Submission Number: PET11-11-00134

On optimum corporate income tax

Kevin Hsu
Securities and Future Bureau, Financial Supervisory Commisson of Taiwan

C.c. Yang
Academia Sinica

Abstract

Due to firms being viewed simply as a production set, modern corporations are largely absent from the theory of optimum taxation. This paper addresses optimum corporate income tax by modeling firms as a "nexus of contracts" in the principal-agent framework. Our model involves three parties – workers, employers, and the government, and has elements of both moral hazard and adverse selection. We derive the socially optimal allocation and implement the optimum via the imposition of a corporate income tax. The corporate income tax schedule derived has the feature whereby the higher the productivity of a firm, the lower the marginal tax rate the firm should face. On the other hand, a more stochastic working environment or more risk-averse worker preference has an ambiguous impact on the optimal marginal tax rate in general. We also provide quantitative results on the optimal structure of marginal tax rates through numerical simulations.

Submitted: February 28, 2011.

On optimum corporate income tax

K.W. Kevin Hsu
Securities and Futures Bureau, Financial Supervisory Commission of Taiwan

C.C. Yang*
Institute of Economics, Academia Sinica
Department of Public Finance, National Chengchi University
Department of Public Finance, Feng Chia University

January 2011

Abstract: Due to firms being viewed simply as a production set, modern corporations are largely absent from the theory of optimum taxation. This paper addresses optimum corporate income tax by modeling firms as a "nexus of contracts" in the principal-agent framework. Our model involves three parties – workers, employers, and the government, and has elements of both moral hazard and adverse selection. We derive the socially optimal allocation and implement the optimum via the imposition of a corporate income tax. The corporate income tax schedule derived has the feature whereby the higher the productivity of a firm, the lower the marginal tax rate the firm should face. On the other hand, a more stochastic working environment or more risk-averse worker preference has an ambiguous impact on the optimal marginal tax rate in general. We also provide quantitative results on the optimal structure of marginal tax rates through numerical simulations.

JEL classification: H26, D21, D82

Key words: Corporate income tax; Optimal taxation; Contract theory

---

* Corresponding author. Email: ccyang@ieas.econ.sinica.edu.tw

1. Introduction

Mirrlees (1971, 1974) pioneered the study of optimum personal income tax by incorporating either adverse selection or moral hazard. This line of research has been flourishing since then.[1]

Phelps (1986, p. 674) offered a critical observation on the literature of optimum income taxation in the middle of 1980s:

> *"Thus far, research in this field has been confined to the taxation of personal income: wages, interest, and rent; the economics of business, or company, income taxation has been left untouched. In the now standard models of optimum income taxation there are no company profits and indeed no companies at all, incorporated or unincorporated; these models are extensions of the competitive general equilibrium model of neoclassical theory."*

Although two decades had elapsed since Phelps's observation, Kopczuk and Slemrod (2006, p. 130) still had more or less the same observation:

> *"Firms are, for the most part, absent from the modern theory of optimal taxation. Their disappearance dates from the foundational models developed by Peter A. Diamond and James A. Mirrlees (1971), in which firms are simply mechanical vehicles for combining productive inputs into output in cost-minimizing proportions."*

Kopczuk and Slemrod (2006, p. 130) went on to lament: "The lack of a theoretical framework that features firms impedes rigorous welfare analysis of a number of important policy issues."

Modern corporations are largely absent from the theory of optimum taxation, let alone the optimal taxation of their profits. This paper takes a step toward filling in the void by addressing the issue of optimum corporate income tax.[2]

---

[1] See, for example, Saez (2001) and Low and Maldoom (2004).

[2] Unlike our theme, Kopczuk and Slemrod (2006) addressed the role of firms in facilitating tax administration and enforcement. See also Kleven et al. (2009).

According to Williamson (1981), an important conceptual barrier to an understanding of modern corporations inherent in the neoclassical theory of the firm is that the theory is devoid of interesting hierarchical features. In this paper we depart from the neoclassical theory and adopt a modern approach in which firms possess a hierarchical feature. More specifically, we follow the economics of contract to view firms as a "nexus of contracts" in the principal-agent framework.[3] As a first step toward an optimum corporate income tax, we model firms in the simplest manner possible: we let a firm consist of a risk-neutral principal (employer) and a risk-averse agent (worker), and apply the classical agency theory as depicted in Gibbons (1997) to their contractual relationship.[4]

Mirrlees's seminal articles of 1971 and 1974 studied optimum personal income tax with adverse selection and moral hazard, respectively. In the setting of adverse selection, individuals are assumed to be heterogeneous in productivity and observable incomes depend *deterministically* on their productivity, which is private information to individuals. In the setting of moral hazard, by contrast, individuals are assumed to be homogeneous and observable incomes depend *randomly* on their effort, which is neither observable nor verifiable. Our model involves elements of both moral hazard and adverse selection in that the effort of the risk-averse worker is neither observable nor verifiable in a stochastic environment, and that a firm's productivity is its private information.

Both the government and the firm try to influence the actions of the worker in our framework, although the latter's influence is direct (via employment contract) while the former's is indirect (via corporate income tax). We derive the socially optimal allocation and implement the optimum via the imposition of a corporate income tax. The corporate

---

[3] See Laffont and Martimort (2002), Bolton and Dewatripont (2005), and Salanie (2005) for surveys of the literature on the economics of contract.

[4] Note that only corporations, and neither proprietorships nor partnerships, are subject to corporate income tax in the United States and some other countries. The principal-agent framework may not be the appropriate model in the case of proprietorships or partnerships.

income tax schedule derived has the feature whereby the higher the productivity of a firm, the lower the marginal tax rate the firm should face. This *analytically* derived *monotonic* feature is in stark contrast to the *numerically* derived *non-monotonic* feature of marginal tax rates in the optimum personal income tax. On the other hand, a more stochastic working environment or more risk-averse worker preference has an ambiguous impact on the optimal marginal tax rate in general. This ambiguous feature also differs from the standard result in the pure moral hazard model, in which a higher risk induces a higher demand for social insurance and so gives rise to a higher optimal marginal tax rate.

The rest of the paper is organized as follows. Section 2 introduces our model. Section 3 derives the socially optimal allocation. Section 4 implements the social optimum via the imposition of a corporate income tax. Section 5 reports important properties of the optimal corporate income tax derived. Section 6 provides numerical simulations on the optimal structure of marginal tax rates and Section 7 concludes.

2. Model

Our model involves three parties: workers, employers, and the government. There are two-tier relations: (i) a worker and an employer form a firm or corporation, and (ii) the government imposes corporate income tax on the firm's profit. Workers are assumed to be homogenous and, as a first step toward an optimum corporate income tax, we abstract from the issue of personal income tax by normalizing the workers' reservation utility to zero. We first address the firm and then the government.

2.1. Firm

We follow the classical agency theory as depicted in Gibbons (1997) to model a firm. The firm consists of a risk-neutral principal (employer) and a risk-averse agent (worker). The worker supplies effort e to produce output of value y with $y = \beta + e + \varepsilon$, where $\beta \geq 0$ is a parameter representing the firm's productivity, and ε is a random variable normally distributed with mean 0 and variance $\sigma^2 > 0$. A higher or lower β may have to do with the employer's ability of running the firm (not modeled explicitly). Worker effort e is neither observable nor verifiable and, therefore, employment contracts cannot be conditional on the choice of worker effort. However, the realization of output y is assumed to be publicly observable and so employment contracts can be conditional on the observable but noisy signal of worker effort, namely, the realized output y.

The employer owns the output produced but shares it with the worker by paying a remuneration r that is linearly contingent on the realized output y with $r = a + by$, where a is the compensation unrelated to worker performance such as salary, and b is the performance-related compensation such as a bonus.[5]

The worker's utility function is given by exponential utility $u(r,e) = -\exp\{-\eta[r - \psi(e)]\}$, where $\eta > 0$ is the coefficient of absolute risk aversion, and $\psi(e)$ stands for the disutility of worker effort with $\psi'(e) > 0$, $\psi''(e) > 0$, $\psi'''(e) \geq 0$ and $\psi''''(e) \geq 0$ for any $e > 0$. An example of $\psi(e)$ satisfying the properties specified is the quadratic cost function $\psi(e) = (1/2)e^2$. It can be shown that the worker's maximizing expected utility is equivalent to maximizing the following certainty equivalent[6]

$$v = a + b(\beta + e) - \psi(e) - \frac{\eta}{2} b^2 \sigma^2 . \qquad (1)$$

The derivation of (1) is built on exponential utility $u(r,e) = -\exp\{-\eta[r - \psi(e)]\}$. However, it can be justified by an approximation argument for a more general utility function.

Given a and b, the worker's maximization of (1) with respect to e yields

$$b = \psi'(e) , \qquad (2)$$

---

[5] For a justification of linear contracts, see Holmstrom and Milgrom (1987). For evidence broadly consistent with the basic tenet of the classical agency theory, see Gibbons (1997).

[6] See Bolton and Dewatripont (2005, Section 4.2).

which gives a positive impact of b upon worker effort e since $\psi''(e) > 0$ by assumption. Substituting (2) for b in (1) and normalizing the worker's reservation utility to zero yields the firm's payment to the worker

$$r(e) = \psi(e) + \frac{\eta}{2}[\psi'(e)]^2 \sigma^2 + \psi'(e)\varepsilon \,. \tag{3}$$

The employer chooses b and hence e via (2) to maximize the post-tax expected profit

$$\pi \equiv E[y - r(e) - T] = (\beta + e) - E[r(e)] - E(T)\,, \tag{4}$$

where E denotes the expectation operator with respect to ε, and T the corporate income tax imposed by the government. We describe T later.

2.2. Government

A firm's productivity β is its private information, unknown to the government. Nevertheless, the government knows that β is distributed according to the distribution $F(\beta)$ on the range $[\underline{\beta}, \overline{\beta}]$ with the density $f(\beta) > 0$ for all $\beta \in [\underline{\beta}, \overline{\beta}]$. We assume the monotone hazard rate: $d\{[1 - F(\beta)]/f(\beta)\}/d\beta \le 0$. This assumption is commonly imposed in the incentive literature (for example, Laffont and Tirole, 1993, p. 66; Besley and Coate, 1995; Martimort and Moreira, 2010) and is satisfied by many distributions.[7] As interpreted by Laffont and Tirole (1993, pp. 66-67), this is basically a decreasing returns assumption.

From the Revelation Principle, the government can ask the firm to reveal its true productivity directly. Moreover, the government can without loss of generality restrict attention to the firm's truthful announcement so that in equilibrium $\hat{\beta} = \beta$ maximizes the type-β firm's post-tax expected profit.

By assumption, the realization of output y is publicly observable and so it is observable by the government as well as the firm. As to the realization of worker

[7] See Bagnoli and Bergstrom (2005) for the detail.

remuneration r, it may not be observable by the government. This unobservable restriction may cause a deviation between the economic profit earned by the firm and the accounting profit imposed according to tax codes (more elaborations later). Following the idea of Laffont and Tirole (1986) and others, we let the corporate income tax schedule imposed be a menu, depending on the firm's type announcement $\hat{\beta}$ and the publicly observable variable y, that is, $T = \Gamma(\hat{\beta}, y)$.[8] The form that $\Gamma(\hat{\beta}, y)$ will take at the optimum is the central focus of this paper.

Since worker utility is normalized to zero, the government's objective is simply to maximize the firm's post-tax expected profit, given that the government must collect a fixed amount of tax revenue from the firm. However, taxation must respect the firm's incentive compatibility (IC) and individual rationality (IR) constraints. In what follows, we derive these two types of constraints in detail.

2.2.1. IC constraint

In their celebrated work involving both moral hazard and adverse selection, Laffont and Tirole (1986, hereafter LT) considered a regulation problem in which the regulated agent is a firm, whose cost is random but observable by the regulator. The firm's expected payoff is given by $Et - \phi(e)$, where t is the net-of-cost transfer paid by the regulator to the firm, and $\phi(e)$ is the disutility of the firm effort with $\phi'(e) > 0$, $\phi''(e) > 0$ and $\phi'''(e) \geq 0$ for any $e > 0$.

In our work involving both moral hazard and adverse selection, we consider a taxation problem in which the taxed agent is a firm, whose output is random but observable by the government. The firm's post-tax expected profit is given by $E(y - T) - \gamma(e)$, where $\gamma(e) \equiv E[r(e)]$, the firm's expected payment to the worker. From (3)

[8] For the value of incorporating the firm's type announcement into the tax schedule, see Melumad and Reichelstein (1989).

$$\gamma(e) = \psi(e) + \frac{\eta}{2}[\psi'(e)]^2 \sigma^2, \tag{5}$$

which consists of two components: the compensation for the disutility that the worker suffers due to effort (the first term), and the compensation for the risk that the worker bears (the second term). By our assumption about $\psi(e)$, we have $\gamma'(e) > 0$, $\gamma''(e) > 0$ and $\gamma'''(e) \geq 0$ for any $e > 0$.

The above analogy suggests that the method of deriving the IC constraint employed by LT in their regulation framework seems applicable to our taxation framework. This turns out to be true. However, unlike LT's single-tier principal-agent model, we consider a two-tier framework in which the government "contracts" with the employer who then contracts with the worker. While all parties are risk neutral in LT, workers are risk averse in our model.

Suppose that a β firm announces $\hat{\beta}$ and makes effort $\hat{e}(\hat{\beta}\big|\beta) \equiv e(\hat{\beta}) + \hat{\beta} - \beta$. Following LT, the set of $[\hat{\beta}, \hat{e}(\hat{\beta}\big|\beta)]$ will be called the "concealment set" for the type-β firm. This set clearly includes the element $[\beta, e(\beta)]$ when $\hat{\beta} = \beta$. We make the same assumption as in LT that deviations in the concealment set are the only possible deviations. We focus on this set since, similar to LT and as will be shown later, the firm will not be induced to deviate outside the concealment set at the optimum.

Ruling out the firm's deviations of $\hat{\beta}$ from β in the concealment set is equivalent to requiring

$$\hat{\beta} = \beta \text{ maximizes } \pi(\hat{\beta}\big|\beta) = \beta + \hat{e}(\hat{\beta}\big|\beta) - E[T(\hat{\beta}\big|\beta)] - \gamma[\hat{e}(\hat{\beta}\big|\beta)], \tag{6}$$

where $T(\hat{\beta}\big|\beta) = \Gamma[\hat{\beta}, \beta + \hat{e}(\hat{\beta}\big|\beta) + \varepsilon]$. This leads to the first-order condition

$$E[\dot{T}(\hat{\beta}\big|\beta)] = \{1 - \gamma'[\hat{e}(\hat{\beta}\big|\beta)]\} \cdot \dot{\hat{e}}(\hat{\beta}\big|\beta), \tag{7}$$

where a dot denotes a derivative with respect to $\hat{\beta}$. Utilizing the definition of $\hat{e}(\hat{\beta}\big|\beta)$ and truth telling $\hat{\beta} = \beta$, (7) yields

$$E[\dot{T}(\beta)] = \{1 - \gamma'[e(\beta)]\} \cdot [\dot{e}(\beta) + 1]. \tag{8}$$

Using $\pi(\beta) = \{\beta + e(\beta) - E[T(\beta)]\} - \gamma[e(\beta)]$, we have

$$\dot{\pi}(\beta)=\{1+\dot{e}(\beta)-E[\dot{T}(\beta)]\}-\gamma'[e(\beta)]\dot{e}(\beta)\,. \tag{9}$$

Putting (8)-(9) together, the first-order condition (7) is equivalent to

$$\dot{\pi}(\beta)=\gamma'[e(\beta)]\,. \tag{10}$$

Utilizing (7)-(8), we show in the Appendix that the local second-order condition for (6) is

$$\dot{e}(\beta)\geq -1\,. \tag{11}$$

Following LT's Appendix B, one can show that the local second-order condition (11) implies the global one.

2.2.2. IR constraint

A firm will cease to exist if its profit is expected to be negative. Thus, the IR constraint is

$$\pi(\beta)\geq 0 \text{ for all } \beta\,. \tag{12}$$

Since $\dot{\pi}(\beta)>0$ according to (5) and (10), (12) can reduce to

$$\pi(\underline{\beta})\geq 0\,. \tag{13}$$

2.2.3. Government problem

To sum up, the government's tax problem is to maximize the objective:

$$\int_{\underline{\beta}}^{\bar{\beta}}\pi(\beta)f(\beta)d\beta$$

subject to the IC constraint (10)-(11), the IR constraint (13), and to collect a fixed amount of tax revenue R, that is,[9]

$$\int_{\underline{\beta}}^{\bar{\beta}}E[T(\beta)]f(\beta)d\beta=\int_{\underline{\beta}}^{\bar{\beta}}\{\beta+e(\beta)-\gamma[e(\beta)]-\pi(\beta)\}f(\beta)d\beta\geq R\,, \tag{14}$$

where the first equality has utilized the definition of $\pi(\beta)$. Since π is decreasing in *E(T)*, it is clear that (14) will be binding at the optimum.

[9] Typically, higher moments of revenue distribution other than the first one are ignored in the formulation of the government budget constraint. This may be justifiable by assuming a "risk-neutral" government; see Yang (1993) for the detail.

Notice that the government's objective function above takes the "utilitarian" form, without weighing different firms differently. In the case where the design of an optimum personal income tax is involved, it is typically assumed that the government possesses the motive of redistribution from the rich (high productivity) to the poor (low productivity). In the case where the design of an optimum corporate income tax is involved, however, it seems unclear at least a priori whether the government should favor firms with high productivity or those with low productivity. As a first approximation, we simply let the government's objective take the "utilitarian" form.

2.3. Timing

The timing of the game is as follows. First, the firm type is realized. Second, the government announces the imposition of a corporate income tax. Third, the employer offers an employment contract to the worker. Fourth, the worker puts forth his or her effort if the contract is accepted. Finally, the output is realized; the worker is paid according to the employment contract offered and the firm pays the corporate income tax imposed.

3. Social optimum

Both the government and the firm try to influence the actions of the worker in our framework, although the latter's influence is directly exerted via the employment contract while the former's is indirectly exerted via the corporate income tax. It is convenient to carry out the analysis in two steps. First, this section derives the socially optimal levels of worker effort for different types of firms. Then, the next section addresses the implementation of the social optimum via the imposition of a corporate income tax.

We ignore the second-order condition (11) for the moment, and verify later that the resulting solution from the relaxed problem satisfies (11) and hence solves the unrelaxed problem.

The Hamiltonian for the government problem is

$$H = \{\pi + \lambda[\beta + e - \gamma(e) - \pi]\} f + \mu(\beta)\gamma'(e)\,, \tag{15}$$

where λ is the Lagrange multiplier associated with the government budget (14), and $\mu(\beta)$ is the co-state variable. We employ π as the state variable and e as the control variable. Applying the Pontryagin principle to (15) yields necessary conditions for an interior optimum

$$\frac{\partial H}{\partial e} = 0 = \lambda[1-\gamma'(e)] f + \mu(\beta)\gamma''(e)\,, \tag{16-1}$$

$$\dot{\mu} = -\frac{\partial H}{\partial \pi} = (\lambda - 1) f\,, \tag{16-2}$$

$$\dot{\pi}(\beta) = \gamma'[e(\beta)]\,. \tag{16-3}$$

Furthermore, $\pi(\overline{\beta})$ is free so that

$$\mu(\overline{\beta}) = 0\,, \tag{16-4}$$

and the transversality conditions for (13) are[10]

$$\mu(\underline{\beta}) \le 0\,,\ \pi(\underline{\beta}) \ge 0\,,\ \mu(\underline{\beta})\pi(\underline{\beta}) = 0\,. \tag{16-5}$$

Integrating (16-2) and using (16-4) gives

$$\mu(\beta) = (1-\lambda)\int_{\beta}^{\overline{\beta}} f(\delta)d\delta = (1-\lambda)[1-F(\beta)]\,. \tag{17}$$

On the basis of (16-5), we consider two possibilities: (i) $\pi(\underline{\beta}) > 0$, and (ii) $\pi(\underline{\beta}) = 0$.

3.1. $\pi(\underline{\beta}) > 0$

$\pi(\underline{\beta}) > 0$ implies that $\mu(\underline{\beta}) = 0$ according to (16-5). From (17), we then obtain

$$\lambda = 1\,. \tag{18}$$

This implies $\mu(\beta) = 0$ from (17). Thus, (16-1) yields

$$\gamma'[e(\beta)] = 1\,, \tag{19}$$

which, by (2) and the definition of $\gamma(e)$ in (5), gives rise to

$$b = \psi'(e) = \frac{1}{1+\eta\sigma^2\psi''(e)}\,. \tag{20}$$

---

10 See Leonard and Long (1992, chapter 7).

Thus, the optimal b chosen by the firm is strictly between zero (providing full insurance) and one (providing full incentives), and it goes down as η or σ^2 goes up. Equation (20) characterizes the optimal tradeoff between insurance and incentives as in the classical agency theory; see Gibbons (1995). The imposition of the corporate income tax should respect this tradeoff in the case of $\pi(\underline{\beta}) > 0$. Note in particular that the socially optimal e implicitly defined by (20) is independent of firm type β.

The multiplier λ in (15) is the shadow price of the government budget. It is known as the "marginal cost of public funds" in the tax literature, since it represents the social cost of raising an additional dollar of revenue from the private sector. LT assume that the social planner can raise revenue from the private sector only through distortionary taxes. As a result, they impose $\lambda > 1$ exogenously.[11] This restriction need not hold in our context.

Substituting (19) in (16-3) yields $\dot{\pi}(\beta) = 1$, which leads to

$$\pi(\beta) = \pi(\underline{\beta}) + (\beta - \underline{\beta}) . \tag{21}$$

This result indicates that the corporate income tax imposed should let the firm capture all the gains from its higher productivity.

When $\lambda = 1$, taxation must entail no distortion by the very definition of λ. This explains why $\mu(\beta)$ in (15) is identically zero so that the IC constraint is de facto not binding in this case.[12] This also explains why the classical tradeoff (20) remains true so that the imposition of the corporate income tax should entail no additional distortion in this case.

3.2. $\pi(\underline{\beta}) = 0$

Substituting (17) for $\mu(\beta)$ in (16-1) gives

---

[11] LT used the notation $1 + \lambda$ to denote the marginal cost of public funds in their paper.

[12] Assuming that utility is quasi-linear and that social welfare is Benthamite (additive), the result of $\lambda = 1$ will also arise in the case of the personal income tax at the optimum; see Salanie (2003, Section 4.2.3).

$$\gamma'[e(\beta)] = 1 - (1 - \frac{1}{\lambda})[\frac{1-F(\beta)}{f(\beta)}]\gamma''[e(\beta)]. \tag{22}$$

Since $\mu(\underline{\beta}) < 0$ in this case, we obtain $\lambda > 1$ from (17). (22) then implies that $\gamma'[e(\beta)] < 1$ at the optimum as long as $\beta < \overline{\beta}$.

Let $e^*(\beta)$ denote the solution to (22) and e^{**} the solution to (19). Since (i) $\gamma'(e^{**}) = 1$ but $\gamma'[e^*(\beta)] < 1$ if $\beta < \overline{\beta}$, and (ii) $\gamma(e)$ is increasing and convex in e according to (5), we have $e^*(\beta) < e^{**}$ as long as $\beta < \overline{\beta}$. That is, the corporate income tax imposed should distort the optimal tradeoff between insurance and incentives in the classical agency theory as given by (20) if $\beta < \overline{\beta}$ and, in particular, it should distort the worker effort downward.

To achieve this depression of worker effort, the incentives that the firm offers to its worker must become smaller as compared with (20). By (2) and the definition of $\gamma[e(\beta)]$ in (5), (22) gives rise to

$$b = \psi'(e) = \frac{1 - (1 - \frac{1}{\lambda})(\frac{1-F}{f})\gamma''(e)}{1 + \eta\sigma^2\psi''(e)}, \tag{23}$$

which is indeed smaller than the b given in (20) as long as $\beta < \overline{\beta}$.

The further depression or downward distortion of worker effort obviously causes an additional efficiency loss, but it deters a high-productivity firm from masquerading as a low-productivity firm and enables the government to extract the information rent enjoyed by the high-productivity firm. This tradeoff between the extraction of information rent and the loss of efficiency is the fundamental tradeoff in the adverse selection model; see Laffont and Martimort (2002, chapter 2). In fact, (22) represents a balance at margin between an additional efficiency loss (i.e., $\lambda\{1 - \gamma'[e(\beta)]\}$ in firm number $f(\beta)d\beta$) and an extra rent extraction (i.e., $(\lambda - 1)\gamma''[e(\beta)]d\beta$ in firm number $1 - F(\beta)$).

From (22), we have

$$[\gamma'' + (1 - \frac{1}{\lambda})(\frac{1-F}{f})\gamma''']\dot{e} = -(1 - \frac{1}{\lambda})\gamma''\frac{d[(1-F)/f]}{d\beta}, \tag{24}$$

which by the assumption of the monotone hazard rate leads to $\dot{e}^*(\beta) \geq 0$. This result implies that the extent of the downward distortion of $e^*(\beta)$ from e^{**} should become smaller as a firm's productivity increases. Indeed, since $F(\overline{\beta}) = 1$ if $\beta = \overline{\beta}$, (22) will reduce to $\gamma'[e(\overline{\beta})] = 1$ so that $e^*(\overline{\beta}) = e^{**}$. This is a standard result in the adverse selection model – no firm will mimic the highest type and hence there should be no distortion for this type.

As far as the "contract" between the government and the firm is concerned, our model is a pure adverse selection problem. However, since it involves the risk-averse worker as well, the optimal taxation must take the moral hazard issue into consideration, as will be seen.

Corresponding to $\dot{e}^*(\beta) \geq 0$, we obtain from (23) that $\dot{b}(\beta) = \psi'' \cdot \dot{e}^*(\beta) \geq 0$; that is, the higher the productivity of a firm, the larger the incentives the firm will offer to its worker at the social optimum. This is in sharp contrast to the case of $\pi(\underline{\beta}) > 0$, in which $e = e^{**}$ for all β; that is, the incentives that the firm offers to its worker are independent of its productivity at the social optimum.

Substituting (22) in (16-3) and using $\pi(\underline{\beta}) = 0$ gives

$$\pi(\beta) = (\beta - \underline{\beta}) - \int_{\underline{\beta}}^{\beta} (1 - \frac{1}{\lambda})[\frac{1 - F(\delta)}{f(\delta)}]\gamma''[e(\delta)]d\delta \,. \tag{25-1}$$

This result differs from (21), indicating that, except for $\beta = \overline{\beta}$, the corporate income tax imposed should not let the firm capture all the gains from its higher productivity. Using $\pi(\underline{\beta}) = 0$, (16-3) also gives

$$\pi(\beta) = \int_{\underline{\beta}}^{\beta} \gamma'[e(\delta)]d\delta \,. \tag{25-2}$$

This formula will be useful later.

3.3. Second-order conditions

If $e = e^{**}$, then $\dot{e}(\beta) = 0$ since $e(\beta) = e^{**}$ is the solution to (19). On the other hand, if $e = e^*(\beta)$, then we have $\dot{e}(\beta) \geq 0$ from (24). We thus prove that the second-order condition (11) holds in both situations. This justifies our formulation of the Hamiltonian (15) without the imposition of (11) in the first place. Note that if the monotone hazard rate is strict, i.e., $d\{[1-F(\beta)]/f(\beta)\}/d\beta<0$ (say, F follows a uniform distribution), then we obtain from (24) that $\dot{e}(\beta)>0$.

Finally, we have from (16)

$$\frac{\partial^2 H}{\partial e^2}=-\lambda f\gamma''+(1-\lambda)(1-F)\gamma'''<0,$$

$$\frac{\partial^2 H}{\partial \pi^2}=0,$$

$$\frac{\partial^2 H}{\partial e\partial \pi}=0.$$

This concavity of the Hamiltonian shows that the result characterized by (16) is the solution to the government's tax problem.

4. Implementation

This section addresses the implementation of the social optimum via the imposition of a corporate income tax.

To implement the social optimum, we must find a menu of corporate income taxes $\Gamma(\hat{\beta},y)$, a function of $\hat{\beta}$ (a firm's type announcement) and output y (an observable), such that:

(i) if $\pi(\underline{\beta})>0$, it is optimal for the type-β firm to announce $\hat{\beta}=\beta$ to the government and choose the bonus b given by (20) so that $e=e^{**}$;

(ii) if $\pi(\underline{\beta})=0$, it is optimal for the type-β firm to announce $\hat{\beta}=\beta$ to the government and choose the bonus b given by (23) so that $e=e^{*}(\beta)$.

Moreover, $\int_{\underline{\beta}}^{\overline{\beta}} E[\Gamma(\hat{\beta},y)]f(\beta)d\beta=R$; that is, the government budget is met.

Note that the so-called "announce $\hat{\beta}=\beta$ to the government" should not be taken literally; it simply means that a type chooses the tax allocation in the menu intended for the type. Below we consider the two possibilities, $\pi(\underline{\beta})>0$ and $\pi(\underline{\beta})=0$, respectively.

4.1. $\pi(\underline{\beta}) > 0$ *with* $e = e^{**}$

Consider a uniform lump-sum tax that meets the government revenue constraint with $\pi(\underline{\beta}) > 0$. With the imposition of the lump-sum tax, the employer's maximization of (4) obviously gives rise to (19), which in turn leads to $e = e^{**}$. Since the lump-sum tax is uniform across firm types, it is plain that $\hat{\beta} = \beta$ will hold weakly in that a type-β firm is indifferent between $\hat{\beta} = \beta$ and $\hat{\beta} \neq \beta$.

4.2. $\pi(\underline{\beta}) = 0$ *with* $e = e^{*}(\beta)$

It is known that the optimal nonlinear personal income tax can be implemented through a menu of linear income taxes. That is, persons with different productivities (types) choose to face different "productivity-specific" linear income taxes intended for them.[13] This is also true for the optimal corporate income tax in the case where $\pi(\underline{\beta}) = 0$, as we show below.

From (25-2), let $\pi^{*}(\beta) = \int_{\underline{\beta}}^{\beta} \gamma'[e^{*}(\delta)]d\delta$. Consider the following menu of corporate income taxes that are linear in the observable output *y*:

$$\Gamma(\hat{\beta}, y) = s^{*}(\hat{\beta}) + \tau^{*}(\hat{\beta})(y - y^{*}), \tag{26}$$

where

$$\tau^{*}(\hat{\beta}) = 1 - \gamma'[e^{*}(\hat{\beta})], \tag{26-1}$$

$$y^{*} = \hat{\beta} + e^{*}(\hat{\beta}), \tag{26-2}$$

$$s^{*}(\hat{\beta}) = y^{*} - \gamma[e^{*}(\hat{\beta})] - \pi^{*}(\hat{\beta}). \tag{26-3}$$

We postpone the interpretation of (26) until the next section.

Facing the menu of corporate income taxes given by (26), a β firm chooses $\hat{\beta}$ and *e* to solve

$$\max\ E\{\beta + e + \varepsilon - \Gamma(\hat{\beta}, y)\} - \gamma(e). \tag{27}$$

---

[13] See, for example, Saez (2001).

Optimization with respect to e gives

$$\tau^*(\hat{\beta}) = 1 - \gamma'(e), \tag{28}$$

which, by the definition of $\tau^*(\hat{\beta})$ in (26-1), clearly leads to $e = e^*(\hat{\beta})$. Substituting $e = e^*(\hat{\beta})$ in (27) and utilizing $\dot{\pi}^*(\hat{\beta}) = \gamma'[e^*(\hat{\beta})]$, optimization with respect to $\hat{\beta}$ yields

$$[\beta + e^*(\hat{\beta}) - y^*]\gamma''[e^*(\hat{\beta})]\dot{e}^*(\hat{\beta}) = 0, \tag{29}$$

which gives rise to $\hat{\beta} = \beta$.

Next, using (29), the second-order condition for (27) requires

$$-\gamma''[e^*(\hat{\beta})]\dot{e}^*(\hat{\beta}) + (\beta - \hat{\beta})\gamma''[e^*(\hat{\beta})]\ddot{e}^*(\hat{\beta}) + (\beta - \hat{\beta})\gamma'''[e^*(\hat{\beta})][\dot{e}^*(\hat{\beta})]^2 \leq 0,$$

where $\ddot{e}^*(\hat{\beta})$ denotes the derivative $\dot{e}^*(\hat{\beta})$ with respect to $\hat{\beta}$. With $\hat{\beta} = \beta$, the above inequality reduces to $\dot{e}^*(\beta) \geq 0$, which is a property of $e^*(\beta)$. Note that $\int_{\underline{\beta}}^{\overline{\beta}} E[\Gamma(\hat{\beta}, y)] f(\beta) d\beta = \int_{\underline{\beta}}^{\overline{\beta}} E[T(\beta)] f(\beta) d\beta = R$ by (14) and (26). This proves the implementation of the optimum $\{\hat{\beta} = \beta, e = e^*(\beta)\}$ by the corporate income tax (26).

Finally, note that the firm's choice of $\hat{\beta}$ and e in (27) is not restricted to being in the concealment set. Thus, similar to LT, we do implement the optimal solution and make deviations outside the concealment set unprofitable for the firm.

4.3. Remark

A socially optimal allocation may be implementable by various, not uniquely determined, tax systems in general (Golosov et al. 2006). For example, in the case of $\pi(\underline{\beta}) > 0$, the socially optimal allocation $e = e^{**}$ can also be implemented by a positive but less than 100% proportional tax rate levied on the expected economic profit $Ey - \gamma(e)$.

Similarly, in the case of $\pi(\underline{\beta}) = 0$, there may exist forms of corporate income taxes other than (26) to implement the socially optimal allocation $e = e^*(\beta)$. Nevertheless, to implement the social optimum in which the extent of the distortion of $e^*(\beta)$ from e^{**} is decreasing in β and reaches no distortion with $e^*(\overline{\beta}) = e^{**}$, it seems that the very feature of the corporate income tax schedule derived (i.e., the marginal tax rate is decreasing in firm type and reaches zero for the highest type) has to hold for other forms of corporate income taxes as well. Comparing (22) with (19), we see that the optimal marginal tax rate as defined by (26-1) is precisely equal to the wedge of the further downward distortion of worker effort from e^{**} as required by the social optimum.

5. Optimum corporate income tax

Depending on whether $\pi(\underline{\beta}) > 0$ or $\pi(\underline{\beta}) = 0$, the socially optimal allocation is different and, as a result, the corporate income tax schedule to implement the social optimum is also different. This is what we have shown in Sections 3 and 4. A question naturally arises: Which situation, $\pi(\underline{\beta}) > 0$ or $\pi(\underline{\beta}) = 0$, is more relevant in the real world? Our model assumes that firm productivity β is exogenously given on the range $[\underline{\beta}, \overline{\beta}]$. However, once we allow for free entry of firms, it is plausible that the lowest productivity $\underline{\beta}$ will be pinned down by the free-entry condition, namely, $\pi(\underline{\beta}) = 0$. Thus, unless there is regulation of entry, it seems that $\pi(\underline{\beta}) = 0$ rather than $\pi(\underline{\beta}) > 0$ is more relevant in the real world.

In this section we focus on the situation where $\pi(\underline{\beta}) = 0$ and report important properties of the corresponding optimal corporate income tax derived in the previous section.

With $\hat{\beta} = \beta$, the optimal corporate income tax (26) can be expressed as a menu of "productivity-specific" linear income taxes

$$\Gamma(\beta, y) = g^*(\beta) + \tau^*(\beta) y, \tag{30}$$

where $\tau^*(\beta)$ is the marginal tax rate and $g^*(\beta)$ the lump-sum grant with

$$\tau^*(\beta) = 1 - \gamma'[e^*(\beta)], \tag{30-1}$$

$$g^*(\beta) = [1 - \tau^*(\beta)][\beta + e^*(\beta)] - \gamma[e^*(\beta)] - \pi^*(\beta). \tag{30-2}$$

From (22), we immediately see that $0 \le \tau^*(\beta) < 1$. Thus, similar to the optimum personal income tax à la Mirrlees (1971), the marginal tax rates for the optimum corporate income tax are restricted to being non-negative and less than 100%.

On the basis of (30), we next address two issues: (i) the shape of the optimal corporate income tax schedule, and (ii) the impact of increasing η or σ^2 on the shape. To be consistent with the numerical simulations later, we focus on the situation where the monotone hazard rate is strict so that $\dot{e}^*(\beta) > 0$.

5.1. Shape of optimal tax schedule

From (30), we have

$$\frac{d\tau^*}{d\beta} = -\gamma''[e^*(\beta)]\dot{e}^* < 0, \tag{31-1}$$

$$\frac{dg^*}{d\tau^*} = -[\beta + e^*(\beta)] < 0, \tag{31-2}$$

$$\frac{d^2 g^*}{d(\tau^*)^2} = \frac{d(dg^*/d\tau^*)/d\beta}{d\tau^*/d\beta} = \frac{-(1+\dot{e}^*)}{d\tau^*/d\beta} > 0. \tag{31-3}$$

(31-2) and (31-3) together show that the lump-sum grant g^* is strictly decreasing and convex with respect to the marginal tax rate τ^*. A firm chooses $(g^*(\beta), \tau^*(\beta))$ according to its productivity β: the higher the β, the lower the marginal tax rate τ^* (see (31-1)) but the higher the lump-sum grant g^* (see (31-2)).

(31-1) shows that the higher the productivity of a firm, the lower the marginal tax rate the firm should face. This feature of $\tau^*(\beta)$ is to implement the social optimum that the downward distortion of $e^*(\beta)$ from e^{**} should be smaller if a firm's productivity β is higher. Indeed, to implement $e^*(\overline{\beta}) = e^{**}$ (no further distortion for the highest type), $\tau^*(\overline{\beta}) = 0$ must hold according to (22) and (26-1).

A central issue in the literature on optimal personal income taxation is the shape of the tax schedule, namely, whether optimal marginal tax rates rise or fall with income in different income ranges. However, few precise analytical results are obtained, except that the marginal tax rate should equal zero at the top.[14] Through numerical simulations, it is found that the optimal marginal tax rate for personal income tax is typically non-monotonic with respect to individual types.[15] In contrast to this non-monotonic numerical finding, (31-1) analytically shows that the optimal marginal tax rate for corporate income tax is monotonically decreasing with respect to firm types.

The marginal tax rate reaching zero at the top characterizes both personal and corporate income taxation at the optimum. This result is not surprising in view of the fact that the design of both personal and corporate taxation involves a tradeoff between rent extraction and efficiency loss. As noted earlier, no agent will mimic the highest type and hence there should be no distortion for this type.

To highlight the main finding, we state:

Proposition 1. *The optimal corporate income tax schedule derived has the feature that the higher the productivity of a firm, the lower the marginal tax rate the firm should face. This feature of taxation is to implement the social optimum whereby the higher the productivity of a firm, the lower the further distortion should be.*

The marginal tax rate τ^* is directly applied to the observable output y according to (30). This appears inconsistent with what we observe in the real world. Corporate income tax codes in practice usually entail allowances for the firm's worker payment as a tax deductible against y. However, the true worker remuneration r is a piece of information internal to the firm and often unobservable to an outsider like the government. Worker compensation could take the form of wages, salaries, or other compensation such as vacation allowances, bonuses, commissions, and fringe benefits. Tax authorities typically impose strict regulations on their deductibility to prevent a firm from "inflating" its

---

[14] For a practical assessment of this theoretical result, see Saez (2001).
[15] See Salanie (2003, chapter 4) for a summary of the literature.

business expenses.[16] These regulations are likely to cause a substantial deviation between the true worker payment and the tax deductible officially allowed. In the real world, there exists the so-called "third party" to verify the business expenses incurred by the firm. However, this raises an extra issue regarding the possible collusion between the firm and the third party.

One can rewrite (30) as

$$\Gamma(\beta, y) = g^{**}(\beta) + \tau^{*}(\beta)\{y - \gamma[e^{*}(\beta)]\}, \tag{30'}$$

where

$$g^{**}(\beta) = [1 - \tau^{*}(\beta)]\{\beta + e^{*}(\beta) - \gamma[e^{*}(\beta)]\} - \pi^{*}(\beta).$$

The above expression indicates that the tax deductible from y is not the true, random worker compensation r but the "mean" of r. This treatment of taxation may not be unreasonable in view of the circumstance where the realization of random variable r is unobservable to the government. It in fact has the flavor of a presumptive tax in that a substitute for the otherwise appropriate tax base is used by the tax authorities, due to the difficulty of measuring, verifying, or monitoring the latter.[17] In any case, the corporate income tax schedule derived here must be viewed in a normative rather than positive way: it depicts what it ought to be (prescription) rather than why it is (description). Interestingly, if the deductible $\gamma[e^{*}(\beta)]$ in (30') were to be replaced with the true worker compensation $r(e)$, then the government would not be able to accomplish the distortion e from e^{**} to implement $e^{*}(\beta)$. This is due to the well-known tax neutrality of economic profit taxation.[18]

5.2. Impact of increasing η or σ^2

---

16 See, for example, IRS (2008, chapter 2) regulation of employees' pay.

17 The practice of presumptive taxation can be seen in developed as well as developing countries. As Slemrod and Yitzhaki (1994, p. 25) put it: "All taxes are presumptive, to some degree."

18 From (4), the employer would become to maximize $(1 - \tau^{*}(\beta))E[y - r(e)]$ and hence $e = e^{**}$ would hold, irrespective of firm type.

How will increasing η or σ^2 impact the marginal tax rates of optimum corporate income tax? This question is interesting since it asks about how shocks in the employer-worker "contract" may impact the government-firm "contract."

In the pure moral hazard model of Mirrlees (1974), the benefit of providing insurance will increase if individuals become more risk averse or environments become more stochastic. As a result of the increased benefit of insurance, imposing higher marginal tax rates on personal income is socially desirable.[19]

When workers become more risk averse or working environments become more stochastic, the benefit of insurance provided to workers will also increase in our context. As a result, it is socially desirable to blunt the worker incentive and lower the worker effort. However, this is only a partial picture in our case.

The impact of variations in η or σ^2 on the firm is summarized by (5), the firm's expected payment to the worker. This payment equals the disutility that worker suffers due to effort plus the risk that the worker bears, which together constitute the real cost of our economy. Since $\tau^*(\beta) = 1 - \gamma'[e^*(\beta)]$, the optimal marginal tax rate $\tau^*(\beta)$ is inversely related to the marginal real cost $\gamma'[e^*(\beta)]$.[20] Increasing η or σ^2 will reduce the optimal $e^*(\beta)$ and hence $\gamma'[e^*(\beta)]$. This reduction stems from the increased benefit of providing insurance, as noted above. However, since $\gamma'(e) = \psi'(e) + \psi'(e)\psi''(e)\eta\sigma^2$ from (5), increasing η or σ^2 will also increase $\gamma'[e^*(\beta)]$ directly. Formally, using (5) and the definition of τ^* in (30-1) yields

$$\frac{d\tau^*}{d(\eta\sigma^2)} = -\gamma''(e^*) \cdot \frac{\partial e^*}{\partial(\eta\sigma^2)} - \psi'\psi'', \tag{32}$$

---

[19] See, for example, Varian (1980) for the detail. In the design of social insurance, Low and Maldoom (2004) emphasized the incentive effect of uncertainty to depress optimal marginal tax rates on personal income, due to the individuals' precautionary behavior. This incentive effect is absent in our model because of the worker's utility function specified in the classical agency theory.

[20] The last term of (22) represents the extent of the further distortion from the second best characterized by (19), and (22) dictates that the higher the marginal real cost, the lower the extent of the further distortion should be. This leads to an inverse relationship between the marginal real cost and the optimal marginal tax rate.

which shows the indirect effect of varying $\eta\sigma^2$ on τ^* through e^* (the first term on the right), and the direct effect on τ^* by shifting $\gamma'(e)$ (the second term). From (22), we have

$$\frac{\partial e^*}{\partial(\eta\sigma^2)} = -\frac{\psi'\psi'' + (1-\frac{1}{\lambda})\frac{1-F}{f}(\psi''^2 + \psi'\psi''')}{\gamma'' + (1-\frac{1}{\lambda})\frac{1-F}{f}\gamma'''} < 0, \tag{33}$$

where $\gamma'' = \psi'' + \eta\sigma^2(\psi''^2 + \psi'\psi''')$ and $\gamma''' = \psi''' + \eta\sigma^2(3\psi''\psi''' + \psi'\psi''''')$, both of which are derived from (5). The indirect effect in (32) is positive since $e^*(\beta)$ will go down as $\eta\sigma^2$ goes up according to (33). However, the direct effect in (32) is negative due to the upward shift in $\gamma'(e)$. The overall effect of varying $\eta\sigma^2$ on τ^* is ambiguous in general.

To sum up, we state:

Proposition 2. *The optimal corporate income tax derived has the feature that the impact of a higher η or σ^2 upon the optimal marginal tax rate is ambiguous in general.*

Nevertheless, consider the popular setting with $\psi(e) = (1/2)e^2$. This setting implies $\gamma''' = 0$ and hence $d\tau^*/d(\eta\sigma^2) = [1-(1/\lambda)](1-F)/f$ from (32) and (33). A higher η or σ^2 thus exerts an unambiguously positive impact on the optimal marginal tax rate of corporate income tax. This result indicates that, as long as the term γ''' in (33) is small enough, the indirect effect of varying $\eta\sigma^2$ on τ^* will be dominating.

6. Numerical simulations

We have shown that the marginal tax rates for the optimum corporate income tax are restricted to being non-negative and less than 100%, and that they are monotonically decreasing with respect to firm productivity and reach zero for the highest type. We have also shown that the impact of a higher η or σ^2 upon the optimal marginal tax rate is ambiguous in general. These are qualitative results. This section provides quantitative results on the optimal structure of marginal tax rates through numerical simulations.

Putting (22) and (30-1) together yields

$$\tau^*(\beta) = (1-\frac{1}{\lambda})[\frac{1-F(\beta)}{f(\beta)}]\gamma''[e^*(\beta)]. \tag{34}$$

To calculate the optimal tax rate $\tau^*(\beta)$ quantitatively according to (34), we must (i) specify $F(\beta)$ and $\psi(e)$, (ii) assign values for the parameters η and σ^2, and (iii) obtain values for both λ and $e^*(\beta)$. We address each of them in turn.

We first consider the case where $F(\beta)$ is a uniform distribution with $\beta \in [0,1]$. The uniform distribution may be a good prior in that it does not discriminate against any type a priori. We also consider the other two distributions of β. Specifically, we let $f(\beta) = \beta^{v-1}(1-\beta)^{w-1} / B(v,w)$, where $B(v,w)$ is a beta function and $v, w > 0$ are its parameters. This is a beta distribution defined on the interval $\beta \in (0,1)$. We choose either $(v,w) = (1,3)$ or $(v,w) = (5,1)$ in our simulations. The distribution of β is skewed to the right in the former set of parameter values, while to the left in the latter set. We examine how the optimal marginal tax rates will be altered when the uniform distribution of β is replaced by the beta distribution. Note that both distributions, uniform and beta, satisfy the property of a strictly monotone hazard rate: $d\{[1-F(\beta)]/f(\beta)\}/d\beta < 0$. We choose $\psi(e) = (1/3)e^3$ in our simulations. This choice is mainly to allow for the impact of $\psi'''(e)$. As noted after Proposition 2, if $\psi(e) = (1/2)e^2$, a higher η or σ^2 would exert a positive impact unambiguously on the optimal marginal tax rate.

In a recent paper, Cohen and Einav (2007) estimated risk preferences from data on deductible choices in auto insurance contracts. When imposing the exponential utility function as in our model, they found $\eta = 0.0031$ for the mean individual. This estimate is comparable with previous estimates in their comparison. Thus, we assign this value for η. As to the value of σ^2, we appeal to the widely cited work of Demsetz and Lehn (1985), which measured the instability of a firm's profit due to variations in stock and accounting returns. Using the data that consists of 511 firms from major sectors of the US economy during the years 1976-80, the authors found that the mean for the standard deviation of monthly stock market rates of return was 0.084, and that the mean for the standard deviation of annual accounting rates of return was 0.055. We simply use the latter value for σ in our illustration. The data used in Demsetz and Lehn (1985) are admittedly old. To remedy this defect, we consider other values for σ. The minimum for the standard deviation of annual accounting rates of return in the sample used by Demsetz and Lehn (1985) is 0.002, while the maximum is 0.320. The difference between them is substantial. To check the sensitivity of our results to variations in σ, we also report the results for these two extreme values. Note that the parameters η and σ^2 always take the form $\eta\sigma^2$ in (34). Thus, our check on the sensitivity to variations in σ also serves as a check on the sensitivity to variation in η.[21]

After specifying $F(\beta)$ and $\psi(e)$ and assigning values for η and σ^2, we utilize (30) and $\pi(\underline{\beta})=0$ to calculate the corresponding $e^*(\underline{\beta})$. Substituting the calculated $e^*(\underline{\beta})$ in (22) with $\beta=\underline{\beta}$ enables us to obtain the value for λ. With the value of λ at hand, we can then obtain $e^*(\beta)$ for all β's on the basis of (22).

(Insert Table 1 about here)

Our benchmark model in the simulations is the case where $F(\beta)$ =uniform, $\psi(e)=(1/3)e^3$, $\eta=0.0031$, and $\sigma=0.055$. Table 1 reports the results obtained. The optimal marginal tax rate τ^* declines monotonically from .25 to .13 to 0 as firm productivity β increases from 0 to .5 to 1. Equation (34) decomposes the determination of τ^* into three parts: $(\lambda-1)/\lambda$, $[1-F(\beta)]/f(\beta)$, and $\gamma''[e^*(\beta)]$. We report the corresponding changes in these three parts as τ^* varies in Table 1. Since $(\lambda-1)/\lambda$ remains constant with respect to β while $\gamma''[e^*(\beta)]$ is increasing in β, the decline in τ^* is obviously attributed to changes in the part $[1-F(\beta)]/f(\beta)$.

(Insert Table 2 about here)

---

21 Without imposing the exponential utility, Cohen and Einav (2007) found $\eta=0.0067$ for the mean individual in their benchmark model. We try this value for η and find little difference in our simulation results.

Table 2 reports the resulting optimal marginal tax rates as we vary σ substantially from that in the benchmark model. As can be seen from the table, the optimal marginal tax rates become higher if σ is lowered from .055 to .002; they become lower if σ is raised from .055 to .32. This implies that, opposite to the result with $\psi(e) = (1/2)e^2$, the direct effect dominates the indirect effect in (32). However, even though the variations in σ are substantial, the corresponding changes in the optimal marginal tax rate τ^* are very slight and remain almost unchanged. The direct and indirect effects of (32) cancel each other out to a large extent as σ varies.

(Insert Table 3 about here)

Table 3 reports the resulting optimal marginal tax rates as we replace a uniform $F(\beta)$ in the benchmark model by the two different beta distributions. In all distributions, $\tau^*(\beta = 0) = 0.25$ and $\tau^*(\beta = 1) = 0$. Relative to the uniform distribution, the optimal marginal tax rates display a sharp decline at low β's in the case of the beta distribution with $(v, w) = (1,3)$, while they display a sharp decline at high β's in the case of the beta distribution with $(v, w) = (5,1)$. Overall, it seems that the optimal structure of marginal tax rates is significantly affected by the distribution of firm productivity.

7. Conclusion

Due to firms being viewed simply as a production set, modern corporations are largely absent from the theory of optimum taxation. In this paper we address optimum corporate income tax by adopting a modern approach wherein firms are viewed as a "nexus of contracts" in the principal-agent framework. Our model involves three parties – workers, employers, and the government, and has elements of both moral hazard and adverse selection. We derive the socially optimal allocation and implement the optimum via the imposition of a corporate income tax. The corporate income tax schedule derived has the feature that the higher the productivity of a firm, the lower the marginal tax rate the firm should face. This *analytically monotonic* feature is in stark contrast to the *numerically non-monotonic* feature in the optimum personal income tax. The corporate income tax derived has also the feature that a more stochastic working environment or more risk-averse worker preference entails an ambiguous impact on the optimal marginal tax rate in general. This ambiguous feature differs from the standard result in the pure moral hazard model, in which a higher risk induces a higher demand for social insurance and so a higher optimal marginal tax rate. We also provide quantitative results on the optimal structure of marginal tax rates through numerical simulations.

As a first step toward an optimum corporate income tax, our modeling of firms is admittedly highly stylized and abstracts from many complications in the real world. For example, instead of a single agent vs. a single principal, a firm may take the organizational form consisting of many agents or principals. As another example, to focus on corporate income tax, we exclude the consideration of personal income tax. However, the co-existence of personal and corporate income tax is often observed in the real world.[22] Despite these and other possible abstractions, it is hoped that our simple model may well serve as a stepping stone for further study on the issue.

[22] For a broad discussion of taxing corporate income, especially in the context of globalization, see Auerbach et al. (2010).

Appendix

Derivation of (11)

From (7), the second-order condition of (6) is

$$\left\{1-\gamma'[\hat{e}(\hat{\beta}\big|\beta)]\right\}\cdot[\ddot{\hat{e}}(\hat{\beta}\big|\beta)]-\gamma''[\hat{e}(\hat{\beta}\big|\beta)]\cdot[\dot{\hat{e}}(\hat{\beta}\big|\beta)]^2-E[\ddot{T}(\hat{\beta}\big|\beta)]\leq 0\,, \tag{A1}$$

where the notation of two dots denotes a second derivative with respect to $\hat{\beta}$.

Using $\hat{e}(\hat{\beta}\big|\beta)\equiv e(\hat{\beta})+\hat{\beta}-\beta$, we have $\dot{\hat{e}}(\hat{\beta}\big|\beta)\equiv\dot{e}(\hat{\beta})+1$ and $\ddot{\hat{e}}(\hat{\beta}\big|\beta)\equiv\ddot{e}(\hat{\beta})$. Thus, with the truth telling $\hat{\beta}=\beta$, (A1) yields

$$\left\{1-\gamma'[e(\beta)]\right\}\cdot[\ddot{e}(\beta)]-\gamma''[e(\big|\beta)]\cdot[\dot{e}(\big|\beta)+1]^2-E[\ddot{T}(\beta)]\leq 0\,. \tag{A2}$$

Next, from (8), we have

$$E[\ddot{T}(\beta)]=\left\{1-\gamma'[e(\beta)]\right\}\cdot[\ddot{e}(\beta)]-\gamma''[e(\big|\beta)]\cdot[\dot{e}(\big|\beta)+1]\dot{e}(\beta)=0\,. \tag{A3}$$

Substituting (A3) in (A2) leads to (11).

References

Auerbach, A.J., Devereux, M.P., Simpson, H., 2010. Taxing corporate income. In: Mirrlees, J., et al. (Eds.), Dimensions of Tax Design: The Mirrlees Review. Oxford University Press, Oxford.

Bagnoli, M., Bergstrom, T., 2005. Log-concave probability and its applications. Economic Theory 26, 445-469.

Besley, T., Coate, S., 1995. The design of income maintenance programmes. Review of Economic Studies 62, 187-221.

Bolton, P., Dewtripont, M., 2005. Contract Theory. MIT Press, Cambridge MA.

Cohen, A., Einav, L., 2007. Estimating risk preferences from deductible choice. American Economic Review 97, 745-788.

Demsetz, H., Lehn, K., 1985. The structure of corporate ownership: causes and consequences. Journal of Political Economy 93, 1155-1177.

Diamond, R.A., Mirrlees, J.A., 1971. Optimal taxation and public production I: production efficiency. American Economic Review 61, 8-27.

Gibbons, R., 1997. Incentives and careers in organization. In: Kreps D.M., Wallis, K.F., (Eds.), Advances in Economic Theory and Econometrics: Theory and Applications, Vol. 2. Cambridge University Press, Cambridge.

Golosov, M., Tsyvinski A., Werning, I., 2006. New dynamic public finance. NBER Macroeconomics Annual 21, 317-363.

Holmstrom, B., Milgrom, P., 1987. Aggregation and linearity in the provision of intertemporal incentives. Econometrica 55, 303-328.

IRS, 2008. Business Expenses, IRS Publication 535. IRS, Washington, D.C.

Kleven, H.J., Kreiner, C.T., Saez, E., 2009. Why can modern government tax so much? An agency model of firms as fiscal intermediaries. Mimeo.

Kopczuk, W., Slemrod, J., 2006. Putting firms into optimal tax theory. American Economic Association Papers and Proceedings 96, 130-134.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The Principal-Agent Model. Princeton University Press, Princeton.

Laffont, J.-J., Tirole, J., 1986. Using cost observation to regulate firms. Journal of Political Economy 94, 614-641.

Laffont, J.-J., Tirole, J., 1993. A Theory of Incentives in Procurement and Regulation. MIT Press, Cambridge MA.

Leonard, D., Long, N.V., 1992. Optimal Contol Theory and Static Optimization in Economics. Cambridge University Press, Cambridge.

Low, H., Maldoom, D., 2004. Optimal taxation, prudence and risk-sharing. Journal of Public Economics 88, 443-464.

Melumad, N.D., Reichelstein, S., 1989. Value of communication in agencies. Journal of Economic Theory 47, 334-368.

Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175-208.

Mirrlees, J.A., 1974. Notes on welfare economics, information and uncertainty. In: Balch, M., McFadden, D., Wu, S., (Eds.), Essays in Equilibrium Behavior under Uncertainty. North-Holland, Amsterdam.

Mortimort, D., Moreira, H., 2010. Common agency and public good provision under asymmetric information. Thoeretical Economics 5, 159-213.

Phelps, E.S., 1986. Profits theory and profits taxation. IMF Staff Papers 33, 674-696.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68, 205-229.

Salanie, B., 2003. Economics of Taxation. MIT Press, Cambridge MA.

Salanie, B., 2005. The Economics of Contract, 2nd Edition. MIT Press, Cambridge MA.

Slemrod, J., Yitzhaki, S., 1994. Analyzing the standard deduction as a presumptive tax. International Tax and Public Finance 1, 25-34.

Varian, H., 1980. Redistributive taxation as social insurance. Journal of Public Economics 14, 49-68.

Williamson, O.E., 1981. The modern corporation: origins, evolution, attributes. Journal of Economic Literature 19, 1537-1568.

Yang, C.C., 1993. Optimal linear income tax with random revenue. Journal of Public Economics 52, 391-401.

Table 1
Optimal marginal tax rates and their determinants in the benchmark model

β	τ^*	e^*	$(\lambda\text{-}1)/\lambda$	$(1\text{-}F)/f$	γ''
0	0.25	0.866	0.144	1	1.733
0.1	0.23	0.878	0.144	0.9	1.757
0.2	0.21	0.891	0.144	0.8	1.783
0.3	0.18	0.904	0.144	0.7	1.809
0.4	0.16	0.917	0.144	0.6	1.835
0.5	0.13	0.930	0.144	0.5	1.861
0.6	0.11	0.944	0.144	0.4	1.888
0.7	0.08	0.957	0.144	0.3	1.916
0.8	0.06	0.971	0.144	0.2	1.944
0.9	0.03	0.985	0.144	0.1	1.972
1	0	0.999	0.144	0	2.001

Table 2
Optimal marginal tax rates under different σ's

β	σ		
	0.055	0.002	0.32
0	0.249981	0.249996	0.249915
0.1	0.228224	0.228237	0.228167
0.2	0.205794	0.205805	0.205745
0.3	0.182673	0.182682	0.182632
0.4	0.158843	0.158851	0.158811
0.5	0.134288	0.134295	0.134263
0.6	0.108990	0.108994	0.108971
0.7	0.082929	0.082933	0.082917
0.8	0.056090	0.056092	0.056082
0.9	0.028453	0.028453	0.028449
1	0	0	0

Table 3
Optimal marginal tax rates under different *F*'s

β	*F*		
	Uniform	Beta (1, 3)	Beta (5, 1)
0	0.250	0.250	0.250
0.1	0.228	0.189	0.250
0.2	0.206	0.137	0.250
0.3	0.183	0.094	0.249
0.4	0.159	0.060	0.248
0.5	0.134	0.035	0.243
0.6	0.109	0.018	0.233
0.7	0.083	0.008	0.213
0.8	0.056	0.002	0.176
0.9	0.028	0.000	0.111
1	0	0	0