

Taxing Top CEO Incomes[†]

By LAURENCE ALES AND CHRISTOPHER SLEET*

We use a firm-CEO assignment framework to model the market for CEO effective labor. In the model's equilibrium, more talented CEOs match with and supply more effort to larger firms. Taxation of CEO incomes affects the equilibrium pricing of CEO effective labor and, hence, spills over and affects firm profits. Absent the ability to tax profits or a direct concern for firm owners, a standard prescription for high marginal income taxes emerges. However, given such an ability or concern, the optimal marginal tax rates are much lower. (JEL D31, H21, M12)

What should the marginal tax rate on top income earners be? Recent research suggests that it should be high, perhaps as high as 70 percent or 80 percent. Underpinning these numbers is the well-known Diamond-Saez formula that relates the optimal marginal tax rate on top incomes to the elasticity of taxable income and a property of the right tail of the earnings distribution. This formula is derived under the assumption that the policymaker's objective is to maximize tax revenues derived from top earners. It abstracts from any positive impact of the efforts of these earners on the incomes of other agents or on tax revenues collected from other sources.¹ Our paper departs from this research by taking seriously the idea that the activities of high-earning CEOs, an important group of top earners, have positive spillovers for others.² We use a firm-CEO assignment framework to model the market for CEO effective labor. Gabaix and Landier (2008) and Terviö (2008) have shown that such a framework is valuable for understanding recent growth in CEO incomes and the interaction of firm and CEO attributes in shaping this growth. We show that in an assignment model (augmented with an intensive CEO effort margin), the taxation

*Ales: Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 (e-mail: ales@cmu.edu); Sleet: Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 (e-mail: csleet@andrew.cmu.edu). We thank Caroline Hoxby and Florian Scheuer and many seminar participants for helpful comments and discussion. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <http://dx.doi.org/10.1257/aer.20151093> to visit the article page for additional materials and author disclosure statement(s).

¹The literature has considered negative impacts such as rent seeking and has modified the basic Diamond-Saez formula accordingly: see Piketty, Saez, and Stantcheva (2014). Saez (2001) also modifies the formula to allow for social concern for top income earners.

²Bakija, Cole, and Heim (2012) report that 60 percent of the top 0.1 percent of earners by income are executives, managers, supervisors, and financial professionals. Our focus on CEOs is also partially motivated by the availability of high-quality uncensored data on the incomes of CEOs. However, we believe that our theoretical and broad quantitative insights carry over to other "superstar" buyer-seller relationships that generate high incomes for sellers. In particular, they are applicable to hedge fund managers and other financial professionals who make complex decisions.

of CEO incomes affects the equilibrium pricing of CEO effective labor and, hence, spills over and affects firm profits.³ If the policymaker has no social concern for profits, then the classic Diamond-Saez formula remains intact. Otherwise, the optimal marginal tax rate on CEO incomes is modified downward. At our benchmark parameterization, a full reform of CEO income and profit taxation entails an optimal marginal tax on top CEO incomes of about 15 percent.

The basic CEO to firm assignment model supposes one-to-one matching of differently talented CEOs to differently sized firms. As noted, we augment this model with an intensive CEO effort margin (and with taxes). The equilibrium features assortative matching of CEO talent with firm size. More talented CEOs match with and supply more effort (and more effective labor) to larger firms. The indivisibility of the CEO position prevents combinations of less talented CEOs replacing more talented ones and equalizing the price for effective CEO labor across firms. On the other hand, competition among similarly talented CEOs for a position prevents any given CEO from extracting all of the surplus from a firm. In equilibrium the price of a unit of CEO effective labor equals the marginal product of CEO effective labor at the firm at which this unit is the last hired. Since the marginal product of CEO effective labor is increasing in firm size, the matching of more talented CEOs to larger firms enhances the dispersion of top CEO incomes. Even if there is relatively little dispersion in CEO talent, large variations in firm size can translate into large variations in top CEO incomes. However, since a CEO is only paid the marginal product of her effective labor on the last unit she sells to her firm (with inframarginal units priced by and paid their marginal product at smaller firms), claimants to firm profits capture some surplus. In this setting, an increase in the marginal tax rate above a threshold income induces an upward adjustment in the pricing schedule for effective labor. This in turn redistributes from firm profits to CEO incomes and, hence, CEO income tax revenues. If the policymaker is concerned only with maximizing income tax revenues, then this redistribution provides a motive for higher marginal income taxes. If, on the other hand, the policymaker is indifferent to the allocation between income tax revenues and firm profits (because the latter can be taxed, because tax receipts and firm claimants are equally valued, or because depressing firm profits has adverse effects on firm creation), then no such redistribution motive for higher marginal taxes exists.

We first consider the optimal linear tax rate across a range of top CEO incomes. The classic Diamond-Saez tax formula relates this tax rate to the elasticity of taxable income and a right tail property of the income distribution. Since this formula assumes the policymaker maximizes revenues from the income taxation of top earners and is silent on how these earners generate income, it applies in our setting, but only if the policymaker attaches no social value to firm profits. An alternative Mirrleesian formula relates the optimal tax rate to the elasticity of worker effort and a right tail property of the worker talent distribution. Absent income effects on CEO effort, it holds if the policymaker places equal weight on CEO income tax revenues and firm profits (for example, because these can be taxed at 100 percent) and, thus,

³As we elaborate, the nature of this spillover is subtle. There is no *marginal mispricing*. Instead, the inframarginal rents accruing to a given firm are impacted by the effort of CEOs at other, smaller firms and, hence, by the taxes placed on these CEOs.

maximizes surplus not paid to CEOs.⁴ Since in our baseline model firm profits are pure rents, 100 percent taxation of profits is optimal. Thus, the Mirrleesian formula is relevant if a comprehensive reform of profit and CEO income taxation is implemented. If institutional constraints or other economic frictions restrain the optimal profit tax rate and the social marginal value of profit, then neither the Diamond-Saez nor the Mirrleesian formulas are valid. In contrast, in the standard public finance setting with heterogeneous workers selling effective labor to a competitive firm, all surplus not paid out to high income workers is captured in taxes. Thus, there is no distinction between maximizing this surplus and maximizing income tax revenues and so both formulas hold.

The simpler linear tax setting just described provides intuition for the analysis of optimal nonlinear taxation. Specifically, and analogous to the linear setting, a conventional-looking optimal tax formula in terms of the CEO income elasticity and the local Pareto coefficient of CEO incomes emerges when zero social weight is placed on firm profits. In this case, the elasticity is adjusted to take into account the effect of a higher tax rate at a given income on the equilibrium pricing schedule for effective labor and, hence, the incomes received by more talented CEOs earning larger amounts at bigger firms.⁵ When the policymaker weights CEO income tax revenues and firm profits equally, then a conventional Mirrleesian optimal tax formula in terms of the CEO's effort elasticity and the Pareto coefficient of CEO talent arises. We use (nonlinear) optimal tax formulas expressed in terms of the underlying (structural) talent and firm size asset distributions to quantitatively characterize optimal taxes on CEOs across a range of high incomes and firm profit weights. To that end we extend an empirical strategy of Terviö (2008) to allow for an intensive effort margin. If a comprehensive reform of income and firm profit taxation is implemented, then income taxes and firm profits are equally weighted and optimal marginal tax rates decline from around 20 percent at an income of \$10 million to about 10 percent at an income of \$100 million. If a partial reform of CEO income taxation occurs holding profit taxes close to their statutory values in the United States of about 60 percent and abstracting from direct concern for profit recipients or impact of profits on firm entry, then optimal tax rates decline from 34 percent to 27 percent over a similar income range. In either case, they are very different from the rates of 70 percent to 80 percent recently advocated in the literature.

The remainder of the paper proceeds as follows. Following a brief literature review, Section I provides our baseline assignment model of CEO incomes and firm profits. It gives an initial characterization and formulation of equilibrium suitable for tax analysis. Section II analyzes the optimal linear tax rate across a range of CEO top incomes. Section III considers optimal nonlinear taxation. It provides formulas that characterize the fully optimal nonlinear tax. Section IV uses these formulas and data on CEO compensation and firm values to calculate the optimal nonlinear tax function for CEOs over a range of incomes and weights on profits. Section V concludes. Online appendices contain proofs and additional details.

⁴It can also be interpreted as adjusting the Diamond-Saez formula to purge out adjustments in CEO income in response to tax rises that come at the expense of firm profits.

⁵Analogous to before, a higher marginal tax rate at a given income causes the incomes of more talented CEOs to rise at the expense of firm profits.

Related Literature.—Our paper contributes to a literature in normative public finance that considers the income tax implications of fiscal spillovers.⁶ Stiglitz (1982) analyzes optimal income taxation with endogenously determined wages. In this framework, diminishing returns with respect to a given skill type's labor input and imperfect substitutability between skill types implies that the wage distribution is endogenous to tax policy. Rothschild and Scheuer (2013) extend this analysis to a rich worker-occupation assignment setting. Ales, Kurnaz, and Sleet (2015) explore the policy implications of technical change in such a setting. In these models, many workers match with an occupation and spillovers operate through the collective effect of worker labor supply and occupation choices on occupational output prices.⁷ Rothschild and Scheuer (2014) extend the framework of these papers by divorcing the private return from an occupation from the social return and, hence, incorporating explicit externalities at the level of occupations. Thus, they introduce a motive for corrective Pigouvian taxation.⁸ In a different direction, Stantcheva (2014) provides a rich model that combines informational frictions between workers and firms as well as between workers and the government. Taxes perturb the menu of contracts offered by firms and, hence, induce redistributions among different worker types (rather than between CEOs and firm owners). None of the preceding papers focuses on top earners or CEOs per se. While we abstract from the externalities assumed in Rothschild and Scheuer (2014) and the private information between firms and workers (CEOs) assumed in Stantcheva (2014), we enrich the analysis with heterogeneity on the side of firms and an assignment structure.

Recent work on the taxation of top earners has emphasized positive fiscal spillovers from the taxation of top earners. In the context of high income earners, Saez, Slemrod, and Giertz (2012) extend the basic Diamond and Saez (2011) formula to accommodate income shifting from the personal to the corporate tax base. Piketty, Saez, and Stantcheva (2014) further extend the formula to incorporate rent seeking by managers. In their setting, managers exert costly effort in bargaining, which to the extent that it allows them to extract resources above their marginal product imposes a negative externality on others and is discouraged by higher taxation. Their formula for the optimal marginal tax rate above a threshold income is similar to the one that we derive in Section II. However, the economics behind these formulas and their connections to the data are quite different. We abstract from rent-seeking, but instead stress the positive role of CEOs in creating firm value which is at center of an important literature on CEO compensation. In a celebrated result, Diamond and Mirrlees (1971) established that Ramsey tax formulas hold (and intermediate goods taxation is undesirable) in a general production setting provided the government has access to a rich enough set of linear commodity taxes and profits are taxed at 100 percent. This echoes our result that the Mirrlees optimal tax formulas hold when profits are taxed at 100 percent. Subsequent contributions by Stiglitz and Dasgupta

⁶We use the term *fiscal spillover* broadly to describe situations in which the taxation of one population of agents or source of income spills over and affects other populations or income sources. These spillovers encompass classical general equilibrium tax incidence effects, externalities, and income shifting across tax bases.

⁷Specifically, workers are perfectly substitutable within an occupation. Imperfect substitutability between workers stems from imperfect substitutability between occupational outputs and the comparative advantage of differently talented workers for different occupations which impedes occupational mobility.

⁸See also Rothschild and Scheuer (2016) and Lockwood, Nathanson, and Weyl (forthcoming).

(1971) and Munk (1978, 1980) explored the consequences of restrictions on profit taxes for optimal commodity taxation. In these papers, Ramsey optimal tax formulas are modified by a concern for the impact of commodity and factor taxes on the revenues collected via the taxation of profits. This parallels our result that marginal taxes on CEO incomes are decreasing in the social weight placed on profits.

By far the most closely related paper to ours is the contemporaneous work of Scheuer and Werning (2015). These authors develop the implications for optimal taxation of an assignment setting similar to ours. They focus on the case in which firm profits receive the same effective social weight as income tax revenues collected from CEOs. For this case they derive formulas that permit the testing of an income tax code for Pareto optimality (with respect to different combinations of weights across CEOs). These formulas are closely related to the ones we provide. In particular, they show that in this case, the Mirrleesian test condition for Pareto optimality expressed in terms of the effort elasticity and the talent distribution is the same as that obtained in the conventional setting without assignment. Moreover, they show that the formula when expressed in terms of the income distribution and CEO elasticities of income is also the same provided the latter elasticities are defined appropriately (as microeconomic elasticities that hold the pricing schedule for CEO effective labor fixed). This is an important observation that potentially leads to an alternative empirical implementation of this test formula. Collectively, Scheuer and Werning (2015) cast these results as neutrality propositions. However, as Scheuer and Werning (2015) point out, while the test formulas themselves are neutral to assignment considerations, the way these formulas are brought to the data and their potential quantitative implications for tax design are not. We focus on the calculation of optimal taxes under different weightings of income taxes and firm profits. As noted above, we obtain optimal tax rates very much lower than those obtained in conventional analyses.

Ales, Bellofatto, and Wang (2015) consider optimal taxation in a Rosen (1982) span of control setting. They identify top earners with managers who match with and control teams of workers. In contrast to our setting, firms have no exogenously given factors and all variations in firm size are attributable to variations in managerial talent. Managerial productivity is enhanced by firm (team) size and this creates a novel incentive for the government to tax firm size and, hence, shape the equilibrium managerial wage distribution. When considering managerial taxation, our paper departs from Ales, Bellofatto, and Wang (2015) and follows Terviö (2008, p. 643) in assuming that “firms are differentiated by important indivisible characteristics that cannot be easily shuffled among(st them).”

Two recent and highly influential papers by Gabaix and Landier (2008) and Terviö (2008) use a competitive assignment framework to understand the determination of top CEO incomes. In this framework, CEO talent and a firm’s (indivisible and nontransferable) assets are complementary and there is assortative matching of CEOs and firms.⁹ Both Gabaix and Landier (2008) and Terviö (2008) emphasize the role of variations in the size of a firm’s assets in the determination of top CEO incomes with the former attributing the rise in these incomes to increases in firm

⁹Terviö (2008) in particular elaborates on the nature of these assets.

size. Our paper augments the sort of competitive assignment models considered by Gabaix and Landier (2008) and Terviö (2008) with an intensive effort margin and income taxation. Consistent with these contributions we find evidence of a thin right tail to the talent distribution. A separate literature looks at the CEO-firm relationship as a moral hazard problem. This literature is very large: see Gayle and Miller (2009) and Gayle, Golan, and Miller (2015) for recent empirical contributions and Edmans and Gabaix (2009), and Frydman and Jenter (2010) for surveys. It shares with Gabaix and Landier (2008) and Terviö (2008) the idea that CEOs create value for firm owners. In contrast to Gabaix and Landier (2008) and Terviö (2008), it focuses instead on the importance of motivating CEOs to exert effort through an appropriately structured compensation package. Our paper also emphasizes the importance of CEO effort in creating value, but unlike the moral hazard literature it abstracts from the structure of CEO pay and focuses upon the level of pay and the role of taxes in shaping equilibrium outcomes.

I. Competitive Assignment of CEOs and Firms

We augment an assignment game of CEOs and firms with CEO effort and taxes on CEO incomes. While our focus is upon the taxation of CEOs, our analysis applies more broadly to the taxation of (high income) sellers and buyers in an assignment setting.

CEOs and Firms.—The population of CEOs is described by a Lebesgue measure on the interval $I = (0, 1]$. Let $h : I \rightarrow \mathbb{R}_+$ give the talent of each CEO with h a smooth and decreasing function.¹⁰ Thus, $v \in I$ provides a ranking of CEOs by talent and the inverse of h is the countercumulative distribution of talent. Let F denote the distribution function of talent and f its density so that $F = 1 - h^{-1}$ and $f = -1/h_v$, where h_v is the derivative of h . The amount of effective labor z^s supplied by a CEO is a multiplicative combination of talent and effort:

$$(1) \quad z^s = h(v)e.$$

We assume that a CEO can sell labor to only one firm in a period.

The utility of a CEO over consumption c and effort e is given by $U : \mathbb{R}_+ \times [0, \bar{e}] \rightarrow \mathbb{R}$, with U strictly concave, twice continuously differentiable on the interior of its domain, strictly increasing in c , and strictly decreasing in e . U is assumed to satisfy the Spence-Mirrlees single crossing property, that is $-\frac{U_e(c, z^s/h)}{hU_c(c, z^s/h)}$ is assumed to be decreasing in h at each (c, z^s, h) . A CEO must pay a tax $T : \mathbb{R}_+ \rightarrow \mathbb{R}$, $T(w) \in (-\infty, w]$, on her earnings. Consequently, if the v th ranked CEO supplies effective labor z^s and earns income w , her after-tax income and, hence, consumption is $c(w) = w - T[w]$ and her utility is

$$(2) \quad U\left(w - T[w], \frac{z^s}{h(v)}\right).$$

¹⁰To allow for unbounded CEO talent, $I = (0, 1]$ is assumed open at 0. All of our results continue to hold if I is set equal to $[0, 1]$.

Finally, we assume that all CEOs have an outside utility option of $\bar{U} > U(0, 0)$.¹¹

A population of firms is also described by a Lebesgue measure on the interval I . Firms are differentiated by the size of their productive, nontransferable, and indivisible assets. These assets could be intangibles such as reputation or goodwill that are difficult to trade, they could be firm-specific intellectual property or they could capture industry-specific aspects of technology that shape the scale of the firm’s operations.¹² Let $S : I \rightarrow \mathbb{R}_+$ give the size of, i.e., the quantity of assets at each firm, with S a smooth and decreasing function. In the context of firms, $v \in I$ provides a ranking by (asset) size and the inverse of S is the countercumulative distribution of size. Let G denote the distribution function of firm size and g its density so that $G = 1 - S^{-1}$ and $g = -1/S_v$, where S_v is the derivative of S . If the v th firm purchases z^b units of effective labor from a CEO and pays the CEO w , then firm claimants (i.e., the owners of $S(v)$) earn profits of

$$(3) \quad V(S(v), z^b) - w,$$

where the surplus function $V : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is assumed to be super-modular, increasing in both arguments and continuously differentiable and concave in z^b . Note that the surplus value $V(S(v), z^b)$ is net of payments to other adjustable inputs. It can be obtained from a richer problem in which the firm buys (a vector of) additional adjustable inputs x at prices p :

$$(4) \quad V(S(v), z^b) := \sup_x W(S(v), z^b, x) - p \cdot x.$$

The input vector x could include adjustable capital and we explicitly extend the model in such a direction in our later quantitative section.^{13, 14}

The Market Assignment Game with Taxes.—Given T and \bar{U} , CEOs and firms play an assignment game. As a precursor to later optimal tax results, we formalize this game and characterize its equilibrium. The analysis is complicated relative to that in Terviö (2008) and Gabaix and Landier (2008) by the inclusion of the intensive effort margin on the side of CEOs and taxes.¹⁵

¹¹The assumption of a common outside utility option \bar{U} ensures that only the least talented CEOs outside option binds in equilibrium. It is stronger than needed. Provided the outside option of CEOs does not increase too strongly with talent, the behavioral response of CEOs to the tax perturbations we consider will continue to be on the CEO effort rather than the career margin and all of our tax analysis goes through.

¹²See Terviö (2008) and the references therein for additional discussion.

¹³Despite the exclusion of adjustable capital, we refer to $S(v)$ as firm v ’s *assets* adding the qualifiers *nontransferable* or *immovable* when needed.

¹⁴The formulation (4) treats the factor prices p as exogenous to CEO behavior. If the adjustable inputs x complemented CEO effective labor and were not in perfectly elastic supply, then increases in CEO effort could raise demand for these inputs and, hence, (equilibrium) prices p . In this way tax policy that deterred CEO effort could have adverse effects beyond those emphasized in this paper.

¹⁵A tractable and important special case assumes quasilinear-constant elasticity CEO preferences $U(c, e) = c - \frac{\varepsilon}{1 + \varepsilon} e^{\frac{1 + \varepsilon}{\varepsilon}}$ and multiplicative firm payoffs $V(S, z) = DSz$. In this special case, absent taxes, the assignment equilibrium is equivalent to one in which there is no intensive effort margin. The introduction of taxes breaks this equivalence. We thank a referee for emphasizing this point.

Let $w : I \rightarrow \mathbb{R}_+$ give the CEO income paid by each firm, $\mu : I \rightarrow I \cup \{u\}$ the talent rank of each firm's CEO (with $\mu(v) = u$ indicating that v is unmatched), and $z : I \rightarrow \mathbb{R}_+$ the quantity of CEO effective labor purchased by each firm. The functions w , μ , and z are assumed to be Lebesgue measurable. In addition, the match function μ is assumed to be measure-preserving, i.e., for all Lebesgue measurable sets $B \subset \mu^{-1}(I)$, $\mathcal{M}[\mu(B)] = \mathcal{M}[B]$, where \mathcal{M} denotes Lebesgue measure. This captures the one-to-one matching of CEOs and firms.

Equilibria in the Assignment Game.—Definition 1 below defines an equilibrium for the assignment game with taxes. The definition requires that no CEO or firm can improve on its equilibrium allocation by unilaterally leaving the market and that there is no CEO-firm pair whose members can make themselves jointly better off by, if necessary, dissolving their equilibrium matches or leaving their current unmatched states, matching together, and choosing a new income-labor combination.

DEFINITION 1: A triple (μ, z, w) is an equilibrium of the firm-CEO assignment game at (T, \bar{U}) , if:

- (i) (**Participation**) each matched firm-CEO pair $(v, \mu(v))$ is better off at their equilibrium allocation than unmatched, i.e., for all $v \in \mu^{-1}(I)$,

$$(5) \quad V(S(v), z(v)) - w(v) \geq 0 \quad \text{and} \quad U\left(w(v) - T[w(v)], \frac{z(v)}{h(\mu(v))}\right) \geq \bar{U};$$

- (ii) (**Stability**) there is no firm v , CEO v' , and allocation (z', w') such that both firm and CEO weakly prefer (z', w') to their equilibrium allocation and at least one strictly prefers it, i.e., there is no v, v', w' , and z' such that:

- (a) if firm v is unmatched in equilibrium, then it obtains a weakly higher payoff from (z', w') than from being unmatched:

$$(6) \quad V(S(v), z') - w' \geq 0 \quad \text{if} \quad \mu(v) = u,$$

or if firm v is matched in equilibrium (with CEO $\mu(v)$), then it obtains a weakly higher payoff from (z', w') than from allocation $(z(v), w(v))$:

$$(7) \quad V(S(v), z') - w' \geq V(S(v), z(v)) - w(v) \quad \text{if} \quad \mu(v) \in I,$$

and

- (b) if CEO v' is unmatched in equilibrium, then it obtains a weakly higher payoff from (z', w') than from being unmatched:

$$(8) \quad U\left(w' - T[w'], \frac{z'}{h(v')}\right) \geq \bar{U} \quad \text{if} \quad v' \notin \mu(I)$$

or if CEO v' is matched in equilibrium (with firm $\hat{v} = \mu^{-1}(v')$), then it obtains a weakly higher payoff from (z', w') than from allocation $(z(\hat{v}), w(\hat{v}))$:

$$(9) \quad U\left(w' - T[w'], \frac{z'}{h(v')}\right) \geq U\left(w(\hat{v}) - T[w(\hat{v})], \frac{z(\hat{v})}{h(v')}\right) \text{ if } \mu(\hat{v}) = v',$$

with at least one of the applicable inequalities above strict;

(iii) (No rents to the least talented CEO) if $1 \in \mu(I)$, then

$$(10) \quad U\left(w(\mu^{-1}(1)) - T[w(\mu^{-1}(1))], \frac{z(\mu^{-1}(1))}{h(1)}\right) = \bar{U}.$$

Note that an implication of the assumption $\bar{U} > U(0, 0)$ is that any CEO-firm match involves the trade of a positive amount of effective labor for a positive income: there are no passive matches in which nothing is done.¹⁶ The “no rents to the least talented CEO” component of the definition may be justified informally by assuming that the interval I of CEOs are the most talented members of a population of strictly greater measure with outside option \bar{U} and that if the least talented CEO obtained a payoff in excess of \bar{U} , then a slightly less talented unmatched person could supply the same effective labor for an income slightly below $w(\mu^{-1}(1))$ to firm $\mu^{-1}(1)$ and make both this firm and herself better off.

We now give a proposition that characterizes equilibria. It shows that there is assortative matching between CEOs and firms who choose to match and gives simple participation and incentive constraints that must be satisfied in equilibrium. The latter conditions require only that each matched CEO (respectively, firm) is better off accepting its equilibrium allocation than the equilibrium allocation of another matched CEO (firm). In addition, the proposition establishes that these conditions are sufficient for stability if the tax function sufficiently penalizes income-labor allocations outside of the range of the equilibrium allocation functions (z, w) . In particular, in this case they ensure that joint deviations in which a CEO-firm pair dissolve their equilibrium matches, rematch with each other, and select a new income-labor allocation cannot make both parties better off. Thus, the (more complicated) stability conditions on firms and CEOs in Definition 1 are decoupled and re-expressed as simple CEO and firm incentive conditions. This is useful for our subsequent tax analysis.

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

$$(11) \quad U\left(w(v) - T[w(v)], \frac{z(v)}{h(v)}\right) \geq \bar{U} \quad \text{and} \quad V(S(v), z(v)) - w(v) \geq 0,$$

¹⁶CEOs must receive a strictly positive income to match with a firm and give up their outside option. Since firms must earn nonnegative revenue, they must contract for a positive amount of effective labor from a CEO.

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

$$(12) \quad U\left(w(v) - T[w(v)], \frac{z(v)}{h(v)}\right) \geq U\left(w(v') - T[w(v')], \frac{z(v')}{h(v)}\right),$$

and

$$(13) \quad V(S(v), z(v)) - w(v) \geq V(S(v), z(v')) - w(v').$$

On the other hand, given \bar{U} , if T, \tilde{v}, z , and w are such that (i) for all $v \in (0, \tilde{v}]$, (11) to (13) hold; (ii) $U\left(w(\tilde{v}) - T[w(\tilde{v})], \frac{z(\tilde{v})}{h(\tilde{v})}\right) = \bar{U}$ and $V(S(\tilde{v}), z(\tilde{v})) - w(\tilde{v}) \geq 0$; and (iii) for all $w' \notin w((0, \tilde{v}])$, $T[w'] = w'$, then (μ, z, w) with μ such that for all $v \in (\tilde{v}, 1]$, $\mu(v) = u$, and for all $v \in (0, \tilde{v}]$, $\mu(v) = v$, is an equilibrium at (T, \bar{U}) .

PROOF:

See online Appendix.

We also note that, by standard arguments, equilibrium effective labor z and CEO income w are non-increasing on $(0, \tilde{v}]$ as are CEO consumption $c(v) = w(v) - T[w(v)]$ and CEO and firm payoffs

$$\Phi(v) := U\left(w - T[w], \frac{z(v)}{h(v)}\right) \quad \text{and} \quad \pi(v) := V(S(v), z(v)) - w(v).$$

Equilibrium CEO Income Determination.—If z and w are differentiable at $v \in (0, \tilde{v}]$, then (13) implies the first-order condition for firms:

$$(14) \quad V_z(S(v), z(v))z_v(v) - w_v(v) = 0.$$

Integrating out (14) gives

$$(15) \quad w(v) = w(\tilde{v}) + \int_v^{\tilde{v}} V_z(S(v'), z(v'))(-z_v(v')) dv'.$$

It follows that each increment of effective labor $-z_v(v')$ above $z(\tilde{v})$ is *priced* at $V_z(S(v'), z(v'))$, i.e., it is paid its marginal product at the firm at which it is the last unit hired. Thus, CEO v receives her marginal product $V_z(S(v), z(v'))$ only on the last unit she supplies. On inframarginal units she is paid $V_z(S(v'), z(v'))$ with firm v collecting the difference $\{V_z(S(v), z(v')) - V_z(S(v'), z(v'))\}(-z_v(v'))$ as profit. Super-modularity of V implies that the *price of effective labor* V_z is increasing in S . Since more talented firms match with larger firms, this is a source of dispersion of CEO income across talent ranks. Even if there is relatively little dispersion in CEO talent, this force can translate a large variation in firm size into a large variation in

top CEO incomes. If, in addition, T is differentiable at $w(v)$ and w and z are differentiable at $v \in (0, \tilde{v}]$, then (12) implies the CEO's first-order condition:

$$(16) \quad U_c \left(w(v) - T[w(v)], \frac{z(v)}{h(v)} \right) w_v(v) \{1 - T_w[w(v)]\} \\ + U_e \left(w(v) - T[w(v)], \frac{z(v)}{h(v)} \right) \frac{z_v(v)}{h(v)} = 0.$$

Combining (14) and (16), totally differentiating with respect to v and denoting the compensated and uncompensated effort elasticities by \mathcal{E}^c and \mathcal{E}^u gives¹⁷

$$(17) \quad \frac{w_v}{w} = \left(\frac{V_z z}{w} \right) \frac{(1 + \mathcal{E}^u) \left(\frac{h_v}{h} \right) + \mathcal{E}^c \frac{V_z S}{V_z} \left(\frac{S_v}{S} \right)}{1 + \mathcal{E}^c \left\{ \frac{T_{ww} w}{1 - T_w} \frac{V_z z}{w} + \frac{V_{zz} z}{V_z} \right\}}.$$

Expression (17) relates equilibrium CEO income variation $\frac{w_v}{w}$ to talent $\frac{h_v}{h}$ and firm size variation $\frac{S_v}{S}$ (across rank). These last two variables contribute to CEO income variation directly and also indirectly through the incentives for greater effort that they create. In much of the paper we specialize to the case in which firm's objective is multiplicative $V(S, z) = DSz$, with D a parameter. If, in addition, the tax function T is locally linear, then equation (17) reduces to

$$(18) \quad \frac{w_v}{w} = \left(\frac{DSz}{w} \right) \left\{ \frac{h_v}{h} + \mathcal{E}^u \frac{h_v}{h} + \mathcal{E}^c \frac{S_v}{S} \right\}.$$

Thus, CEO income variation is related via the price of effective labor DS to effective labor variation $\frac{z_v}{z} = \frac{h_v}{h} + \mathcal{E}^u \frac{h_v}{h} + \mathcal{E}^c \frac{S_v}{S}$. Equation (18) further decomposes CEO income variation into a part due to variation in the talent rent accruing to CEOs $\left(\frac{DSz}{w} \right) \frac{h_v}{h}$ and a part due to variation in CEO effort $\left(\frac{DSz}{w} \right) \frac{e_v}{e} = \left(\frac{DSz}{w} \right) \times \left(\mathcal{E}^u \frac{h_v}{h} + \mathcal{E}^c \frac{S_v}{S} \right)$.¹⁸ In the assignment models of Terviö (2008) and Gabaix and Landier (2008) there is no intensive effort margin, all CEO income variation is attributable to variation in talent rents and (18) further reduces to $\frac{w_v}{w} = \frac{DSz}{w} \frac{h_v}{h}$. Alternatively, in the standard labor supply model used in optimal taxation there is an

¹⁷ Here the compensated and uncompensated elasticities are given by $\mathcal{E}^c = \frac{U_e/e}{-2 \frac{U_e}{U_c} U_{ce} + \left(\frac{U_e}{U_c} \right)^2 U_{cc} + U_{ee}}$ and $\mathcal{E}^u = \frac{U_e/e - \left(\frac{U_e}{U_c} \right)^2 U_{cc} + \left(\frac{U_e}{U_c} \right) U_{ce}}{-2 \frac{U_e}{U_c} U_{ce} + \left(\frac{U_e}{U_c} \right)^2 U_{cc} + U_{ee}}$ with each U_x and U_{xy} , $x, y \in \{c, e\}$, giving the relevant partial first and second derivatives of U .

¹⁸ Note that the uncompensated elasticity attaches to talent variation and the compensated elasticity to firm asset variation. Local variation in talent modifies the return to effort and CEO consumption; local variation in firm asset size modifies only the return to effort. Firm claimants not the CEO capture additional firm surplus attributable to local variation in S .

intensive effort margin, but workers capture all of the surplus (with no part accruing to owners of a firm asset S). In this case (18) reduces to $\frac{w_v}{w} = \frac{h_v}{h} + \mathcal{E}^u \frac{h_v}{h}$ and CEO income variation is not enhanced by variation in firm size.

Optimal income tax formulas are often expressed as functions of the (local) Pareto coefficients for talent or income. Anticipating these formulas it is useful to re-express (18) in these terms. Recall that F and f denote the talent distribution and density and G and g the firm asset distribution and density. Let M and m denote the distribution and density of CEO incomes. The corresponding local Pareto coefficients for CEO talent, firm size, and CEO income are: $\alpha_h(v) := \frac{h(v)f(h(v))}{1 - F(h(v))}$, $\alpha_S(v) := \frac{S(v)g(S(v))}{1 - G(S(v))}$, and $\alpha_w(v) := \frac{w(v)m(w(v))}{1 - M(w(v))}$. Using $f(h(v)) = -\frac{1}{h_v}(v)$, $g(S(v)) = -\frac{1}{S_v}(v)$, $m(w(v)) = -\frac{1}{w_v}(v)$, and multiplying by the countercumulative distributions (18) becomes

$$(19) \quad \frac{1}{\alpha_w} = \frac{DS_z}{w} \left\{ (1 + \mathcal{E}^u) \frac{1}{\alpha_h} + \mathcal{E}^c \frac{1}{\alpha_S} \right\}.$$

Similarly, letting α_π denote the local Pareto coefficient for firm profits, the envelope condition for firms $\pi_v(v) = DS_v(v)z(v)$ can be rewritten as

$$(20) \quad \frac{1}{\alpha_\pi} = \frac{DS_z}{\pi} \frac{1}{\alpha_S}.$$

Together (19) and (20) permit recovery of the local Pareto coefficients for (unobservable) CEO talent and firm asset size from the corresponding coefficients for CEO incomes and firm profits. In particular, (19) and (20) imply

$$(21) \quad (1 + \mathcal{E}^u) \frac{1}{\alpha_h} = \frac{w}{DS_z} \frac{1}{\alpha_w} - \mathcal{E}^c \frac{\pi}{DS_z} \frac{1}{\alpha_\pi}.$$

Larger values for the reciprocal of the Pareto coefficient in the right tail of a distribution indicate a thicker or fatter tail.¹⁹ Thus, (21) relates the tail thickness of the CEO talent distribution to those of the CEO income and firm profit distributions. Notice, in particular that an observed fat CEO income tail need not imply that the underlying CEO talent tail is fat. Mechanically from (21), $\frac{1}{\alpha_w}$ may be large (a fat CEO income tail) and $\frac{1}{\alpha_h}$ small (a thin CEO talent tail) if $\frac{w}{DS_z}$ is small and/or $\frac{1}{\alpha_\pi}$ large. As described previously, CEOs may be dispersed across a large interval of high incomes not because of large variations in CEO talent, but because competition among firms for CEO talent translates large variations in firm size into large CEO income variation.

¹⁹A distribution has a heavy (right) tail if its density has no more than limiting exponential decay and is fat tailed if its density has limiting geometric decay. Fat tailed distributions have finite limiting Pareto coefficients. Thinner tailed distributions have infinite limiting Pareto coefficients.

A. The Effect of Taxes on CEO Income and Profits

We now describe how the marginal tax rate impacts equilibrium CEO incomes and firm profits. To develop intuition, we start with a simple setting in which taxes are linear above a threshold income w_0 :

$$(22) \quad T[w] := T[w_0] + \tau(w - w_0) \quad w \in [w_0, \infty).$$

We consider the consequences of variation in τ (keeping w_0 , $T[w_0]$, and \bar{U} fixed throughout). We refer to $1 - \tau$ as the *retention rate* (above w_0). To decompose the effects of changes in the marginal tax on CEO incomes and to consider different equilibria parameterized by the retention rate, it will be convenient to re-express CEO income as a function of effective labor and to sometimes make its dependence on taxes explicit. Thus, we define $\omega(z; 1 - \tau)$ to be the income of a CEO supplying effective labor z when the retention rate is $1 - \tau$. Similarly, we sometimes make the dependence of equilibrium effective labor on the tax rate explicit and define $z(v; 1 - \tau)$ to be the equilibrium effective labor of CEO v given retention rate $1 - \tau$.

For simplicity, assume that firm surplus is multiplicative in S and z , i.e., $V(S, z) = DSz$. Then, as described previously, in equilibrium firm v pays its marginal product $DS(v)$ for the last (and only for the last) unit of effective labor it hires. Since larger firms have larger marginal products and hire more effective CEO labor, it follows that the *prices* paid for successively higher increments of effective labor are higher. Moreover, the total income earned by a CEO supplying effective labor z depends upon the prices paid for each incremental unit of effective labor up to z and this in turn depends upon the identity of the firms for whom these incremental units were the last hired. Specifically, in equilibrium CEO v earns

$$(23) \quad \omega(z(v; 1 - \tau); 1 - \tau) = \omega(z_0; 1 - \tau) + \int_{z_0}^{z(v; 1 - \tau)} DS(\nu(z'; 1 - \tau)) dz',$$

where z_0 is the effective labor of the least talented CEO earning income weakly more than w_0 and $\nu(\cdot, 1 - \tau)$ is the inverse of $z(\cdot, 1 - \tau)$ with $\nu(z'; 1 - \tau)$ giving the equilibrium rank of the CEO exerting effective labor z' . The overall impact of a rise in $1 - \tau$ on the income of CEO v is thus²⁰

$$(24) \quad \frac{dw}{d(1 - \tau)}(v) = \frac{\partial \omega}{\partial z}(z(v)) \frac{\partial z}{\partial(1 - \tau)}(v) + \frac{\partial \omega}{\partial(1 - \tau)}(z(v)) \\ = DS(v) \frac{\partial z}{\partial(1 - \tau)}(v) + \int_{z_0}^{z(v)} \left\{ DS_\nu(\nu(z')) \frac{\partial \nu}{\partial(1 - \tau)}(z') \right\} dz',$$

where to simplify notation dependence on $1 - \tau$ is omitted from the functions. The first term on the right-hand side of both equalities gives the impact of the retention rate change on CEO v 's effective labor and, hence, income holding the equilibrium

²⁰Equation (24) uses $\frac{\partial \omega}{\partial(1 - \tau)}(z_0, 1 - \tau) = 0$: see the online Appendix for the derivation.

pricing schedule for effective labor fixed at $\omega(\cdot; 1 - \tau)$. Provided income effects on CEO effort are not too strong, this term is positive.²¹ The second term gives the impact of the tax change on the income paid to the CEO supplying effective labor $z(v; 1 - \tau)$. This term is negative (again provided income effects on CEO effort are not too strong). To see why consider the impact of a rise in $1 - \tau$ that induces CEOs to work harder. Each incremental unit of effective labor between z_0 and a given $z(v; 1 - \tau)$ then becomes associated with a less talented CEO matched to a smaller firm with a lower marginal product of effective labor. As a result, the prices paid for these incremental units fall, as does the income paid to the CEO supplying $z(v; 1 - \tau)$.

Next consider the impact of a retention rate rise on the profit of firm v , $\pi(v; 1 - \tau) = DS(v)z(v; 1 - \tau) - \omega(z(v; 1 - \tau); 1 - \tau)$. Application of an envelope theorem to the firm's problem implies that

$$\begin{aligned}
 (25) \quad \frac{d\pi}{d(1 - \tau)}(v) &= -\frac{\partial\omega}{\partial(1 - \tau)}(z(v)) \\
 &= -\int_{z_0}^{z(v)} \left\{ DS_v(v(z')) \frac{\partial v}{\partial(1 - \tau)}(z') \right\} dz' > 0.
 \end{aligned}$$

A retention rate rise induces a downward adjustment in the schedule $\omega(\cdot; 1 - \tau)$ causing firm profits to rise. As our discussion below highlights, this spillover from the CEO income tax rate to firm profits is a force for lower optimal tax rates relative to conventional formulas.

Let $\mathcal{E}_w(v)$ and $\mathcal{E}_\pi(v)$ denote the elasticities of CEO income and firm profit with respect to the retention rate (for the v th ranked CEO and firm at retention rate $1 - \tau$). In the online Appendix, we use (24) and (25) to derive explicit expressions for these elasticities. In general, these expressions are complicated. However, they are much simplified in the case of quasilinear/constant elasticity CEO preferences. Specifically, if CEO preferences are $U(c, w) = c - \frac{\mathcal{E}}{1 + \mathcal{E}} e^{\frac{1 + \mathcal{E}}{\mathcal{E}}}$, with $\mathcal{E} > 0$ the elasticity of CEO effort, then $\mathcal{E}_w(v)$ and $\mathcal{E}_\pi(v)$ are given by

$$(26a) \quad \mathcal{E}_w(v) = \frac{DS(v)z(v)}{w(v)} \mathcal{E} - \left(\frac{\pi(v) - \pi_0}{w(v)} \right) \mathcal{E}$$

and

$$(26b) \quad \mathcal{E}_\pi(v) = \left(\frac{\pi(v) - \pi_0}{\pi(v)} \right) \mathcal{E},$$

²¹ More precisely, if CEOs' uncompensated behavioral elasticity of effort is positive, then a rise in $1 - \tau$ induces CEOs to work harder.

where π_0 is the profit of the smallest firm paying its CEO at least w_0 , $w(v) = \omega(z(v))$, and $\frac{\partial \omega}{\partial(1-\tau)} = -\left(\frac{\pi(v) - \pi_0}{\pi(v)}\right) \mathcal{E}$. Note that (26a) can be rewritten as $\mathcal{E}_w(v) = \left(1 + \frac{\pi_0}{w(v)}\right) \mathcal{E} > \mathcal{E}$ and that $\mathcal{E}_\pi(v) = \left(\frac{\pi(v) - \pi_0}{\pi(v)}\right) \mathcal{E} \in (0, \mathcal{E})$.²²

II. Optimal Linear Taxes

As a precursor to analysis of optimal *nonlinear* taxes, we consider the problem of a policymaker selecting an optimal linear tax function over a range of top incomes. The simple tax formulas in this case directly connect our results to related formulas in Saez (2001); Diamond and Saez (2011); and Piketty, Saez, and Stantcheva (2014). They also highlight the role of spillovers from CEO income taxation to firm profits and, hence, that of the social marginal value of profit in shaping and modifying conventional optimal tax formulas.

The Policymaker's Problem.—Assume that the policymaker is restricted to CEO income tax functions in the class (22) and that she selects a marginal tax rate τ to maximize a weighted sum of income tax revenues and firm profits:

$$(27) \quad \sup_{\tau \in [0, 1]} \tau \int_0^{v_0} \{w(v; 1 - \tau) - w_0\} dv + \chi \int_0^{v_0} \pi(v; 1 - \tau) dv,$$

keeping \bar{U} and w_0 fixed.²³

Interpreting χ .—The weight χ is the social marginal value of (aggregate) profit and an important parameter in our analysis. If the planner cares only about tax revenues and the social value of profit stems entirely from its role as a source of such revenues, then χ corresponds to the profit tax rate τ^F . Under this interpretation, the policymaker in (27) is simply selecting the CEO income tax rate τ to maximize overall tax revenues given the profit tax. In our assignment model, profits are pure rent and placing taxes upon them is nondistortionary. Consequently, the fully optimal tax system in which both CEO and profit taxes are chosen is one in which profits are taxed at 100 percent and $\chi = 1$. Analysis of optimal CEO income taxation with χ fixed at a value of less than 1 corresponds in our environment to a partial reform of the tax system (with the profit tax rate fixed at a suboptimal level). Alternatively, it reflects a fully optimal outcome in an extended model featuring economic frictions

²²In the language of Scheuer and Werning (2015), $\mathcal{E}_w(v)$ and $\mathcal{E}_\pi(v)$ are macro-elasticities that describe the equilibrium response of a given CEO's income to changes in the retention rate. Aggregates of these elasticities feature in the optimal tax equations of the next section. In particular, the elasticity of taxable income that is emphasized in the public finance literature and that features in the classical Diamond-Saez optimal tax equation is an aggregate of the (macro)elasticities $\mathcal{E}_w(v)$. Scheuer and Werning (2015) make the distinction between such macro-elasticities and micro-elasticities that isolate the response of an individual's income to a tax change holding everyone else's behavior fixed. They show that test conditions for the Pareto optimality of a nonlinear tax function are of a standard form when cast in terms of such micro-elasticities.

²³Here, v_0 denotes the rank of the least talented CEO earning at least w_0 . Note that the choice of τ does not affect the incomes of CEOs earning less than w_0 , which evolve from $w(\bar{v})$ according to (18) with $w(\bar{v})$ and $z(\bar{v})$ determined to ensure that firm \bar{v} maximizes profits subject to the \bar{v} CEO receiving utility \bar{U} . Hence, it does not affect v_0 .

or institutional constraints that restrict both profit taxes and the social marginal value of profit.²⁴

We briefly elaborate the implications of such frictions for the interpretation of χ and the determination of profit taxes. Assume first that firms can divert profits to owners before taxation, but that for every pretax dollar diverted the owner receives only $1 - t$ dollars. For example, suppose that a firm can realize profits in a foreign tax jurisdiction through transfer pricing. If foreign corporate and dividend income taxes equal t and owners do not repatriate profit income (because, say, they partly reside and consume in the foreign country), then the firm can divert profits to owners at the rate $1 - t$.²⁵ Alternatively, suppose that a firm can conceal profit transfers to owners as a deductible business expense at a cost of t dollars lost per dollar concealed. In these situations, each firm v , after determination of $y(v)$ and $w(v)$, will select the pretax transfer p to owners to solve

$$\max_{p \in [0, V(S(v), y(v)) - w(v)]} (1 - t)p + (1 - \chi)\{V(S(v), y(v)) - w(v) - p\}.$$

If the profit tax exceeds t , then the firm will divert all profit to owners prior to the application of the profit tax and, consequently, this tax and χ are effectively bounded by t . If the amount received by firm owners as a function of the amount diverted to them is concave rather than linear,²⁶ then the social marginal value of profit χ at the optimum no longer equals the profit tax τ^F . Rather, $\chi = \tau^F (1 - \theta(\tau^F)) \leq \tau^F \leq 1$, where $\theta(\tau^F)$ is the fraction of profit optimally transferred to owners prior to taxation at the profit tax.

It is natural to conjecture that the disincentive to create firms might also deter high rates of profit taxation and, hence, lower the social marginal value of profit χ to a planner concerned only with maximizing tax revenues. If, however, profits relax firm entry conditions, then profit has social value beyond its direct role as a base for taxation. In some situations firm entry is consistent with the optimal social marginal value of profit equaling that of tax revenues (even if the tax on firm profit is optimally set below 1). We develop the implications of the firm entry margin for the effective social marginal value of profit in the online Appendix.²⁷

The weight χ may also incorporate direct social concern for firm claimants. For example, if χ^F is the welfare weight placed on firm claimant incomes (inclusive of any lump sum transfers) and a unit welfare weight is placed on tax revenues,

²⁴A literature exploring the optimal structure of commodity and other taxation under exogenous restrictions on profit taxes developed in the 1970s and 1980s. Prominent contributions include Stiglitz and Dasgupta (1971) and Munk (1978, 1980). In these papers Ramsey optimal tax formulas are modified by a concern for the impact of commodity and factor taxes on the revenues collected from the taxation of profits. In particular, if the profit tax is constrained to be less than 100 percent, this concern is less important and a higher commodity tax may be warranted. This parallels our result that if $\chi < 1$, then marginal taxes on CEO incomes are optimally higher.

²⁵Zucman (2014) describes and attempts to measure the extent of such tax avoidance and evasion under current tax regimes. He estimates that about 20 percent of US corporate profits are realized overseas and about 80 percent of offshore accounts held by US citizens are not declared to US tax authorities.

²⁶This concavity could reflect progressive foreign taxation, high costs of overseas consumption for some owners, or convex costs of concealing transfers.

²⁷In our baseline assignment model, a fixed population of firms learns their asset size $S(v)$ and chooses whether or not to hire a CEO and operate. Only the profit of the smallest entering firm is relevant for the firm entry condition. Profits at larger firms are not and the social marginal value of profit equals the profit tax. Further, all entry costs are absorbed into V and are, thus, implicitly treated as being tax deductible.

then, absent other economic frictions, the effective weight placed on profits is $\chi = \tau^F + \chi^F(1 - \tau^F)$. If τ^F is (constrained to be) less than 1, then the effective social weight on profits is enhanced by direct concern for firm claimants and is entirely due to such concern if $\tau^F = 0$.

Further Assumptions.—The objective in (27) places no direct weight on CEOs. In our quantitative analysis in Section IV, we focus upon CEOs earning incomes above \$500,000. The zero welfare weight placed on CEOs parallels the treatment of high income earners in the optimal tax analyses of Diamond and Saez (2011) and Piketty, Saez, and Stantcheva (2014). In the online Appendix we extend our analysis to allow for positive weighting of CEOs. We also abstract from any effects of τ on the after-tax labor income of (socially valued) non-CEO workers. Implicitly, this supposes that such workers do not pay the tax either because it is specific to CEOs or because they earn incomes below the threshold w_0 and that the incomes earned by non-CEO workers are not affected by the effort of CEOs. Thus, non-CEOs are only affected by τ to the extent that they benefit from tax revenues or have a claim on firm profits.²⁸

Aggregate Elasticities and Tail Coefficients.—To state the optimal tax formulas implied by (27), it is useful to introduce a number of definitions. Let W denote the total income of CEOs earning more than the threshold income w_0 when the retention rate is $1 - \tau$ and Π the corresponding total firm profit. For notational ease we suppress the dependence of W , Π , and other variables on $1 - \tau$ for the remainder of this section. Define the *aggregate* elasticities at the optimum by

$$(28) \quad \mathcal{E}_W = \frac{1 - \tau}{W} \frac{\partial W}{\partial(1 - \tau)} \quad \text{and} \quad \mathcal{E}_\Pi = \frac{1 - \tau}{\Pi} \frac{\partial \Pi}{\partial(1 - \tau)}.$$

To the extent that high income earners are identified with CEOs, \mathcal{E}_W is the model counterpart of the elasticity of taxable income (ETI) for high earners emphasized in the empirical public finance literature.²⁹ These elasticities are aggregates of the individual level elasticities given in the preceding section: $\mathcal{E}_W = \int_0^{v_0} \frac{w}{W} \mathcal{E}_w(v) dv$ and $\mathcal{E}_\Pi = \int_0^{v_0} \frac{\pi}{\Pi} \mathcal{E}_\pi(v) dv$. Both incorporate adjustment in the equilibrium CEO income and effective labor schedules. By our previous discussion, \mathcal{E}_Π is positive under reasonable economic restrictions. In particular, in the quasilinear/constant elasticity setting using (26) and (28) we have

$$(29) \quad \mathcal{E}_W = \left\{ 1 + \frac{\pi_0}{W} \right\} \mathcal{E} \quad \text{and} \quad \mathcal{E}_\Pi = \frac{1}{A_\Pi} \mathcal{E},$$

²⁸If (4) holds, non-CEO labor is an adjustable input that complements CEO effective labor and the elasticity of such labor supply is not perfect, then higher tax rates on CEOs that deter CEO effort depress demand for non-CEO labor and, hence, the equilibrium non-CEO wage making non-CEOs worse off. Including such an effect would create an additional force for lower marginal taxation of CEOs.

²⁹See Saez, Slemrod, and Giertz (2012) for an extensive discussion of the role of ETI in public finance.

where $A_{\Pi} = \frac{\Pi}{\Delta\Pi}$, with $\Delta\Pi = \Pi - \pi_0$, and as before \mathcal{E} is the CEO's effort elasticity. Let $A_W := \frac{W}{\Delta W}$ where $\Delta W := W - w_0$. We refer to A_W and A_{Π} as the tail coefficients of CEO income and profit. It is easy to verify that

$$\frac{1}{A_W} = \int_0^{v_0} \frac{w(v)}{W} \frac{1}{\alpha_w(v)} dv.$$

In particular, if CEO income has a Paretian right tail above w_0 , then $A_W = \alpha_w$, where α_w is the constant Pareto coefficient.

Optimal Tax Formulas.—With the preceding definitions in place, a simple formula for the optimal marginal income tax rate τ^* is available. Rearranging the first-order condition from (27) and using the definitions given above yields

$$(30) \quad \tau^* = \frac{1 - \chi A_W^* \frac{\Pi^*}{W^*} \mathcal{E}_{\Pi}^*}{1 + A_W^* \mathcal{E}_W^*},$$

where $*$ denotes optimal values. Formula (30) contrasts with the standard expression derived by Saez (2001) and emphasized by Diamond and Saez (2011):

$$(31) \quad \tau^{Saez} = \frac{1}{1 + A_W^* \mathcal{E}_W^*}.$$

The logic behind (30) extends that behind (31) to include concern for the spillover from CEO income taxes to profits.³⁰

Given $A_W^* \frac{\Pi^*}{W^*} \mathcal{E}_{\Pi}^* > 0$, formulas (30) and (31) give the same value for the optimal tax rate only if profits receive no weight in the policymaker objective ($\chi = 0$). More generally, if $\chi > 0$, then the depressing effect of τ on profits creates a motive for lower marginal taxes.

Optimal Taxes when $\chi = 0$.—Diamond and Saez (2011) and Saez, Slemrod, and Giertz (2012) use formula (31) to provide guidance on the optimal taxation of top earners. As noted, this optimal tax formula emerges as the appropriate one in our assignment model as well (only) if $\chi = 0$. Following the discussion earlier in this section, this is an extreme case requiring no ability to tax profits, no impact of profits on firm entry conditions, and no concern for the recipients of firm profits. However, as a benchmark for subsequent calculations, we evaluate (31).

Various authors proceed as if $A_W \mathcal{E}_W$ is relatively stable in the face of marginal tax rate changes and use empirical evaluations of $A_W \mathcal{E}_W$ in US data (and at a prevailing allocation) to determine or at least approximate its value at the optimum. Our model is consistent with such a strategy if CEO preferences are of the quasilinear/constant elasticity form, firm surplus is multiplicative, and the tax rate is linear

³⁰Note formula (30) does not rely on the assignment framework. It is valid whenever there is a spillover from the CEO income tax rate to firm profits and, in this sense, is quite general. The assignment model supplies the mechanism underlying this spillover and facilitates our subsequent empirical strategy for evaluating optimal non-linear taxes.

above a threshold. Based on prior empirical analyses of the general population (of top earners), Diamond and Saez (2011) and Saez, Slemrod, and Giertz (2012) set A_W equal to 1.5 and \mathcal{E}_W to 0.25 implying a top tax rate of 72.7 percent.³¹ It is possible that the population of top earning CEOs is different from the general population of top earners. We estimate A_W for the CEO population using the Standard and Poor's ExecuComp database and find that it is stable above an income of about \$12 million (in 2011 US\$) and equal to 2.1. There is limited direct evidence on \mathcal{E}_W for top CEOs. Frydman and Molloy (2011) report a strong negative correlation between top marginal tax rates and aggregate CEO incomes in the United States. However, they also estimate a small contemporaneous response of CEO incomes to tax reforms in the cross section. Based on this, they reject a value of \mathcal{E}_W for CEOs above 0.2; their largest point estimate is about 0.094. Goolsbee (2000) studies data from 1991 to 1995 and rejects an elasticity above 0.4. If \mathcal{E}_W is set equal to the 0.25 value proposed by Diamond and Saez (2011) and Saez, Slemrod, and Giertz (2012), and A_W is set to 2.1, then (31) implies a top tax rate of 65.6 percent. Lower values for \mathcal{E}_W would imply higher marginal tax rates. In particular, if $\mathcal{E}_W = 0.1$, then the top tax rate is 82.6 percent.

Fiscal Spillovers.—Saez, Slemrod, and Giertz (2012) and Piketty, Saez, and Stantcheva (2014) emphasize positive fiscal spillovers from income tax rates to other tax bases and modify formula (31) accordingly. These spillovers create motives for even higher marginal tax rates than those reported above. Specifically, Saez, Slemrod, and Giertz (2012) consider shifts of income from the personal to corporate tax bases in response to higher marginal tax rates. They assume that 50 percent of income is shifted and that this shifted income is taxed at 30 percent. Their modified version of (31) then implies that the optimal tax rate rises from 72.7 percent to 76.8 percent. Piketty, Saez, and Stantcheva (2014) assume that income earners can engage in rent-seeking at the expense of tax revenues. Higher marginal tax rates deter such rent-seeking. They suggest that the elasticity of earnings from rent-seeking with respect to the retention rate might be at least 0.3, implying an optimal top marginal tax rate of 83 percent.³²

Much of the CEO literature is, however, consistent with a positive impact of CEO effective labor on firm profits and, hence, a negative spillover from top income taxation to firm profits. In particular, as we have stressed, this is true of the CEO assignment model augmented with CEO effort. We focus upon this.³³ At the same time, the literature has given little guidance on the size of this spillover and, in particular, on the magnitude of \mathcal{E}_{Π} . In the online Appendix we provide simple regressions of the dividend-GDP ratio on top marginal income tax rates (and corporate tax

³¹Saez, Slemrod, and Giertz (2012) argue that the best available estimates for the (long-run) elasticity of earnings with respect to the retention rate range between 0.12 and 0.4. They select a value of 0.25 as the midpoint of these.

³²Stantcheva (2014) considers a model in which neither the policymaker nor firms observe workers types. Firms offer menus of contracts and workers of different types select among them. Taxes perturb this menu, hence inducing further redistributions amongst worker types (rather than between CEOs and firm owners as occurs in our setting). This introduces a term, analogous to $-\chi \frac{\Pi^*}{W^*} \mathcal{E}_{\Pi}$, that captures the social desirability of such redistribution into her optimal tax formulas.

³³In doing so, we abstract from sensitivity of rent seeking to taxation. Equally, we abstract from other positive benefits of CEO effort such as job creation for non-CEOs.

rates) for the United States that are consistent with a small positive value for \mathcal{E}_Π .³⁴ These regressions do not, of course, establish a casual relationship between income tax rates and firm profits. Moreover, 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. Empirical measures of profit must be appropriately adjusted. In Section IV, we pursue a different approach and obtain implications of a calibrated assignment model for optimal (nonlinear) CEO income taxation. This approach implicitly characterizes spillovers to profit.

A Restatement of the Optimal Tax Formula.—We now restrict attention to the case of quasilinear and constant elasticity CEO preferences and a multiplicative firm surplus function. In this setting, we derive an alternative version of the optimal tax formula in terms of the primitive talent distribution and effort elasticity for the case $\chi = 1$. This formula anticipates the optimal nonlinear tax formulas derived in the next section (without the assumption of quasilinearity or a constant elasticity).

The first-order condition for τ^* in (27) may be organized as

$$(32) \quad \frac{\tau^*}{1 - \tau^*} (W^* \mathcal{E}_W^* + \Pi^* \mathcal{E}_\Pi^*) + \left(\frac{\chi - \tau^*}{1 - \tau^*} \right) \Pi^* \mathcal{E}_\Pi^* - \Delta W^* = 0.$$

The first term gives the marginal behavioral impact on tax revenues of a change in $1 - \tau$ if the *entire* firm surplus $R^* = \int_0^{v_0} r^* dv$, $r^*(v) = DS(v)z(v; 1 - \tau^*)$, is taxed at the rate τ^* . The second term adjusts the first to take into account the fact that profits are only taxed at (or valued at) the rate χ . The third term gives the mechanical effect on tax revenues of a change in $1 - \tau$ holding the distribution of CEO incomes fixed. If $\chi = 1$ and CEO preferences are of the quasilinear and constant elasticity form, then $\Delta W^* - \left(\frac{\chi - \tau^*}{1 - \tau^*} \right) \Pi^* \mathcal{E}_\Pi^* = \Delta W^* - \Pi^* \mathcal{E}_\Pi^* = (1 + \mathcal{E}) \int_0^{v_0} r^* \frac{1}{\alpha_h} dv$, where, as before, α_h is the local Pareto coefficient of the talent distribution and \mathcal{E} the effort elasticity. In addition, the first term in (32) is simply $\frac{\tau^*}{1 - \tau^*} R^* \mathcal{E}$. Consequently, when $\chi = 1$ and CEO preferences are of the quasilinear and constant elasticity form, equation (32) implies

$$(33) \quad \tau^* = \frac{1}{1 + \frac{\mathcal{E}}{1 + \mathcal{E}} \frac{1}{\int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv}}.$$

If the talent distribution has a Paretian right tail, then (33) further reduces to

$$(34) \quad \tau^* = \frac{1}{1 + \frac{\mathcal{E}}{1 + \mathcal{E}} \alpha_h}.$$

³⁴Note that even if \mathcal{E}_Π^* is small, the impact of τ on profits can be large *relative to its impact on CEO income tax revenues* if $\frac{\Pi^*}{W^*}$ is large enough. In this case, for a large enough value of χ , the depressing effect of τ on firm profits can still translate into a significant depressing effect on optimal tax rates.

Formulas (33) and (34) anticipate optimal nonlinear tax formulas derived in the next section (without the assumption of quasilinearity or a constant elasticity). Strikingly, these formulas hold in standard (i.e., nonassignment) labor market settings in which there are no spillovers to profits. Thus, when $\chi = 1$ the standard optimal tax formula expressed in terms of talents (33) holds, but the standard formula expressed in terms of incomes (31) does not and when $\chi = 0$ the situation is reversed. For alternative χ values neither standard formula holds. The logic behind these results is as follows. The standard formula expressed in terms of incomes (31) is valid when the policymaker seeks to maximize income tax revenues; the standard formula expressed in terms of talents (34) is valid when the policymaker seeks to maximize the total surplus not captured by CEOs (or, more generally, high income earners). In the conventional optimal tax setting, total income tax revenue equals total surplus not captured by high income earners and so both formulas hold. But in the assignment setting some surplus is paid as profit to firm claimants. Hence, either the policymaker is maximizing income tax revenues or she is maximizing total surplus not captured by CEOs (or she is maximizing a weighted sum of the two) and so at most one of the formulas holds.

Note that in the conventional labor supply setting (33) is consistent with a high optimal marginal tax rate. In particular, in this setting $r^* = w^*$ and $(1 + \mathcal{E}) \frac{1}{\alpha_h} = \frac{1}{\alpha_w^*}$ implying $\frac{1}{1 + \mathcal{E}} \frac{1}{\int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv} = \frac{1}{\int_0^{v_0} \frac{w^*}{W^*} \frac{1}{\alpha_w^*} dv} = A_W^*$. In addition, $\mathcal{E} = \mathcal{E}_W^*$ and so $\frac{\mathcal{E}}{1 + \mathcal{E}} \frac{1}{\int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv} = A_W^* \mathcal{E}_W^*$. As noted previously, empirical evaluations of $A_W^* \mathcal{E}_W^*$ are relatively small implying a high tax rate. In contrast, in the assignment setting the talent distribution has a thinner right tail than the income distribution (see equation (21)), and so α_h is larger than α_w^* (and larger than would be implied by attempts to infer α_h from the income data using the standard labor market model). Thus, in the assignment model, $\frac{1}{1 + \mathcal{E}} \frac{1}{\int_0^{v_0} \frac{r^*}{R^*} \frac{1}{\alpha_h} dv} > \frac{1}{\int_0^{v_0} \frac{w^*}{W^*} \frac{1}{\alpha_w^*} dv} = A_W^*$ and (33) is consistent with a lower optimal tax rate than (31).

III. Optimal Nonlinear Taxation

We now generalize our earlier results and consider the policymaker’s optimal choice of nonlinear tax function over (all) CEO incomes. In the general nonlinear setting, the simplest and most direct way of deriving optimal nonlinear tax formulas is to formulate the policymaker’s problem as a mechanism design problem and then recover optimal taxes from the associated first-order conditions. This gives optimal formulas in terms of effort elasticities and (local) Pareto coefficients of the talent and firm asset distributions. We subsequently derive formulas in terms of the CEO income and firm profit distributions and the CEO income elasticity via direct perturbation of the tax function. Note that the latter is complicated relative to Saez (2001) by the endogeneity of the CEO income schedule.

The Policymaker’s Mechanism Design Problem.—It is convenient to reformulate a tax equilibrium in terms of a tuple (\tilde{v}, z, w, Φ) , where Φ gives the CEO’s utility

with $\Phi(v) := U(c(v), z(v)/h(v))$. Let $\mathcal{C}(\phi, z/h)$ be the consumption of a CEO when her utility is ϕ and her effort z/h . We focus on smooth allocations and relax the global CEO and firm incentive constraints (12) and (13), replacing them with, respectively, the CEO's envelope condition and the firm's first-order condition.³⁵ To further simplify matters and to align our work with Diamond and Saez (2011), in the main text we (continue to) focus on the case in which CEOs receive zero welfare weight.³⁶ Let T^0 denote tax revenues (or more generally social surplus) generated by unmatched CEOs. The policymaker's problem can then be formulated as the optimal control problem:³⁷

(35)

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

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

$$(36) \quad \Phi_v(v) = -U_e \left(\mathcal{C} \left[\Phi(v), \frac{z(v)}{h(v)} \right], \frac{z(v)}{h(v)} \right) \frac{z(v)}{h(v)} \frac{h_v(v)}{h(v)}$$

and

$$(37) \quad w_v(v) = V_z(S(v), z(v))z_v(v).$$

The Optimal Tax Formula in Terms of Primitive Distributions.—After manipulating the first-order and co-state equations from the optimal control problem (35), the following optimality condition emerges for all $v \in (0, \tilde{v}]$:³⁸

(38)

$$V_z + \frac{U_e/h}{U_c} = -\frac{p^\Phi}{h} \left\{ \left[U_{ec} \left(-\frac{U_e}{U_c} \right) + U_{ee} \right] \frac{z}{h} + U_e \right\} \frac{-h_v}{h} + (1 - \chi)vV_{zS}(-S_v),$$

where p^Φ is the co-state associated with CEO utility Φ . This condition captures the marginal benefits and costs associated with a small change in CEO v 's effective labor (holding her utility fixed). It has a very natural interpretation. The left-hand side of (38) gives the marginal benefit of the CEO's (compensated) labor supply increase. It consists of the marginal increase in firm surplus V_z less the additional consumption needed to maintain the CEO at her previous utility level $U_e h/U_c$. Relative to the standard Mirrlees model the only modification is that the marginal product of

³⁵In our later calculations we check for monotonicity and an absence of bunching ex post at the optimum, ensuring these local conditions are sufficient for incentive compatibility.

³⁶The general case in which CEOs receive nonnegative welfare weights is analyzed in the online Appendix.

³⁷The case $\chi = 1$ is somewhat easier to solve since w no longer appears in (35). It can be formulated as an optimal control problem with two rather than three state variables.

³⁸All functions below are understood to be evaluated at (the allocation associated with) the CEO's rank v . To economize on notation the v argument is dropped.

effective labor equals V_z rather than 1. The terms on the right-hand side capture two sources of welfare loss associated with a small increase in CEO v 's effective labor. The first is the (standard) increment to the information rent paid to more productive CEOs in order that the CEOs' incentive compatibility conditions continue to hold. The second term is novel to the assignment setting. It can be interpreted as the welfare loss associated with a redistribution from CEO income tax revenues to firm profits at firms ranked above v . The economic forces underlying this redistribution, though viewed from the perspective of a local perturbation in labor supply, rather than a global perturbation in the marginal tax rate, are essentially the same as in the linear tax setting. Recall that a unit of effective labor is priced by (i.e., paid the marginal product of) the firm for which it is the last unit hired. An increase in CEO effective labor at firm v implies that the additional effective labor is priced by firm v rather than a slightly higher ranked and more productive firm. Thus, its (shadow) price falls and the incomes paid to all greater CEO effective labor supplies are correspondingly reduced. In the associated tax equilibrium, firm profits rise, tax revenues collected from CEOs fall and, if $\chi < 1$, welfare falls. If, on the other hand, $\chi = 1$, because, for example, the policymaker can tax firm profits at 100 percent, then this redistributive effect has no impact on welfare. In this case, the policy objective is to maximize surplus extracted from CEOs, and (38) reduces to the standard optimality condition in Mirrlees models (with zero weight on top earners and modulo the change to the marginal product of effective labor on the left-hand side).

Together the CEOs' and the firms' first-order conditions (14) and (16) imply

$$(39) \quad (1 - T_w[w])V_z h U_c = -U_e.$$

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 pretax return on effort equates to the right-hand side of (38). If $\chi = 1$, then the right-hand side of (38) equals the usual marginal informational rents term from the Mirrlees model and combining (38) and (39) and the definition of α_h , the standard formula for optimal marginal tax rates is obtained:

$$(40) \quad T_w[w] = \frac{1}{1 + \frac{1}{\tilde{p}^\Phi} \frac{\mathcal{E}^c}{1 + \mathcal{E}^u} \alpha_h},$$

where \tilde{p}^Φ is the normalized co-state $\frac{U_c}{1 - F(h)} p^\Phi$. More generally, using (38) and (39) and the definitions of α_h and α_S , the optimal marginal tax rate is

$$(41) \quad T_w[w] = \frac{1 + (1 - \chi) \frac{1}{\tilde{p}^\Phi} \frac{\mathcal{E}^c}{1 + \mathcal{E}^u} \frac{V_z S}{V_z} \frac{\alpha_h}{\alpha_S}}{1 + \frac{1}{\tilde{p}^\Phi} \frac{\mathcal{E}^c}{1 + \mathcal{E}^u} \alpha_h}.$$

Intuitively, when $\chi < 1$ and the policymaker places more weight on CEO income tax revenues than firm profits, then the second marginal cost term in (38) becomes

relevant. In this case, consistent with the intuition described above, the policymaker sets a higher marginal tax rate on a CEO's income in order to reduce that CEO's effective labor supply, modify the (implicit) pricing schedule for effective labor and, hence, redistribute from firm profits to the incomes of more talented CEOs. In doing so the policymaker achieves her goal of collecting more tax revenues from CEOs.

The Optimal Tax Formula in Terms of the Income Distribution.—In the linear tax setting we presented an optimal tax formula in terms of the elasticities and tail coefficients of CEO incomes and firm profits. This formula was obtained from a direct perturbation of the optimal tax function. Such perturbations are more complicated in the nonlinear setting. We pursue such a perturbation now and, hence, relate equations (40) and (41) to nonlinear tax formulas in terms of induced CEO incomes and (spillovers to) firm profits. We focus on the quasilinear/constant elasticity CEO preferences and multiplicative firm objective case.

The perturbation we consider involves a small modification to the marginal tax rate over a small interval of incomes starting at income w_0 . Following Saez (2001), its impact on tax revenues can be decomposed into *mechanical* and *behavioral* parts. The first of these equals $1 - M(w_0)$; it gives the revenue response to the perturbation holding the CEO income schedule w fixed. The second component is the behavioral part. This is more complicated than in standard models since it incorporates the impact of the tax change on the equilibrium schedule of incomes. It may be expressed compactly as

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

where $\tilde{\mathcal{E}}_w(w_0)$ is a weighted elasticity of CEO incomes at and above w_0 with respect to the local retention rate $1 - T_w[w_0]$. The precise formula for this elasticity is given in the online Appendix. In addition to the usual *local* effect on tax revenues caused by the CEO at w_0 working harder in response to a higher retention rate, this elasticity also incorporates a *global* effect on revenues collected from CEOs earning more than w_0 . This is caused by the downward shift in the equilibrium income schedule discussed previously.

When $\chi = 0$, the policymaker is concerned only with maximizing CEO income tax revenues. In this case the sum of the mechanical and behavioral impacts derived above must equal 0 at the optimum and so combining terms:

$$T_w^*[w_0] = \frac{1}{1 + \alpha_w(v_0)\tilde{\mathcal{E}}_w(w_0)},$$

where v_0 is the rank of the CEO earning w_0 and $\alpha_w(v_0)$ is the corresponding local Pareto coefficient of income.

For χ values greater than 0, the policymaker is also concerned with the spillover of marginal tax rates to firm profits. Consequently, the behavioral term is augmented with an extra component that captures the enhancing effect of a higher retention rate

at w_0 on the profits of firms ranked above v_0 . Again this enhancement is due to the reduction in CEO incomes at each effective labor supply above z_0 . It is given by

$$(42) \quad -\chi \frac{1}{1 - T_w[w_0]} \frac{\mathcal{E} \frac{S_0 z_0}{w_0}}{1 + \mathcal{E} \mathcal{T}(w_0) \frac{S_0 z_0}{w_0}} \frac{\alpha_w(v_0)}{\alpha_S(v_0)} \\ \times \int_0^{v_0} \exp \left\{ - \int_v^{v_0} \frac{\frac{S_v(v')}{S(v')} \mathcal{E} \mathcal{T}(w^*(v')) \frac{S(v') z^*(v')}{w(v')}}{1 + \mathcal{E} \mathcal{T}(w^*(v')) \frac{S(v') z^*(v')}{w^*(v')}} dv' \right\} dv,$$

where $\mathcal{T}(w) = \frac{w T_{ww}[w]}{1 - T_w[w]}$ is the elasticity of the marginal retention function. The term (42) is the analogue for the nonlinear setting of the spillover term $\chi \frac{\Pi}{1 - \tau} \mathcal{E}_{\Pi}$ in the derivation of (30). Adding it to the other behavioral and mechanical terms leads to an optimal tax equation for the nonlinear setting analogous to (30).

IV. Quantitative Evaluation of Optimal Nonlinear Taxes

Equation (41) may be used to compute optimal marginal taxes over a range of CEO incomes. We specialize the analysis to quasilinear/constant elasticity CEO preferences and a multiplicative firm surplus function. In this case (41) may be rewritten as

$$(43) \quad T_w^*[w^*(v)] = \frac{1 + (1 - \chi) \frac{\mathcal{E}}{1 + \mathcal{E}} \frac{\alpha_h(v)}{\alpha_S(v)}}{1 + \frac{\mathcal{E}}{1 + \mathcal{E}} \alpha_h(v)}.$$

To quantitatively evaluate the implications of this formula for optimal taxes values for the weight χ , the Pareto coefficients α_h and α_S and the elasticity \mathcal{E} are required. We discuss the choice of χ first.

A. Selecting Values for χ

As described in Section II the weight χ is the social marginal value of profit. In our baseline model, it corresponds simply to the profit tax rate, τ^F . We consider a range of values for χ between the empirically observed profit tax rate in the United States and the optimal profit rate of 1. In the United States, corporate profits over much of the tax schedule are taxed at 35 percent. These are augmented by state level taxes and by dividend taxes placed on disbursed profits. The OECD reports that the overall tax rates on dividend income in the United States was 57.6 percent in 2015.³⁹

³⁹This value includes the marginal statutory corporate income tax rate on distributed profits and the sum of the maximum federal personal and average state income tax rate on dividends. Data taken from OECD Tax Database,

Since it is possible that statutory rates overstate the effective tax rate on profits, we consider values of χ as low as 0.4. Note that χ also incorporates direct social concern for profit claimants and, in more elaborate models, the social value of relaxing constraints on firm entry. These considerations raise χ above the profit tax rate and in some models increase χ to 1.

The extent of social concern for firm claimants is a normative and ethical question, though one likely influenced by the incomes of firm claimants. Ownership of corporate equities extends well beyond top CEOs (to whom we continue to attach zero weight) and to those with much lower incomes. To see this, we use data extracted from the *Survey of Consumer Finances* (SCF) for 2013. The SCF provides a measure of directly and indirectly held equities (equities in stock mutual funds, IRAs/Keoghs, and other managed assets). The total value of equities held by households in the SCF equals \$15,904 billion.⁴⁰ Overall, 65 percent of the reported value of equity in the SCF is in the hands of households with incomes of less than \$500,000. The median household in the SCF reports an income of about \$51,700 and equity (held directly and indirectly) valued at \$33,000 and constituting 40 percent of its financial wealth.

B. Recovering Measures of α_h and α_S

In our firm-CEO assignment model, equations (19) and (20) relate the Pareto coefficients for the CEO talent distribution and firm assets to those for CEO incomes and firm profits. We seek to use these equations to determine α_h and α_S . There are two complications in doing so. First use of (19) and (20) requires measurement of DSz , i.e., measurement of firm surplus after payment to (non-CEO) adjustable inputs, and, hence, measurement of economic profit. A firm's market capitalization combines the capitalized value of such surpluses (net of payments to its CEO) with the value of the firm's adjustable capital. Recovery of the value of firm surpluses, thus requires disentangling these from the value of adjustable capital. Furthermore, the surplus from a given application of CEO effective labor may be realized over time. Specifically, a given set of CEO decisions may have a long lasting impact on a firm's stream of surpluses. DSz corresponds to the value of this stream rather than the contemporaneous value of surplus. To handle these issues, we follow the procedure of Terviö which requires introducing two parameters describing the share of gross surplus paid to adjustable capital and the rate of decay of CEO effective labor on surplus. We select the parameter values suggested by Terviö (2008) and undertake sensitivity analysis around them. A second issue in applying (19) and (20) concerns the decomposition of effective labor into its talent and effort components. This hinges on the elasticity of CEO effort. We make conservative choices in this regard and undertake sensitivity analysis around them. Our choices are consistent with modest values for the elasticity of taxable income and with moderate variation in effort across the population of CEOs.

Corporate and Capital Income Taxes: Table II.4. Historical values for the United States range from 51.7 percent in 2003 to 88.7 percent in 1981. The average value for OECD countries in 2015 is 43.1 percent.

⁴⁰The *Flow of Funds* reports that in 2013 the total market value of domestic US corporations is equal to \$27,183 billion (Table L.223, Lines 2 + 3). Total direct and indirect ownership of the household sector amounts to \$19,502 billion (Table L.223, Lines 6 + 17). Additional amounts are held by pension funds (\$4,888 billion) (Table L.223, Lines 14 + 15 + 16) and insurance and life insurance companies (\$2,053 billion) (Table L.223, Lines 12 + 13).

Connecting α_n and α_S to Firm Market Capitalization and CEO Income Data.— Along the lines of Terviö (2008), we first provide a simple dynamic extension of the environment in Section I that accommodates productivity growth and long-lived effects of CEO effort. Specifically, we assume that (i) firms are infinitely lived and firm productivity grows at a constant and common rate g ; (ii) CEOs marginal utility of consumption decays at a steady rate g over time; and (iii) the CEO's outside utility option is constant. In addition, we assume that the tax function in successive periods is linear above a productivity-adjusted threshold income $(1 + g)^t w_0$ and that the tax levied at this threshold grows at a rate g , i.e., $T_t[(1 + g)^t w_0] = (1 + g)^t T[w_0]$. This assumption implies that if a CEO's income grows at a rate g and has an initial value in excess of w_0 , then the CEO's tax liabilities also grow at this rate. In addition, we assume that the impact of CEO effective labor exerted at t has a long-lived effect on firm surplus that decays at rate λ . The latter assumption implies that if a firm buys a stream of effective labor $\{z_r\}_{r=-\infty}^{\infty}$ from a sequence of CEOs, the effective labor applied at date t within the firm is

$$(44) \quad Z_t = \lambda \sum_{i=0}^{\infty} \frac{z_{t-i}}{(1 + \lambda)^{i+1}}.$$

These assumptions correspond to (and extend) the “strong stationarity conditions” of Terviö (2008). Combined with quasilinear-constant elasticity CEO preferences and a multiplicative firm surplus, they ensure a stationary equilibrium in which firm surpluses and CEO incomes scale up by a factor of $1 + g$ in each period and the effective labor supplies of CEOs remain constant.

Assume that $\{CSZ_t(1 + g)^t\}^{1-\theta} k_t^\theta$ gives the firm payoff at date t net of payment to and after maximization over all adjustable inputs except capital. If r denotes the (after-tax) rental cost of adjustable capital, then a firm with asset S using effective labor Z_t obtains a surplus at t of

$$(45) \quad \hat{V}_t(S, Z_t) := \max_{k_t \geq 0} \{CSZ_t(1 + g)^t\}^{1-\theta} k_t^\theta - rk_t.$$

After maximization over k_t (and with appropriate choice of the constant C and units for S), the date t surplus has the multiplicative form $\hat{V}_t(S, Z_t) = (1 + g)^t SZ_t$. Hence, the present discounted value of firm surpluses at date t is

$$(46) \quad (1 + g)^t \sum_{j=0}^{\infty} B^j SZ_{t+j},$$

where $B = \frac{1+g}{1+r}$ is the growth-adjusted firm discount factor. Since the effective labor z_t supplied by a CEO at t has a long-lasting effect, the (present discounted value of the) surplus it generates is

$$(47) \quad V_t(S, z_t) := (1 + g)^t DSz_t, \quad \text{with } D := \frac{\lambda}{1 + \lambda - B}.$$

In a stationary equilibrium, the income received by the v th CEO grows at the rate g , i.e., $w_t(v) = (1 + g)^t w(v)$ and the v th firm chooses v' to maximize $(1 + g)^t \{DS(v)z(v') - w(v')\}$. As in previous sections, its first-order condition implies

$$(48) \quad \frac{w_v}{w} = \left(\frac{DSz}{w}\right) \frac{z_v}{z}.$$

Equation (48) is central to our two-step identification strategy. The first step (as in Terviö 2008) is to relate firm surplus DSz to (observed) market capitalization. The second step (new to this paper) is to isolate the component of effective labor variation z_v/z that is due to talent variation and substitute out that part which is not. We describe these steps next.

In the stationary equilibrium, the v th firm matches with the v th CEO and hires a constant amount of effective labor $z(v)$. Thus, the effective labor used by the firm at t equals that supplied by the CEO at t and $Z_t(v) = z(v)$. Consequently, the capitalized value of the v th firm's profits at date 0 is

$$(49) \quad P(v) := \frac{S(v)z(v) - w(v)}{1 - B}.$$

The market capitalization of the v th firm at date 0 augments $P(v)$ with the date 0 adjustable capital choice $k_0(v)$ and is given by $Q(v) := k_0(v) + P(v)$. Date 0 adjustable capital $k_0(v)$ may be recovered from (45). Substituting this optimal choice and the definition of $P(v)$ in (49) into the definition for the market capitalization $Q(v)$ implies

$$(50) \quad Q(v) = \left(\frac{1}{\Xi}\right) \frac{S(v)z(v)}{1 - B} - \frac{w(v)}{1 - B},$$

where $\Xi := \frac{1 - \theta}{1 - \theta + \frac{\theta}{r}(1 - B)}$ inflates the firm surplus to account for adjustable capital. Hence, the firm's first-order condition (48) can be re-expressed as

$$(51) \quad \frac{w_v}{w} = \left(\frac{DSz}{w}\right) \frac{z_v}{z} = D\Xi \left(\frac{w + (1 - B)Q}{w}\right) \frac{z_v}{z},$$

where (50) is used to replace the unobserved firm surplus Sz with observed market capitalization Q and obtain the second equality. In our setting, CEO effective labor variation z_v/z is attributable partly to variation in talent h_v/h and partly to variation in effort e_v/e . The latter in turn is related to variation in the return to effort which depends on CEO talent and firm asset size. Specifically, and as in the static case, if the tax system in period 0 is linear across high incomes, then the v th CEO's first-order condition implies

$$(52) \quad \frac{z_v}{z} = \frac{h_v}{h} + \frac{e_v}{e} = (1 + \mathcal{E}) \frac{h_v}{h} + \mathcal{E} \frac{S_v}{S}.$$

Together (51) and (52) relate CEO income variation w_v/w to CEO talent h_v/h and firm asset size variation S_v/S . It remains to eliminate $\mathcal{E}S_v/S$ from this, i.e., to eliminate that part of effort variation attributable to (unobserved) firm asset variation. This is done by totally differentiating (50) to obtain an expression for S_v/S and combining the result with (51) and (52) to give a differential equation for CEO talent in terms of observable CEO income and firm market capitalization. Re-expressing this in terms of local Pareto coefficients (with α_Q the local Pareto coefficient for firm market capitalization), yields

$$(53) \quad \frac{1}{\alpha_h} = \mathcal{N} \left(\frac{w}{w + (1-B)Q} \right) \frac{1}{\alpha_w} + \mathcal{P} \left(\frac{(1-B)Q}{w + (1-B)Q} \right) \frac{1}{\alpha_Q},$$

where $\mathcal{N} > 0 > \mathcal{P}$ are constants depending on the parameters λ , r , g , θ , and \mathcal{E} . Equation (53) is the empirically operational version of (21). It implies that the reciprocal Pareto coefficient for CEO talent is a weighted sum of the coefficients for CEO income and firm market capitalization (rather than CEO income and profit as in (21)). It is distinct from analogous expressions in Terviö (2008) in that it purges out local variation in CEO income due to variation in effort. In particular, and analogous to the weighting of the profit Pareto coefficient in (21), the weight $\mathcal{P} = -\frac{\mathcal{E}}{1+\mathcal{E}}$ on the reciprocal Pareto coefficient for firm market capitalization, $\frac{1}{\alpha_Q}$, is negative. Heuristically, greater variation in market capitalization is associated with greater variation in firm asset size and a correspondingly greater contribution of CEO effort variation to firm effective labor variation. This in turn implies a smaller role for CEO talent variation in explaining CEO effective labor and income variation.

Estimates of α_S are also needed. A derivation very similar to that underpinning (53) yields

$$(54) \quad \frac{1}{\alpha_S} = \mathcal{M} \left(\frac{w}{w + (1-B)Q} \right) \frac{1}{\alpha_w} + \left(\frac{(1-B)Q}{w + (1-B)Q} \right) \frac{1}{\alpha_Q},$$

where $\mathcal{M} < 0$ depends upon parameters.

Connecting CEO Income and Firm Market Capitalization Data.—We use (53) and (54) to recover estimates of the local Pareto coefficients for talent and firm size. This requires prior estimation of the local Pareto coefficients for CEO compensation and firm market capitalization along with calculation of the parameters B , \mathcal{N} , \mathcal{P} , and \mathcal{M} . We use CEO compensation data for the year 2011 taken from the Standard and Poor's ExecuComp database. 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. The value of options received as compensation is calculated by determining the profit obtained at the time the options are exercised. We restrict our sample to CEOs with reported income above 2011 US\$500,000. The final sample contains 1,683 CEOs. We compute firm market capitalizations using data on the number of shares outstanding and average monthly share price contained in the Center for Research in Security Prices (CRSP) database. The model predicts a perfect ordering between CEO income and market

capitalization. This relationship, known as Roberts' Law, is robust in the data, but is obviously not perfect. To bring the model to the data, we order by CEO income. We then impute corresponding market capitalizations by estimating the following log-linear relationship:

$$\log Q_i = \beta_0 + \beta_1 \log w_i + \varepsilon_i,$$

where w_i is the income of the i th CEO and Q_i is the market capitalization of the firm the i th CEO manages. Using the estimated coefficients $(\hat{\beta}_0, \hat{\beta}_1)$, market capitalization values are set equal to⁴¹

$$\log \hat{Q}_i = \hat{\beta}_0 + \hat{\beta}_1 \log w_i.$$

We find that the right tail of the CEO income and imputed market capitalization distributions are well described by Pareto distributions. We fit Pareto distributions to both and, hence, recover estimates of α_w and α_Q .

Selecting Parameter Values.—We follow Terviö (2008) and select values for $g = 0.025$, $r = 0.05$, $\lambda = 0.5$, and $\theta = 0.4$. Robustness tests of our results with respect to these parameters are performed in Section IVD. We choose a benchmark value of $\mathcal{E} = 1/15$, but also consider other elasticity values in the range $1/20$ to $1/10$. Collectively, these choices pin down the parameters B , \mathcal{N} , and \mathcal{P} in (53) and \mathcal{M} in (54). Note that the data and the requirement that α_h be positive place some restrictions on parameter choices. In particular, given the benchmark parameterization of Terviö (2008), positivity of the talent Pareto coefficient requires that the elasticity \mathcal{E} be below 0.12.⁴² Our choices for \mathcal{E} are consistent with this bound. If, in addition to our CEO preference assumptions, the tax rate is linear above w_0 , then $\mathcal{E}_W = \Gamma \mathcal{E}$, where $\Gamma = 1 + \frac{D\Xi(1-B)Q(0) + (D\Xi - 1)w(0)}{W}$. Our data and benchmark parameter choices imply a value of $\Gamma = 1.53$ and, hence, a value of \mathcal{E}_W approximately equal to 0.1. The alternative values for \mathcal{E} we consider imply a range for \mathcal{E}_W between 0.0765 and 0.153. These values align with the estimates of Frydman and Molloy (2011). They obtain point estimates of \mathcal{E}_W below 0.1 and reject values above 0.2. Our implied values for \mathcal{E}_W are also in line with those of Gruber and Saez (2002) who obtain an elasticity of broad income (before deductions) of 0.12 for the general population and 0.17 for high income earners. They are more conservative than the 0.25 value chosen by Diamond and Saez (2011). We also compute the entire optimal equilibrium allocation under our benchmark parameterization.⁴³ Our calculations imply that effort variation from the lowest to the highest ranked CEO is

⁴¹ We experimented with a wide variety of smoothing techniques, in all cases the magnitude and behavior of the estimated α_h and the computed optimal tax rates are little changed. Ordering by firm market capitalization and then smoothing to obtain imputed CEO compensation leads to somewhat higher values for the local Pareto coefficients of talent. This, in turn, implies even lower values for optimal tax rates than we obtain.

⁴² Intuitively, the relatively fat empirical right tail for market capitalization implies a relatively rapid increase in S and, hence, price of effective labor across rank. The empirical right tail for CEO income then requires that CEO effective labor (the product of talent and effort) does not increase too quickly with rank. In particular, for the talent Pareto coefficient to remain positive (and talent to decrease with v), effort cannot be too responsive to the return to effort.

⁴³ Full details are in the online Appendix.

of the order of 40 percent. Bandiera et al. (2011) is one of the few studies to explore effort variation among CEOs. They study a sample of Italian CEOs and report variation in hours worked of about 50 percent between the CEOs at the ninetieth and the tenth percentile rank by hours worked. They also confirm a positive relationship between CEO hours and firm productivity. Our benchmark parameterization, thus, generates variation in effort of a similar order of magnitude to the variation in hours found by Bandiera et al. (2011) albeit at the optimum rather than the prevailing (Italian) tax system.^{44,45}

Recall from (29) that under our assumptions, $\mathcal{E}_{\Pi} = \frac{1}{A_{\Pi}} \mathcal{E}$. We recover profits from (49) and estimate a value for $\frac{1}{A_{\Pi}}$ equal to 0.53. Hence, under our benchmark \mathcal{E} choice, $\mathcal{E}_{\Pi} = 0.035$. This small positive elasticity is of a similar order of magnitude to the profit elasticity estimates obtained from simple regressions described in Section II. Overall, our assumptions regarding the effort elasticity are relatively conservative by the standards of the literature, but generate implications for effort variation and the aggregate profit elasticity that are broadly consistent with the data.

Empirical Characterization of the Tail of the Talent Distribution.—The calculated local Pareto coefficients for CEO talent under our benchmark parameterization are plotted in Figure 1. They show a sharp escalation consistent with a thin right tail to the talent distribution. They are drastically different from those for CEO incomes which were stable and consistent with a right Pareto tail to the income distribution with a Pareto coefficient of 2.1. These findings are in line with Terviö (2008) and Gabaix and Landier (2008) who provide corroborating evidence (in models without CEO effort or taxes) that the talent distribution is thin tailed.⁴⁶

C. Results

Substituting our empirical values for α_h and α_S into (43) along with values for \mathcal{E} and χ gives optimal marginal tax rates as function of CEO rank v .⁴⁷ Figure 2 shows these tax rates for the effort elasticity $\mathcal{E} = 1/15$ and for various values of χ . Those for $\chi = 0.6$ can be interpreted as optimal in the absence of a reform of current US profit taxes (assuming statutory rates are paid and with no additional concern for profits beyond their role as a direct source of revenue); those for $\chi = 1$ as optimal under a full reform of CEO income and firm profit taxation. The figure makes apparent that marginal tax rates are falling with χ and for $\chi = 1$ are low across all CEO ranks.

⁴⁴Bandiera et al. (2011) find that a CEO in the ninetieth percentile (in terms of hours worked) works 20 hours (and about 50 percent) more than a CEO in the tenth percentile. They find a positive relationship between CEO hours and a firm's labor productivity. Similar results are found in Bandiera, Prat, and Sadun (2013) looking at Indian manufacturing firms. Finally, Bandiera et al. (2014) looks at a smaller sample of recent CEOs (mostly in smaller firms) in the United States. This study also documents a positive relationship between firm size and hours worked.

⁴⁵To the extent that the optimal tax system is regressive and the actual one linear or progressive over top incomes, it is likely to enhance variation in effort and hours.

⁴⁶Previous versions of the paper formally estimated the tail properties of the talent distribution. For all estimators considered the distribution was determined to be thin tailed and best characterized by a Weibull-like distribution.

⁴⁷Calculation of optimal marginal tax rates as a function of v using (43) does not require calculation of the entire optimal allocation. In the online Appendix we compute the optimal allocation under a benchmark parameterization. We also undertake several counterfactual exercises suggested by a referee that indicate the respective roles of talent and effort in inducing equilibrium CEO income variation.

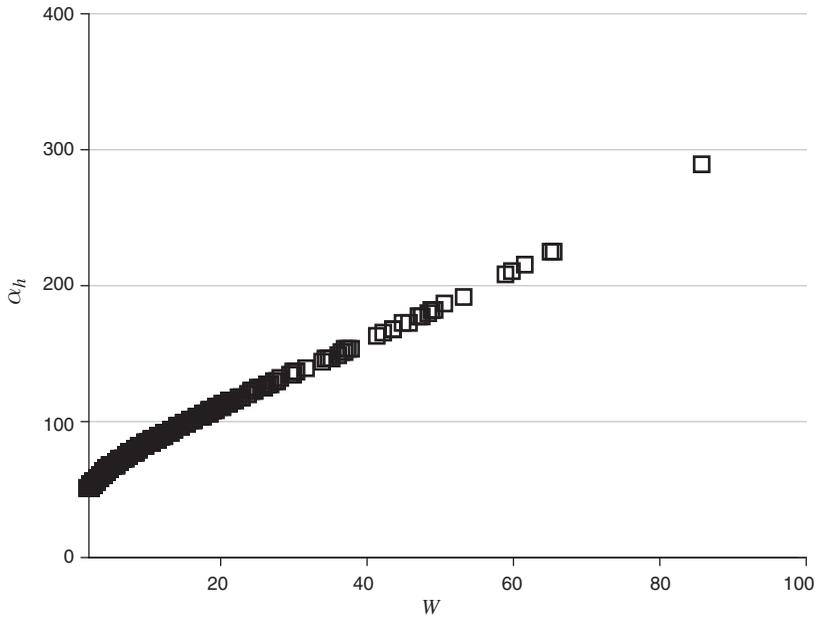


FIGURE 1. ESTIMATES OF THE LOCAL PARETO COEFFICIENT FOR TALENT α_h AS A FUNCTION OF CORRESPONDING CEO INCOME

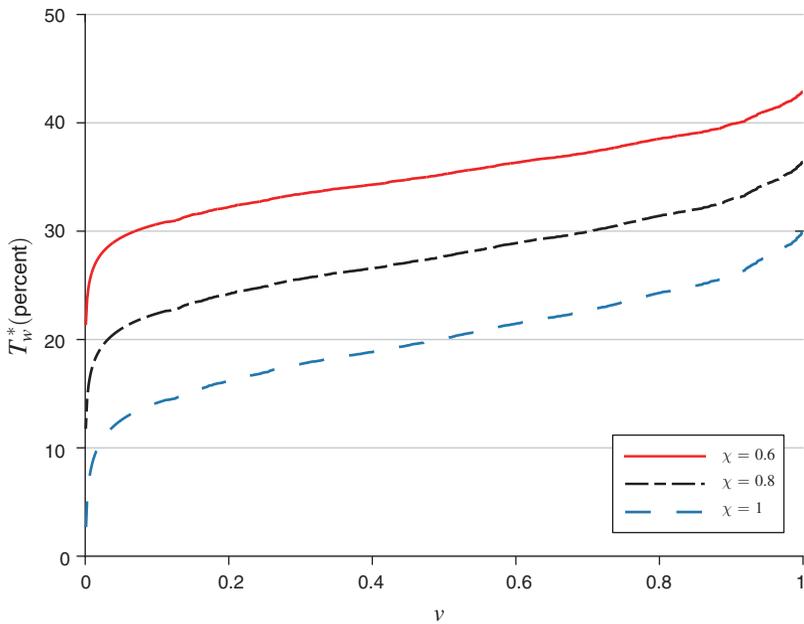


FIGURE 2. MARGINAL TAX RATES AS A FUNCTION OF ν

It is useful to relate these optimal marginal tax rates to equilibrium CEO incomes rather than rank. However, if $\chi = 1$, only the difference between a CEO's income and that of the least talented active CEO, $\Delta w^*(\nu) = w^*(\nu) - w^*(\tilde{\nu})$, is determined in the optimal tax equilibrium. The income $w^*(\tilde{\nu})$ is not determined, because the

TABLE 1—MARGINAL TAX RATES (Percent)

	Δw^* (in millions)		
	30	40	60
$\chi = 0.4$	40.2	39.3	38.4
$\chi = 0.6$	32.2	31.2	30.1
$\chi = 0.8$	24.2	23.0	21.8
$\chi = 1$	16.2	14.9	13.4

Notes: Optimal nonlinear tax rates as function of societal weight χ and income level. Δw^* measured in millions of 2011 US\$.

policymaker is indifferent between the realization of surplus as profit at the smallest active firm or as tax revenues taken from the least talented CEO. The policymaker is constrained only by the requirements that $w^*(\tilde{v}) - T[w^*(\tilde{v})]$ is above the consumption level needed to keep the CEO in the market and that firm profit is nonnegative. For the $\chi < 1$ case, CEO income is only determined up to the inability to place a lump sum tax on the smallest firm's profit. Given this indeterminacy, Table 1 reports optimal marginal tax rates as functions of $\Delta w^*(v) = w^*(v) - w^*(\tilde{v})$.⁴⁸ The table highlights that over large ranges of high (but thinly populated) CEO incomes optimal marginal tax rates are declining in both income and χ . At χ equal to 0.6 marginal tax rates are slightly above 30 percent on high incomes and declining in income, while at $\chi = 1$ they average 15 percent and are declining.

D. Robustness

Focusing on the case $\chi = 0.8$, we now undertake robustness analysis with respect to the parameters λ , θ , and g . Recall that these parameters affect tax rates via their impact on the estimated values of α_h and α_s .⁴⁹ Table 2 reports results. Qualitatively, the responses of tax rates are as expected. For example, a higher value of $\mathcal{E} = 0.1$ (implying $\mathcal{E}_w = 0.153$) implies that optimal tax rates fall below 7 percent. On the other hand, a higher value of θ implies a lower share of CEO effective labor in the determination of firm market capitalization. This in turn lowers the distortionary effect of taxes on firm profits and leads to higher optimal marginal tax rates. Although variation in parameters affects optimal taxes, in all cases considered computed optimal marginal tax rates are much lower than the rates of 70 or 80 percent proposed in the literature.

V. Conclusion

This paper considers the optimal taxation of top earning CEOs. To that end it extends optimal tax theory to an assignment setting in which firms buy CEO labor and some of the surplus generated in production accrues to firm claimants. The classic Diamond-Saez formula continues to prescribe very high marginal tax rates on top CEO incomes of over 70 percent and sometimes 80 percent, but it is only

⁴⁸ $\Delta w^*(\tilde{v})$ is obtained by integrating out (17) using the values of the local Pareto coefficients previously obtained. If $w^*(\tilde{v})$ is normalized to the empirical value of $w(\tilde{v})$, then $w^*(v) = \Delta w^*(v) + \0.5 million.

⁴⁹ Specifically, in equations (53) and (54): $\mathcal{N} := \frac{1}{D\Xi} - \frac{\mathcal{E}}{1 + \mathcal{E}}$, $\mathcal{P} := -\frac{\mathcal{E}}{1 + \mathcal{E}}$, and $\mathcal{M} = 1 - \frac{1}{D\Xi}$.

TABLE 2—MARGINAL TAX RATES (Percent): ROBUSTNESS ANALYSIS

Parameter	Values	Δw^* (in millions)		
		30	40	60
θ	0.2	16.1	15.1	13.8
	0.4	24.2	23.0	21.8
	0.6	36.5	35.4	34
g	0.015	14.0	12.9	11.5
	0.025	24.2	23.0	21.8
	0.035	40.1	39.2	37.8
λ	0.1	32.0	31.0	29.7
	0.5	24.2	23.0	21.8
	2.0	22.5	21.3	20.1
\mathcal{E}	1/10	6.8	5.6	4.2
	1/15	24.2	23.0	21.8
	1/20	36.1	35.0	33.8

Notes: Δw^* measured in millions of 2011 US\$. For each parameter (θ , g , λ) the middle row represents our benchmark calculations as reported in Table 1.

applicable if the policymaker has no social concern for profits. This in turn requires that the planner has no ability to tax profits, no direct social concern for firm claimants, and that depressing profits has no adverse effects on firm entry. In the nonlinear tax setting, a Mirrlees formula (expressed in terms of the CEO effort elasticity and a weighted tail coefficient of the CEO talent distribution) is valid if a comprehensive reform of CEO income and profit taxation is pursued. Our quantitative analysis suggests that the right tail of the CEO talent distribution is thin and that the optimal marginal tax in this case is around 16 percent on incomes above \$30 million or so and lower on higher incomes. If profit taxation was left as is in the United States and only CEO income taxation was reformed (and there was no direct social concern for firm claimants or impact of aggregate profits on firm entry), then the optimal marginal tax rate would be around 32 percent and 40 percent on incomes above \$30 million.

Our paper integrates an assignment model into a normative public economics framework. In this context it has emphasized the impact of CEO income taxes on firm profits and profit tax revenues. However, its broader message is that empirical assessment of the fiscal spillovers associated with top earners is critical for determining the top tax rate. More remains to be done in this direction. First, the limits to profit taxation (and, hence, the extent of the spillover from CEO effort to firm profit tax revenues) remain to be further explored and quantified. Potentially, profit taxes interact with financing frictions, encourage firms to realize profits outside of a tax jurisdiction, and deter creation of the firm asset S in the first place. Further quantitative analysis of CEO income and profit taxation that takes explicit account of these various factors would be valuable. Second, our model abstracts from any impact of CEO effort on the demand for and, hence, wages of (socially) valued workers. Thus, it omits a potential motive for even lower marginal tax rates on top incomes than those reported here.⁵⁰ Analysis of this effect requires embedding the CEO

⁵⁰In the opposite direction, we have omitted pure rent-seeking on the part of CEOs of the sort emphasized by Piketty, Saez, and Stantcheva (2014), a force for higher marginal tax rates.

assignment model into a general equilibrium framework that explicitly incorporates workers.⁵¹ Finally it remains to explore the quantitative implications of assignment for taxation in other top earner settings, such as entertainers, athletes, and entrepreneurs. We leave analysis of these issues for future work.

REFERENCES

- Ales, Laurence, Andres Bellofatto, and Jessie Wang.** 2015. "Taxing Atlas: Executive Compensation, Firm Size and Their Impact on Optimal Top Income Tax Returns." <http://www.andrew.cmu.edu/user/ales/abw-taxingatlas.pdf> (accessed May 20, 2016).
- Ales, Laurence, Musab Kurnaz, and Christopher Sleet.** 2015. "Technical Change, Wage Inequality, and Taxes." *American Economic Review* 105 (10): 3061–101.
- Ales, Laurence, and Christopher Sleet.** 2016. "Taxing Top CEO Incomes: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.20151093>.
- Bakija, Jon, Adam Cole, and Bradley T. Heim.** 2012. "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data." <http://web.williams.edu/Economics/wp/BakijaColeHeimJobsIncomeGrowthTopEarners.pdf> (accessed May 20, 2016).
- Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun.** 2011. "What Do CEOs Do?" Center for Economic and Policy Research Discussion Paper DP8235.
- Bandiera, Oriana, Andrea Prat, and Raffaella Sadun.** 2013. "Managing Firms in an Emerging Economy: Evidence from the Time Use of Indian CEOs." <http://www.hbs.edu/faculty/Pages/item.aspx?num=45037> (accessed May 20, 2016).
- Bandiera, Oriana, Andrea Prat, Raffaella Sadun, and Julie Wulf.** 2014. "Span of Control and Span of Attention." Harvard Business School Working Paper 12-053.
- Diamond, Peter A., and James A. Mirrlees.** 1971. "Optimal Taxation and Public Production II: Tax Rules." *American Economic Review* 61 (3): 261–78.
- Diamond, Peter, and Emmanuel Saez.** 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendations." *Journal of Economic Perspectives* 25 (4): 165–90.
- Edmans, Alex, and Xavier Gabaix.** 2009. "Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories." *European Financial Management* 15 (3): 486–96.
- Frydman, Carola, and Dirk Jenter.** 2010. "CEO Compensation." *Annual Review of Financial Economics* 2 (1): 75–102.
- Frydman, Carola, and Raven S. Molloy.** 2011. "Does Tax Policy Affect Executive Compensation? Evidence from Postwar Tax Reforms." *Journal of Public Economics* 95 (11–12): 1425–37.
- Gabaix, Xavier, and Augustin Landier.** 2008. "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123 (1): 49–100.
- Gayle, George-Levi, Limor Golan, and Robert A. Miller.** 2015. "Promotion, Turnover, and Compensation in the Executive Labor Market." *Econometrica* 83 (6): 2293–369.
- Gayle, George-Levi, and Robert A. Miller.** 2009. "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" *American Economic Review* 99 (5): 1740–69.
- Goolsbee, Austan.** 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy* 108 (2): 352–78.
- Gruber, Jon, and Emmanuel Saez.** 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1): 1–32.
- Lockwood, Benjamin, Charles Nathanson, and E. Glen Weyl.** Forthcoming. "Taxation and the Allocation of Talent." *Journal of Political Economy*.
- Munk, Knud Jorgen.** 1978. "Optimal Taxation and Pure Profit." *Scandinavian Journal of Economics* 80 (1): 1–19.
- Munk, Knud Jorgen.** 1980. "Optimal Taxation with Some Non-Taxable Commodities." *Review of Economic Studies* 47 (4): 755–65.
- Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva.** 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6 (1): 230–71.
- Rosen, Sherwin.** 1982. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics* 13 (2): 311–23.
- Rothschild, Casey, and Florian Scheuer.** 2013. "Redistributive Taxation in the Roy Model." *Quarterly Journal of Economics* 128 (2): 623–68.

⁵¹ See Scheuer and Werning (2015) for qualitative work in this direction.

- Rothschild, Casey, and Florian Scheuer.** 2014. "A Theory of Income Taxation under Multidimensional Skill Heterogeneity." <http://web.stanford.edu/~scheuer/multidim.pdf> (accessed May 20, 2016).
- Rothschild, Casey, and Florian Scheuer.** 2016. "Optimal Taxation with Rent-Seeking." *Review of Economic Studies* 83 (3): 1225–62.
- Saez, Emmanuel.** 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1): 205–29.
- Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz.** 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50 (1): 3–50.
- Scheuer, Florian, and Ivan Werning.** 2015. "The Taxation of Superstars." National Bureau of Economic Research Working Paper 21323.
- Stantcheva, Stefanie.** 2014. "Optimal Income Taxation with Adverse Selection in the Labour Market." *Review of Economic Studies* 81 (3): 1296–329.
- Stiglitz, Joseph E.** 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics* 17 (2): 213–40.
- Stiglitz, Joseph E., and P. Dasgupta.** 1971. "Differential Taxation, Public Goods and Economic Efficiency." *Review of Economic Studies* 38 (114): 151–74.
- Terviö, Marko.** 2008. "The Difference that CEOs Make: An Assignment Model Approach." *American Economic Review* 98 (3): 642–68.
- Zucman, Gabriel.** 2014. "Taxing across Borders: Tracking Personal Wealth and Corporate Profits." *Journal of Economic Perspectives* 28 (4): 121–48.