \right\} \right] du,$$

while integrating (IV.5) gives

$$(IV.10) \quad p^w(v) = -(1 - \chi)v.$$

Combining conditions (IV.3) to (IV.6) and (IV.10) implies:

$$(IV.11) \quad V_z + \frac{U_e/h}{U_c} = -\frac{p^\Phi}{h} \left\{ \left[U_{ec} \left(-\frac{U_e}{U_c} \right) + U_{ee} \right] \frac{z}{h} + U_e \right\} \frac{-h_v}{h} + (1 - \chi)v V_{zS}(-S_v).$$

As described in the main text, this optimality condition captures the marginal benefits and costs associated with a small change in CEO v 's effective labor (holding her utility fixed). Together the CEOs' and the firms' first order conditions imply:

$$(IV.12) \quad (1 - T_w[w])V_z h U_c = -U_e.$$

The optimal marginal tax rate must be set to align a CEO's private return to effort with the social return. Since the CEO's effective wage coincides with the firm's marginal product, this reduces to ensuring that her pre-tax return on effort equates to the right hand side of (IV.11). If $\chi = 1$, then the right hand side of (IV.11) equals the usual marginal informational rents term from the Mirrlees model and combining (IV.11) and (IV.12) and the definition of α_h the standard formula for optimal marginal tax rates obtains:

$$(IV.13) \quad T_w[w] = \frac{1}{1 + \frac{1}{\bar{p}^\Phi} \frac{\mathcal{E}^c}{1+\mathcal{E}^u} \alpha_h},$$

where $\bar{p}^\Phi = \frac{U_c p^\Phi}{1-F(h)}$ is the normalized co-state. More generally, using (IV.11) and (IV.12) and the definitions of α_h and α_S , the optimal marginal tax rate is:

$$(IV.14) \quad T_w[w] = \frac{1 + (1 - \chi) \frac{1}{\bar{p}^\Phi} \frac{\mathcal{E}^c}{1+\mathcal{E}^u} \frac{V_{zS} S}{V_z} \frac{\alpha_h}{\alpha_S}}{1 + \frac{1}{\bar{p}^\Phi} \frac{\mathcal{E}^c}{1+\mathcal{E}^u} \alpha_h}.$$

SPECIAL UTILITY CASES

In the case of quasilinear/constant elasticity $U(c, e) = c - \frac{\mathcal{E}}{1+\mathcal{E}} e^{\frac{1+\mathcal{E}}{\mathcal{E}}}$ and a multiplicative firm surplus DSz , (IV.14) reduces to:

$$(IV.15) \quad T_w[w] = \frac{1 + (1 - \chi) \frac{\mathcal{E}}{1+\mathcal{E}} \frac{\alpha_h}{\alpha_S}}{1 + \frac{\mathcal{E}}{1+\mathcal{E}} \alpha_h}.$$

If instead CEO utility has the log-constant compensated elasticity form $U(c, e) = \log c - \log \left(1 + \frac{\mathcal{E}}{1+\mathcal{E}} e^{\frac{1+\mathcal{E}^c}{\mathcal{E}^c}} \right)$, then (IV.14) becomes:

$$(IV.16) \quad T_w[w] = \frac{1 + (1 - \chi) \frac{A_c - 1}{A_c} \frac{\mathcal{E}^c}{1+\mathcal{E}^u} \frac{\alpha_h}{\alpha_S}}{1 + \frac{A_c - 1}{A_c} \frac{\mathcal{E}^c}{1+\mathcal{E}^u} \alpha_h},$$

where $A_c = \int_0^v c(v') dv' / \int_0^v \{c(v') - c(v)\} dv'$ is the tail coefficient of consumption and c the consumption allocation. In Section IV, the quantitative section of this paper, we focus on the quasilinear/constant elasticity CEO preference case. If preferences are modified to be log-constant compensated elasticity, the formula (53) in the paper for recovering α_h from the data is unchanged up to the modification of the coefficients \mathcal{N} and \mathcal{P} to

$$\mathcal{N} := \frac{1}{D\mathcal{E}} - \frac{\mathcal{E}^c}{1 + \mathcal{E}^u} \quad \text{and} \quad \mathcal{P} := -\frac{\mathcal{E}^c}{1 + \mathcal{E}^u}.$$

Saez (2001) considers utility functions of the log-constant compensated elasticity form and selects compensated elasticity values of \mathcal{E}^c equal to 0.25 and 0.5. At high incomes \mathcal{E}^u is close to zero (and less than \mathcal{E}^c). These values thus imply values for $\frac{\mathcal{E}^c}{1+\mathcal{E}^u}$ that are greater than the values implied by our calibration for $\frac{\mathcal{E}}{1+\mathcal{E}}$. Equation (55) then implies even larger values of α_h than our calibration. Turning to (IV.16) evaluated at $\chi = 1$, it follows that tax rates are reduced by the increased values for $\frac{\mathcal{E}^c}{1+\mathcal{E}^u}\alpha_h$, but increased by $\frac{A_c-1}{A_c} < 1$.

DETERMINING \tilde{v}

Substituting the definition of the Hamiltonian and (IV.8) into (IV.7) gives:

$$(IV.17) \quad -p^\Phi(\tilde{v})U_e(\tilde{v})\frac{z(\tilde{v})}{h(\tilde{v})}\frac{h_v(\tilde{v})}{h(\tilde{v})} + \chi V(S(\tilde{v}), z(\tilde{v})) - T^0 \\ + (1 - \chi)w(\tilde{v}) - \mathcal{C}\left[\bar{U}, \frac{z(\tilde{v})}{h(\tilde{v})}\right] + q^V V_S(S(\tilde{v}), z(\tilde{v}))S_v(\tilde{v}) \geq 0.$$

If $\chi = 1$, then the preceding expression reduces to:

$$(IV.18) \quad V(S(\tilde{v}), z(\tilde{v})) - \mathcal{C}\left[\bar{U}, \frac{z(\tilde{v})}{h(\tilde{v})}\right] - T^0 - p^\Phi(\tilde{v})U_e(\tilde{v})\frac{z(\tilde{v})}{h(\tilde{v})}\frac{h_v(\tilde{v})}{h(\tilde{v})} \\ + q^V V_S(S(\tilde{v}), z(\tilde{v}))S_v(\tilde{v}) \geq 0.$$

The term $V(S(\tilde{v}), z(\tilde{v})) - \mathcal{C}\left[\bar{U}, z(\tilde{v})/h(\tilde{v})\right] - T^0$ captures the direct loss of social surplus from shutting down the \tilde{v} firm. The remaining terms capture the benefits to closing the firm down in terms of reduced rents to firms and CEOs ranked above \tilde{v} . Note that (IV.18) does not pin down $w(\tilde{v})$, which may take any value in $[0, V(S(\tilde{v}), z(\tilde{v}))]$. When $\chi = 1$, the policymaker does not care whether the social surplus at \tilde{v} is realized as firm profit $V(S(\tilde{v}), z(\tilde{v})) - w(\tilde{v})$ (perhaps to be captured by taxation on claimants) or realized as CEO income tax revenues $T[w(\tilde{v})] = w(\tilde{v}) - \mathcal{C}\left[\bar{U}, z(\tilde{v})/h(\tilde{v})\right]$. From (IV.8) and (IV.10), $q^V = -p^w(\tilde{v}) = (1 - \chi)\tilde{v} \geq 0$. Thus, if $\chi < 1$, then $q^V > 0$, $V(S(\tilde{v}), z(\tilde{v})) - w(\tilde{v}) = 0$ and (IV.17) reduces to:

$$(IV.19) \quad V(S(\tilde{v}), z(\tilde{v})) - \mathcal{C}\left[\bar{U}, \frac{z(\tilde{v})}{h(\tilde{v})}\right] - T^0 + (1 - \chi)\tilde{v}V_S(S(\tilde{v}), z(\tilde{v}))S_v(\tilde{v}) \\ - p^\Phi(\tilde{v})U_e(\tilde{v})\frac{z(\tilde{v})}{h(\tilde{v})}\frac{h_v(\tilde{v})}{h(\tilde{v})} \geq 0.$$

Inequality (IV.19) is interpreted similarly to (IV.18). Now, however, firm \tilde{v} generates no profit; everything above the consumption amount necessary to give CEO \tilde{v} utility \bar{U} is captured by the policymaker in the form of taxation on CEO \tilde{v} .

A. *Firm Entry and the Marginal Social Value of Profit*

The preceding analysis treats the marginal social value of aggregate profit χ as a fixed parameter. In Section II in the paper we discuss the interpretation and determination of χ . If the policymaker is solely concerned with maximizing tax revenues, then, in our baseline model, χ equals the tax on profit τ^F . Since profit is pure rent, it is optimal to select τ^F equal to 100% and, hence, χ equal to 1. In Section II we also describe how the firm's ability to privately transfer resources to firm owners may restrain profit taxation and the marginal social value of profit below 1. It is natural to conjecture that in a richer model with a firm entry margin, the disincentive to create firms also deters high profit taxation and, hence, leads to a lower marginal social value of profit. In fact, while a firm entry margin can create a rationale for lower rates of profit taxation, it can also introduce an additional motive for valuing profit: after-tax profits retained by firms relax firm entry conditions. Consequently, the marginal social value of aggregate profit can exceed the profit tax rate. If firm entry conditions involve aggregate profit and there are no restrictions on the level of CEO incomes and taxes, then the marginal social value of profit can equal one independent of the profit tax.

We develop two simple extensions of our model that highlight the issues.² Through out the body of the paper and previously in this section we assumed a fixed population of candidate firms (because, for example, a fixed population of candidate entrepreneurs draw a finite number of business ideas from a given quality distribution) and a firm entry condition:

$$V(S(v), z(v)) - w(v) \geq 0.$$

This is equivalent to assuming that firm entry costs are fully deductible and that a firm knows its potential size $S(v)$ before entering. Since firm profit is pure rent, then, as noted, the optimal profit tax and the corresponding marginal social value of profit are both one. If instead firms pay a common and non-deductible entry cost, then the relevant entry condition becomes:

$$(1 - \tau^F)\{V(S(v), z(v)) - w(v)\} \geq k > 0.$$

It remains the case that the marginal social value of profit χ equals the profit tax τ^F . Thus, restrictions on profit taxes, say from profit diversion, continue to restrain the marginal social value of profit at the optimum. Absent such restrictions (and absent restrictions on the income received $w(\tilde{v})$ and taxes paid $T[w(\tilde{v})]$ by the least talented active CEO and, hence, the *level* of the CEO income and tax schedules), the optimal CEO consumption and effective labor allocation equals that from a model with a unit shadow social weight on profit.

²We thank a referee for encouraging us to explore this issue.

The tax/wage implementation requires setting the profit tax (arbitrarily close) to 100% and the pre-tax income and taxes of the least talented CEO to arbitrarily small (negative) numbers. Any additional restriction on the level of CEO incomes or taxes prevents this optimum from being achieved. Further details are available on request.

Consider next the following alternative model of entry.³ Before knowing its potential asset size S , a (candidate) firm draws a nondeductible entry cost k from a distribution L . If the firm chooses to enter, then it pays its cost k and draws an asset size S from a distribution G . If a firm draws $S(v)$, then it makes after-tax profit $(1 - \tau^F)\{V(S(v), y(v)) - w(v)\}$ if it chooses to produce and zero otherwise. A firm chooses to enter and pay the cost k if:

$$(1 - \tau^F) \int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv - k \geq 0,$$

where as before $V(S(\bar{v}), y(\bar{v})) - w(\bar{v}) = 0$. Given τ^F , y and w , there will be a cost threshold \hat{k} such that the $L(\hat{k})$ firms with costs below this threshold choose to enter the market and the rest do not and:

$$(1 - \tau^F) \int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv - \tilde{k} = 0.$$

The central difference between this model and that in the last paragraph is that now “aggregate” profit, i.e. ex ante expected profit $\int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv$, rather than the profit of the smallest producing firm $V(S(\bar{v}), z(\bar{v})) - w(\bar{v})$ relaxes the entry constraint. Total tax revenues are:

$$L(\hat{k}) \left\{ \tau^F \int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv + \int_0^{\bar{v}} T[w(v)] dv \right\},$$

where to simplify the analysis tax revenues collected from unassigned CEOs are abstracted from. The objective may be rewritten as:

$$L(\hat{k}) \left\{ \tau^F \int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv + \int_0^{\bar{v}} [w(v) - c(v)] dv \right\}.$$

Substituting the binding entry constraint into the objective, the latter may be re-expressed (on the space of feasible allocations) as:

$$L(\hat{k}) \left\{ \int_0^{\bar{v}} \{V(S(v), y(v)) - w(v)\} dv + \int_0^{\bar{v}} [w(v) - c(v)] dv - \hat{k} \right\}.$$

Note that this re-specified objective places equal weight on firm profits and CEO

³For a related model, see Scheuer and Werning (2015).

tax revenues. Given an optimal choice of \hat{k} , y and c and \hat{v} may be chosen to maximize:

$$\int_0^{\hat{v}} \{V(S(v), y(v)) - c(v)\} dv,$$

subject to the CEO incentive and participation constraints. Thus, firm profits and CEO income tax revenues are weighted equally, with the (normalized) marginal social value of profit set to 1. The (c, y) allocation pins down w_v via the firms' first order conditions, while τ^F and $w(\hat{v})$ are chosen to satisfy:

$$(1 - \tau^F) \int_0^{\hat{v}} \left\{ V(S(v), y(v)) - w(\hat{v}) - \int_v^{\hat{v}} w_v(v') dv' \right\} dv - \hat{k} = 0$$

and $V(S(\hat{v}), y(\hat{v})) - w(\hat{v}) = 0$. Thus, if the firm entry constraint binds at $\hat{k} > 0$, the optimal value of $\tau^F < 1$.

Proceeding slightly differently, if $L(\hat{k})\psi$ is the multiplier on the firm entry condition, then, after normalizing by $L(\hat{k})$, the planner's Lagrangian inclusive of the firm entry condition implies that the shadow social weight on firm profit is:

$$\chi = \tau^F + \psi(1 - \tau^F).$$

Thus, the marginal social value of profit is enhanced by its impact on the firm entry constraint. The first order condition for $w(\hat{v})$ is:

$$(1 - \tau^F) - \psi(1 - \tau^F) = 0,$$

which implies that $\psi = 1$. Hence, up to the determination of \hat{k} , the model is equivalent to one with equal social weighting of profit and CEO income tax revenues ($\chi = 1$) and no explicit firm entry constraint. Note that this is true regardless of the choice of τ^F . That $\chi = 1$ stems from the fact that the planner can transfer resources from tax revenues to firm profits one for one by adjusting the *level* of taxes and firm wages.

In summary, relaxation of the firm entry constraint by aggregate firm profits can enhance their social value beyond their direct contribution to tax revenues. Hence, the marginal social value of profit can exceed the profit tax rate. If aggregate profits enter the firm entry constraint (and there are no binding restrictions on $w(\hat{v})$ and $T[w(\hat{v})]$ and, hence, the *level* of CEO incomes and income taxes), then the marginal social value of profit can equal that of taxes.

V. Connection with Gabaix and Landier (2008)

Our estimates of the talent distribution of CEOs are consistent with the evidence presented in Gabaix and Landier (2008). In this paper, the authors calibrate the tail index of the distribution of talent using evidence on the distribution of firm size and the pay to firm-size elasticity. Their evidence is consistent

with prior work that characterizes the distribution of firm size as a Pareto distribution with coefficient approximately equal to 1 and with “Roberts’ Law” (the empirical regularity relating CEO compensation and firm size). These two facts imply a negative tail index of CEO talent consistent with a Weibull distribution. A previous version of our paper estimated the shape parameter of a GEV distribution fitted on CEO talent. For that case estimates of the shape parameter also implied a Weibull distribution in our setting. Our current approach is complementary to the one in Gabaix and Landier (2008). First, Gabaix and Landier (2008) use the data directly to guide them in selecting market capitalization as an empirical counterpart to S . We follow Tervio (2008) in using economic theory to connect firm market capitalization to firm surplus. This theory explicitly allows for long lasting effects of CEO labor and for adjustable capital. Second, Gabaix and Landier (2008) focus on asymptotic tail properties of the CEO talent distribution. They appeal to extreme value theory to justify functional form restrictions on this distribution. We instead completely characterize the right tail of the CEO talent and firm asset distributions. Knowledge of both these distributions is needed for the computation of the non-linear tax schedule for values of $\chi \neq 1$. Third, with respect to the model of Gabaix and Landier (2008), we allow for elastic labor. Finally, there are differences in the data used. Specifically, we use measures of CEO compensation calculated using the value of options when exercised; Gabaix and Landier (2008) use the value of options when granted.

VI. Mechanical and Behavioral Impacts of Local Marginal Tax Rate Changes

Here we compute the mechanical and behavioral impacts of marginal tax rate changes in our assignment setting. We focus on the quasilinear/constant elasticity CEO utility-multiplicative firm payoff setting. Let T denote a twice continuously differentiable tax function and consider the following perturbed function:

$$\tilde{T}(w) = \begin{cases} T[w] & w \in [0, w_0) \\ T[w] + \sqrt{\delta}[w - w_0] & w \in [w_0, w_0 + \sqrt{\delta}) \\ T[w] + \delta & w \in [w_0 + \sqrt{\delta}, \infty) \end{cases}.$$

Let \hat{w} and \hat{z} denote the initial equilibrium schedules for income and effective labor and let \tilde{w} and \tilde{z} denote the equilibrium schedules occurring after the tax perturbation. Let v_0 and z_0 be given by $w(v_0) = w_0$ and $z_0 = z(v_0)$.

CEOs (and firms) can be partitioned into four groups which we label groups 0,1,2 and 3 respectively. Group 0 consists of the least talented CEOs with types in $[v_0, 1]$. Their behavior is unaffected by the tax perturbation. Group 1 CEOs with types in $[v_1(\delta), v_0)$ bunch at the kink point in the tax schedule. They earn w_0 and supply effective labor z_0 . The threshold $v_1(\delta)$ satisfies:

$$(1 - T_w[w_0] - \sqrt{\delta})S(v_1(\delta))h(v_1(\delta)) = \left(\frac{z_0}{h(v_1(\delta))} \right)^{\frac{1}{\epsilon}}.$$

Group 2 CEOs have types $(v_2(\delta), v_1(\delta))$ earn incomes between w_0 and $w_0 + \sqrt{\delta}$ and pay the higher marginal tax $T_w[w] + \sqrt{\delta}$. Their first order condition for effective labor supply is given by:

$$(1 - T_w[\tilde{w}(v; \delta)] - \sqrt{\delta})S(v)h(v) = \left(\frac{\tilde{z}(v; \delta)}{h(v)} \right)^{\frac{1}{\varepsilon}},$$

where the notation makes the dependence of the functions \tilde{w} and \tilde{z} on δ explicit. The threshold $v_2(\delta)$ is given by:

$$v_2(\delta) = \sup\{v : \tilde{w}(v; \delta) \geq w_0 + \sqrt{\delta}\}.$$

CEOs in group 3 have ranks $(0, v_2(\delta))$. Their marginal taxes are determined by the original tax schedule and their first order conditions are given by:

$$(1 - T_w[\tilde{w}(v; \delta)])S(v)h(v) = \left(\frac{\tilde{z}(v; \delta)}{h(v)} \right)^{\frac{1}{\varepsilon}}.$$

At $v_2(\delta)$ there is a discontinuity in \tilde{w} and \tilde{z} . Firms and CEOs at and arbitrarily close to $v_2(\delta)$ must be optimizing. In particular, this implies that:

$$S(v_2)\tilde{z}(v_2(\delta); \delta) - \tilde{w}(v_2(\delta); \delta) = S(v_2(\delta); \delta)\tilde{z}_+(v_2(\delta); \delta) - \tilde{w}_+(v_2(\delta); \delta),$$

where $\tilde{z}_+(v_2(\delta))$ and $\tilde{w}_+(v_2(\delta))$ are, respectively, the right limits of \tilde{z} and \tilde{w} at $v_2(\delta)$ (i.e. the limits of effective labor supplied and incomes earned by group 2 CEOs).

The government's revenues after the perturbation are given by:

(VI.1)

$$R(\delta) := \int_0^{v_2(\delta)} \{T[\tilde{w}(v; \delta)] + \delta\}dv + \int_{v_2(\delta)}^{v_0} \{T[\tilde{w}(v; \delta)] + \sqrt{\delta}[\tilde{w}(v; \delta) - w_0]\}dv.$$

MECHANICAL EFFECT

The mechanical effect is obtained from the terms:

$$\int_0^{v_2(\delta)} \delta dv + \int_{v_2(\delta)}^{v_0} \sqrt{\delta}[\tilde{w}(v; \delta) - w_0]dv.$$

Totally differentiating this with respect to δ and setting δ to zero gives:

$$1 - M(w_0).$$

BEHAVIORAL EFFECT

Next we turn to the behavioral effect.⁴ It is obtained from the remaining terms in (VI.1):

$$(VI.2) \quad \int_0^{v_2(\delta)} T[\tilde{w}(v; \delta)] dv + \int_{v_2(\delta)}^{v_0} T[\tilde{w}(v; \delta)] dv.$$

Totally differentiating this with respect to δ and evaluating the limit as δ converges to 0 yields:

$$(VI.3) \quad \begin{aligned} & - \frac{T_w[w_0]}{1 - T_w[w_0]} \frac{\mathcal{E} \frac{S_0 z_0}{w_0}}{1 + \mathcal{E} \mathcal{T}[w_0] \frac{S_0 z_0}{w_0}} m(w_0) w_0 \\ & + \frac{T_w[w_0]}{1 - T_w[w_0]} \frac{\mathcal{E} \frac{S_0 z_0}{w_0}}{1 + \mathcal{E} \mathcal{T}[w_0] \frac{S_0 z_0}{w_0}} m(w_0) w_0 \frac{1}{g(S_0) S_0 T_w[w_0]} \\ & \times \int_0^{v_0} \frac{T_w[\hat{w}(v)]}{1 + \mathcal{E} \mathcal{T}[\hat{w}(v)] \frac{S(v) \hat{z}(v)}{\hat{w}(v)}} \exp \left\{ - \int_v^{v_0} \frac{\frac{S_v(v')}{S(v')} \mathcal{E} \mathcal{T}[\hat{w}(v')]}{1 + \mathcal{E} \mathcal{T}[\hat{w}(v')] \frac{S(v') \hat{z}(v')}{\hat{w}(v')}}} dv' \right\} dv. \end{aligned}$$

Despite its complexity this expression has a straightforward interpretation. A higher marginal tax “at” w_0 induces CEOs at this income to work less hard causing a reduction in revenues. This effect is captured by the term on the first line of (VI.3). But it also raises the incomes of more talented CEOs with ranks $v \in (0, v_0)$ and, hence, the revenues collected from them. This is captured by the term spread across the second and third lines of (VI.3).

Adding the mechanical term as well, setting the sum to zero and rearranging gives the optimal marginal tax rate for $\chi = 0$,

$$T_w[w] = \frac{1}{1 + \frac{m(w_0)w_0}{1-M(w_0)} \tilde{\mathcal{E}}_w(w_0)},$$

⁴In this appendix, we use the term “behavioral effect” to describe the overall effect of the tax rate change on CEO incomes and, hence, tax revenues. It consists of individual effective labor responses and an equilibrium response of the CEO income schedule.

where $\tilde{\mathcal{E}}_w(w_0)$ is defined as:

$$(VI.4) \quad \tilde{\mathcal{E}}_w(w_0) := \frac{\varepsilon \frac{S_0 z_0}{w_0}}{1 + \mathcal{E}\mathcal{T}[w_0] \frac{S_0 z_0}{w_0}} \left\{ 1 - \frac{1}{g(S_0)S_0} \frac{1}{T_w[w_0]} \int_0^{v_0} \frac{T_w[\hat{w}(v)]}{1 + \mathcal{E}\mathcal{T}[\hat{w}(v)] \frac{S(v)\hat{z}(v)}{\hat{w}(v)}} \exp \left\{ - \int_v^{v_0} \frac{S_v(v')}{S(v')} \mathcal{E}\mathcal{T}[\hat{w}(v')] \frac{S(v')\hat{z}(v')}{\hat{w}(v')} dv' \right\} dv \right\}.$$

Further manipulation establishes that at the $\chi = 0$ optimum,

$$\tilde{\mathcal{E}}_w(w_0) := \frac{\varepsilon \frac{S_0 z_0}{w_0}}{1 + \mathcal{E}\mathcal{T}(v_0) \frac{S_0 z_0}{w_0}} \left\{ 1 - \frac{1 - G(S_0)}{g(S_0)S_0} \right\}.$$

When $\chi > 0$, the impact of the marginal tax rate change on firm profits is also relevant. The impact of such a change is obtained from:

$$(VI.5) \quad \lim_{\delta \downarrow 0} \frac{\partial}{\partial \delta} \int_0^{v_0} \tilde{\pi}(v; \delta) dv,$$

where $\tilde{\pi}(v; \delta) := S(v)\hat{z}(v; \delta) - \tilde{w}(v; \delta)$. Calculating the limit in (VI.5) yields:

$$(VI.6) \quad m(w_0) \frac{1}{g(S_0)S_0} \frac{\mathcal{E}S_0 z_0}{1 - T_w[w_0]} \frac{1}{1 + \mathcal{E}\mathcal{T}[w_0] \frac{S_0 z_0}{w_0}} \times \int_0^{v_0} \exp \left\{ - \int_v^{v_0} \frac{S_v(v')}{S(v')} \mathcal{E}\mathcal{T}[\hat{w}(v')] \frac{S(v')\hat{z}(v')}{\hat{w}(v')} dv' \right\} dv.$$

The term in (VI.6) is the analogue of $\chi \frac{\Pi}{1-\tau} \mathcal{E}_{\Pi}$ in the derivation of (30) in the paper. Adding it to the other behavioral and mechanical terms leads to an optimal tax equation for the non-linear setting analogous to (30).

VII. Optimal Affine Tax Rates When $\chi = 0$

In this section we calculate the top affine tax rate for CEOs when $\chi = 0$. Recall that if the policymaker attaches no social weight to profits ($\chi = 0$) or CEOs, then in the affine setting the optimal tax rate on top earning CEOs is given by the Diamond-Saez formula:

$$(VII.1) \quad \tau^* = \frac{1}{1 + \mathcal{E}_W^* A_W^*}.$$

The values of the elasticity \mathcal{E}_W^* and the tail coefficient A_W^* are those arising in the optimal equilibrium. In general, they are endogenous and jointly determined with the optimal policy. However, under the assumption of quasilinear/constant elasticity CEO preferences, a multiplicative firm objective and linearity of taxes in incomes above a threshold, the product $\mathcal{E}_W A_W$ is independent

of the marginal tax rate (see Section II). Below we combine existing evidence on elasticities of taxable income and our own estimates of A_W in US data to obtain an estimate of the product $\mathcal{E}_W A_W$. We then recover the tax rate implied by formula (VII.1).

SELECTING A VALUE FOR \mathcal{E}_W

There is limited direct evidence on \mathcal{E}_W for CEO's. Bakija, Cole, and Heim (2012) estimate a fairly large elasticity of taxable income with respect to the retention rate of 0.7 for the top 0.1 per cent of US income earners using tax return data. In addition, they find that executives, managers, supervisors and financial professionals account for 60 per cent of the top 0.1 per cent income earners. Time series evidence shows a strong negative correlation between top marginal tax rates and CEO incomes in the US. However, regressions provided by Frydman and Malloy (2011) indicate a small contemporaneous response of CEO incomes to tax reforms. They reject a value of \mathcal{E}_W above 0.2. Goolsbee (2000) studies data from 1991 to 1995 and rejects an elasticity above 0.4. In the context of top income earners (but not necessarily CEOs), Diamond and Saez (2011) select a value for \mathcal{E}_W of 0.25. Given this range of values we use multiple values for \mathcal{E}_W . We proceed cautiously and use the Diamond-Saez value of 0.25 as an upper bound. We also use a more conservative value of $\mathcal{E}_W = 0.1$.

RECOVERING A_W FROM CEO INCOME DATA

The model assumes a continuum of CEOs and firms. In the data there are, of course, a finite number of each. To connect the model to the data, we will treat CEO income (and later firm market size) data as if it is a noisy and incomplete realization of a continuum economy. We will then fit a distribution to the tail of this data and use this to derive tax policy implications for the corresponding (continuum) economy. In doing this we are implicitly assuming that the resulting policy implications are approximately optimal for the (repeated) draws of large finite firm and CEO populations occurring in the US.

In this appendix, we compute A_W using CEO compensation data from the Standard and Poor's ExecuComp database for the year 2011.⁵ The measure of compensation considered includes the amounts received by a CEO (within a fiscal year) from salary, bonus, restricted stock grants and an evaluation of long term incentive pay. This last item is mostly comprised of options. The value of options received as compensation can be calculated either by evaluating at the time they are granted (using the Black-Scholes formula) or by determining the profit obtained at the time the options are exercised. This last approach is used

⁵We have computed estimates of A_W for all years from 1947 to 2011 using data from Frydman and Sacks (2010) and from ExecuComp. The estimates display fluctuations over time however the (time series average) of the Pareto coefficient is 2.23 well within the 95 per cent confidence intervals for the 2011 value. Details are available on request.

by the IRS to determine the taxable amount and our benchmark results follow suite.⁶

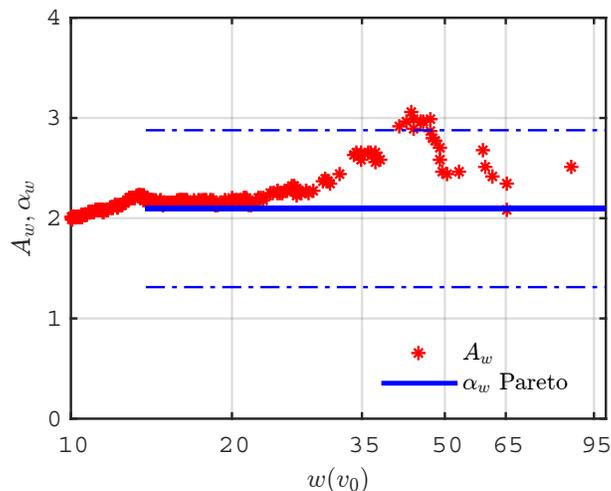


Figure VII.1. : A_W denotes the tail coefficient for the income distribution. Figure VII.1 displays the level of the Pareto coefficient for income α_w above the threshold value of 13.8 millions. 95 per cent confidence intervals for α_w are also plotted.

Recall that $A_W = \frac{W}{\Delta W}$ and if the right tail of the CEO income distribution is Pareto above a threshold income \bar{w} , then A_W is constant and equal to the (constant) Pareto coefficient α_w of the distribution above this threshold. Non-parametric calculations of A_W indicate that it is indeed quite stable above an income of \$10 million or so in our data (see Figure VII.1). Thus, we fit a Pareto distribution to the right tail of the CEO income distribution. We use a two step procedure of Clauset, Shalizi, and Newman (2009). This entails first estimating α_w by maximum likelihood at each fixed \bar{w} and then selecting the \bar{w} value (and corresponding α_w estimate) that maximizes a Kolmogorov-Smirnov goodness-of-fit statistic. An estimate for α_w of 2.1 with a 95 per cent confidence interval equal to (1.13 3.06) (in 2011 data) is obtained. The threshold \bar{w} is estimated to be 13.8 million dollars. We thus compute a top optimal marginal tax rate for CEOs in a continuum economy in which $A_W = 2.1$.

The estimated value of 2.1 for α_w is inline with the numbers typically used to describe the right tail of the income distribution in the taxation literature. Saez

⁶We have also computed estimates of A_W using the former measures of CEO income. They lead to slightly higher estimates of A_W . Details are available on request.

(2001) reports an estimated value of α_w equal to 2, while Diamond and Saez (2011) and Piketty, Saez, and Stantcheva (2014) assume a slightly lower value of 1.5. This suggests that the right tail of the US CEO income distribution is not very different from the tail of the general population's income distribution.⁷

COMPUTED OPTIMAL TAX

Together a value of A_W equal to 2.1 and an elasticity of taxable income \mathcal{E}_W equal to 0.25 imply an optimal tax rate on top CEO incomes (above the threshold \bar{w} of about \$14 million) of approximately 56 per cent. The more conservative taxable income elasticity $\mathcal{E}_W = 0.1$ together with $A_W = 2.1$ implies a marginal income tax on incomes above the threshold of close to 76 per cent. The latter combination of low elasticity and somewhat higher Pareto coefficient generates a top tax rate in line with those reported by Diamond and Saez (2011). Table VII.1 reports optimal marginal tax rates for different income thresholds w_0 based on these estimated coefficients, the coefficients defining the confidence interval and the empirical A_W values displayed in Figure VII.1 for \mathcal{E}_W equal to 0.1 and 0.25.

Table VII.1—: Marginal Tax Rates (per cent): Affine Tax Rates

Tail Properties	Elasticity	w_0 (in millions)		
	\mathcal{E}_W	30	40	60
(I) Based on Pareto Distribution	$\frac{1}{4}$		65.6 [57.3 76.8]	
	$\frac{1}{10}$		82.7 [77.0 89.2]	
(II) Based on A_W	$\frac{1}{4}$	62.6	65.0	61.9
	$\frac{1}{10}$	80.7	82.6	80.2

Notes: w_0 measured in millions of 2011 USD. Row (I) in the table displays rates based on a Paretian right tail of the income distribution. The rate is constant above a threshold value of 13.8 million dollars. Row (II) displays results based on the empirical tail coefficients of the income distribution.

A. Comparisons of the CEO Tail Coefficient to that in the General Population

We compare our estimate of A_W from the population of CEOs to estimates obtained from the entire population. Saez (2001) plots $\frac{1}{w_0} \int_{w_0}^{\infty} wm(w)dw$ (p. 211) and observes that its value between incomes of \$100 thousand and about \$30 million in 1993 in current dollars is about 2. Hence, he infers that A_W is about 2. Alvaredo, Atkinson, Piketty and Saez (2013), using the World Top Incomes Database (WTID), report a value of 1.6 for A_W in the US in 2013. Diamond and

⁷In Section VII.A we estimate the Pareto coefficient for earned (labor) income amongst the general population. Our estimates for the period since 1990 are in the neighborhood of 2.

Saez (2011) and Piketty, Saez, and Stantcheva (2014) both use values of 1.5 in their analyses. These values are a little below the value of 2.1 for A_W that we find in our sample of CEOs. However, it should be noted that estimates of A_W for the general population refer to the distribution of all income irrespective of source. For example the definition of income in the WTID includes not only *wages, salaries and pensions* (which is the quantity of interest in the optimal tax analysis of this paper) but also: *entrepreneurial income, dividends, interest income and rents*. In addition these additional categories are of progressively more important for high income quantiles.

In the direction of correcting the WTID estimates for source, we use the data available in the WTID to obtain the tail parameter for earned income (*wages, salaries and pensions*). This data is, however, aggregated. We use the following strategy to purge non-labor income. Suppose total income y is the sum of earned income w and other sources of income z . We assume that w is distributed at the top according to a Pareto distribution with unknown tail parameter α_w . In addition we assume that there exist a strictly monotone relationship between w and y , so that ordering individuals by w or y will yield the same ranking. The WTID reports by percentile threshold (the thresholds are: 90th, 95th, 99th, 99.5th, 99.9th and 99.99th) both the fraction of total income due to earned income and the conditional average total income for that income group. We assume that for all individuals in a given income group i , earned income is related to total income according to: $w = \rho_i \cdot y$. Given information on the average y within an income group i we can then recover the conditional average for earned compensation \bar{w}_i within the group. If the income distribution has a Pareto tail, it follows that the threshold earned income value for group i , \underline{w}_i , is related to \bar{w}_i according to:

$$\bar{w}_i = \frac{\alpha_w}{\alpha_w - 1} \underline{w}_i.$$

The Pareto assumption further implies:

$$\underline{w}_i = \frac{\bar{w}}{(1 - P_i)^{\frac{1}{\alpha_w}}},$$

where P_i is the fraction of agents below \underline{w}_i . Then considering two percentile categories and simplifying we obtain an estimate for α_w equal to:

$$(VII.2) \quad \alpha_w = \frac{\log\left(\frac{1-P_j}{1-P_i}\right)}{\log\left(\frac{\bar{w}_i}{\bar{w}_j}\right)}.$$

In Figure VII.2 we plot our estimates of α_w from the WTID. We also plot our estimates of the tail parameter using CEO compensation data (in blue) and the Pareto tail parameter reported for the entire WTID (in red). As noted earlier

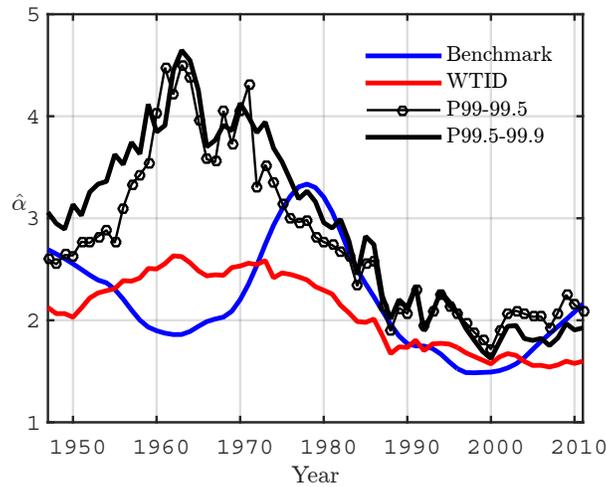


Figure VII.2. : Estimates of α_w . Benchmark refers to the estimates from CEO data using options exercised. WTID from Alvaredo, Atkinson, Piketty and Saez (2013). P99-99.5 and P99.5-99.9 computed using equation (VII.2). Compensation values from 1947 to 1991 are from Frydman and Sacks (2010). From 1992 onwards from ExecuComp.

the coefficient using CEO data displays a more compact distribution than that obtained using the entire income distribution as reported in the WTID. However as we control for the sources of income focusing on earned income the difference becomes smaller. This is particularly evident in the earlier part of the sample. In terms of historical patterns all three approaches display a stretching out of the distribution starting from the 1970s through to 2000.

VIII. Optimal Tax Allocations

This appendix describes the allocation associated with the optimal tax function at a benchmark parametrization. It provides additional information on the incidence of taxation at the optimum and undertakes counterfactual exercises that explore the role of effort and talent in generating dispersion in CEO incomes.

A. Preliminaries

We specialize the optimal control problem described in the body of the paper to the case of quasi-linear/constant elasticity CEO preferences and a multiplica-

tive firm objective. Specifically, assume:

$$V(S(v), z(v)) := DS(v)z(v)$$

and

$$\Phi(v) := U\left(c(v), \frac{z(v)}{h(v)}\right) = c(v) - \frac{1}{1 + \frac{1}{\varepsilon}} \left(\frac{z(v)}{h(v)}\right)^{1 + \frac{1}{\varepsilon}},$$

so that

$$(VIII.1) \quad C[\Phi(v), z(v)/h(v)] = \Phi(v) + \frac{1}{1 + \frac{1}{\varepsilon}} \left(\frac{z(v)}{h(v)}\right)^{1 + \frac{1}{\varepsilon}}.$$

The optimal control problem is then:⁸

$$(VIII.2) \quad \sup_z \int_0^1 \left\{ \chi DS(v)z(v) + (1 - \chi)w(v) - \Phi(v) - \frac{1}{1 + \frac{1}{\varepsilon}} \left(\frac{z(v)}{h(v)}\right)^{1 + \frac{1}{\varepsilon}} \right\} dv$$

subject to:

$$(VIII.3) \quad \Phi_v(v) = \left(\frac{z(v)}{h(v)}\right)^{1 + \frac{1}{\varepsilon}} \frac{h_v(v)}{h(v)}$$

$$(VIII.4) \quad w_v(v) = DS(v)z_v$$

$$(VIII.5) \quad \Phi(1) \geq \bar{U} \quad \text{and} \quad DS(1)z(1) \geq w(1).$$

This is a standard optimal control problem with three state variables (Φ, w, z) and one control z_v .

SOLVING THE OPTIMAL CONTROL PROBLEM

We solve a scaled version of the optimal control problem in (VIII.2) using the numerical solver DIDO version 7.3.7.⁹ The bulk of the parameter values utilized are those presented in the paper. In addition, we set $\chi = 0.8$ and $\bar{U} = 2$. This value of \bar{U} is low enough that all firms and CEOs are active. Alternative values of \bar{U} that ensure the full set of firms and CEOs are active affect the level of CEO consumption, but not the optimal effort allocation or optimal marginal taxes.

B. The Optimal CEO Effort Allocation

VARIATION IN EFFORT AT THE OPTIMUM

The optimal effort allocation is displayed in Figure VIII.1. Effort is decreasing

⁸Under the parameter values we consider $\bar{v} = 1$.

⁹For details on the solution algorithm, refer to Ross and Fahroo (2003).

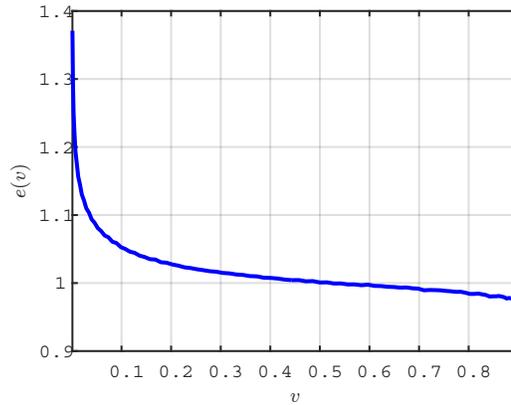


Figure VIII.1. : Effort.

and convex in v up until the median CEO at which point the profile becomes concave. Overall the top CEO's effort is about 40 per cent more than that of the bottom. Care must be taken in interpreting these numbers since they are impacted by our functional form choices. However, the effort variation we find is of a similar order of magnitude to the hours variation amongst top CEOs reported by Bandiera, Prat, and Sadun (2011).¹⁰ Next we document the impact of optimal effort and effective labor variation on firm surplus and on CEO compensating differentials measured in dollars.

IMPACT OF CEO EFFORT AND TALENT ON FIRM SURPLUS

We next perform two counterfactuals that highlight the roles of CEO effort and talent in generating firm output at the optimum (under our parameterization). First, we fix the effort of all CEOs above the median at the level of the median CEO's effort and recompute firm surplus. Second, we fix the talent of all CEOs above the median at that of the median, but keep their efforts fixed at their levels under the optimal allocation. Results are shown in Figure VIII.2 as a fraction of optimal surplus. For the top firm, having a CEO exerting the same amount of effort as the median CEO (a CEO working in a company roughly 100 times smaller) causes a reduction in surplus of approximately 27 per cent. For the top firm, having a CEO with the same talent as the median (but exerting the higher effort of a highly talented CEO at the optimum), leads to a reduction in surplus of 17 per cent. Overall holding effort fixed at the median level has

¹⁰These authors report that when CEOs are ranked by hours worked, then those at the 90th percentile work on average 20 hours (or roughly 50 per cent more) than those at the 10th percentile. They also document a positive relationship between recorded CEO effort and firm labor productivity.

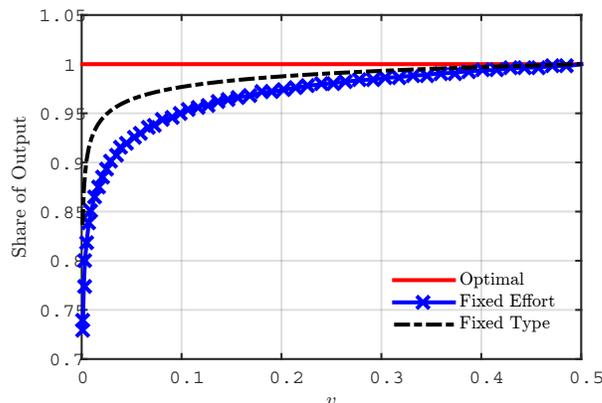


Figure VIII.2. : The Impact of effort and talent on firm surplus.

slightly larger effect than holding talent fixed at this level. Reducing the effort elasticity \mathcal{E} further below $1/15$ brings the effect of holding talent fixed closer to that from holding effort fixed and eventually reverses the relative size of these effects.

COMPENSATING DIFFERENTIALS

Define the compensating differential:

$$\Delta V(v) := -v \left(\frac{z(0.5)}{h(0.5)} \right) + v \left(\frac{z(v)}{h(v)} \right).$$

$\Delta V(v)$ gives the extra consumption needed to make the median CEO indifferent between her equilibrium allocation and an alternative allocation in which she supplies the effective labor of the v -th CEO. Since $\Delta V(v)$ is net-of-tax, it understates the increase in gross pay firm v would need to pay to induce the median CEO to supply CEO v 's effective labor.¹¹ Figure VIII.3 plots $\frac{\Delta V(v)}{w(v)}$, i.e. the compensating differential normalized by the equilibrium income of the v -th CEO $w(v)$. For the median CEO to be indifferent between her equilibrium allocation and that of an allocation in which supplies the effective labor of a top ranked CEO, her (after-tax) income would have to increase by an amount equal to 7 times that of the v -th CEO's gross pay. Thus, it is extremely costly to motivate the median CEO sufficiently that she replicates the top CEO's performance.

¹¹The CEO's gross pay would need to increase by $\Delta w(v) = C^{-1}[\Delta V(v) + c(0.5)] - w(0.5)$, where $C(w) := w - T[w]$ and $c(0.5)$ is the equilibrium consumption of the median CEO.

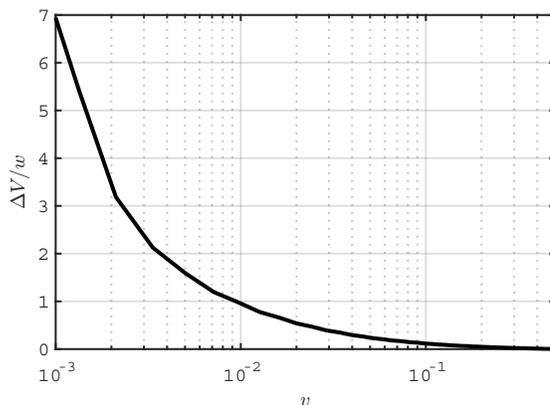


Figure VIII.3. : Compensating Differentials.

TAX COUNTERFACTUAL AND BURDEN

We next analyze the incidence of taxation upon CEOs and firms. The optimal tax system is of the form:

$$T[w] = T[w_0] + \int_{w_0}^w T_w[w']dw'$$

We consider a counterfactual in which T_w is set to zero at all w' .¹² We compute the corresponding equilibrium allocation and, hence, calculate the consumption gains for CEOs and profit gains for firms relative to the optimal allocation in which positive marginal taxes are used to collect additional tax revenue. For CEOs we then compute the proportional consumption gain from switching to the zero marginal tax equilibrium. We display these changes for different CEOs in the distribution in Figure VIII.4a. By definition the welfare impact of such a switch is zero for the CEO at the bottom since competition maintains her utility at \bar{U} . More talented CEOs work harder in response to the reduction in marginal tax rates and capture some of the increase in surplus in the form of additional utility (see the envelope condition (VIII.3)). These gains amount to a 3 per cent of consumption gain for the most talented CEOs. As noted throughout the paper, lower marginal tax rates on CEOs raise firm profits. Figure VIII.4b shows that

¹²The lowest ranked active firm must pay its CEO enough that her after-tax income and, hence, consumption (and effective labor) is sufficient to attain her reservation utility \bar{U} . The lump sum tax $T[w_0]$ imposed on the lowest ranked active CEO is used by the policymaker to extract all profit from the smallest active firm. The ability to impose this tax has little effect on the CEO's utility, since competition always forces the lowest ranked CEO's utility to \bar{U} . However, it has a large proportional effect on the lowest ranked firm's profit. In this counterfactual we focus on the tax burden associated with non-negative marginal taxes.

such increases are proportionally greater in smaller, lower profit firms, whose lower income CEOs face higher marginal taxes at the optimum.

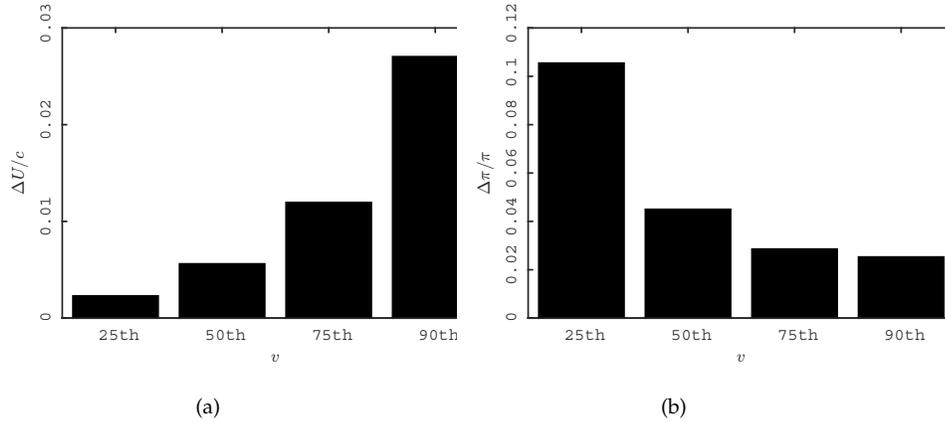


Figure VIII.4. : Tax Burden Calculations.

APPENDIX REFERENCES

- Alvaredo, F., A. Atkinson, T. Piketty, and E. Saez (2013). The Top 1 per cent in International and Historical Perspective. *Journal of Economic Perspectives* 27(3), 3–20.
- Bakija, J., A. Cole, and T. Heim. (2012). Jobs and income growth of top earners and the causes of changing income inequality: Evidence from US tax return data. Working Paper, Williams College.
- Bandiera, O., A. Prat, and R. Sadun. (2013). Managing firms in an emerging economy: Evidence from the time use of Indian CEOs. Working Paper.
- Clauset, A., C. Shalizi, and M. Newman (2009). Power law distributions in empirical data. *SIAM Review* 51(4), 661–703.
- Diamond, P. and E. Saez (2011). The Case for a Progressive Tax: From Basic Research to Policy Recommendations. *Journal of Economic Perspectives* 25(4), 165–190.
- Frydman, C. and R. Malloy. (2011). Does tax policy affect executive compensation? Evidence from post war tax reforms. *Journal of Public Economics* 95(11-12), 1425–1437.
- Frydman, C. and R. Sacks. (2010). Executive compensation: A new view from a long term perspective, 1936-2005. *Review of Financial Studies* 23(5), 2099–2138.

- Gabaix, X. and A. Landier. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics* 123(5), 49–100.
- Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. *Journal of Political Economy* 108(2), 352–378.
- Klenke, A. (2008). *Probability theory: A comprehensive course*. Springer-Verlag.
- Piketty, T., E. Saez and S. Stantcheva. (2014). Optimal Taxation of Top Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy* 6(1), 230–271.
- Ross, I. and F. Fahroo. (2003). Legendre pseudispectral approximations of optimal control problems. *New Trends in Nonlinear Dynamics and Control and their Applications*, p.p. 327–342. Springer.
- Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies* 68(1), 205–229.
- Terviö, M. (2008). The difference that CEOs make: An assignment approach. *American Economic Review* 98(3), 642–668.
- Scheuer, F. and I. Werning. (2015) The Taxation of Superstars. NBER Working Paper No. 21323.