PET
7/27/06
10:15 AM

An Agency Theory of Dividend Taxation*

Raj Chetty and Emmanuel Saez

UC Berkeley and NBER

July 25, 2006

Abstract

This paper analyzes the efficiency consequences of dividend taxation in a simple agency model of corporate behavior where managers have a higher preference for retained earnings than shareholders. In this model, a dividend tax cut leads to an increase in dividend payments in the short run. Firms where managers place more weight on profit maximization are more likely to raise dividends in response to the tax cut. Dividend tax cuts can generate large efficiency gains because they help correct agency issues. In the long run, however, a dividend tax cut can lead shareholders to reoptimize their contract with managers. As a result, the dividend tax cut has smaller effects on dividend payments and economic efficiency in the long-run than in the short-run. In contrast with the benchmark old view and new view models of dividend taxation, our model matches the empirical findings from the 2003 dividend tax cut, showing that there was an immediate response of dividends concentrated in firms with strong principals or large top executive share ownership.

*Raj Chetty, chetty@econ.Berkeley.edu, Emmanuel Saez, saez@econ.berkeley.edu, University of California, Department of Economics, 549 Evans Hall #3880, Berkeley, CA 94720. We thank Alan Auerbach, Martin Feldstein, Roger Gordon, Kevin Hassett, and James Poterba for very helpful comments and discussions. Financial support from National Science Foundation grants SES-0134946 and SES-0452605 is gratefully acknowledged.

1 Introduction

There are two leading theories of dividend taxation and corporate behavior: the "old view" (Harberger 1966, Feldstein 1970, Poterba and Summers 1985) and the "new view" (Auerbach 1979, Bradford 1981, King 1977). The old view assumes that marginal investment is financed by the external capital market through new equity issues. Under this assumption, the taxation of dividends raises the cost of capital and, as a result, has a negative effect on corporate investment and dividend payouts. The new view assumes that marginal investment is financed from the firm's retained earnings. In this case, the dividend tax rate does not affect the cost of capital because the dividend tax applies equally to current and future distributions. Therefore, the dividend tax rate does not affect the investment and dividend payout decisions of the firm.

There has been a controversial debate in the empirical literature trying to test the old view and new view predictions by estimating the responses of dividends to taxes. Feldstein (1970) and Poterba and Summers (1985), using time series evidence from the United Kingdom found that, consistent with the old view, dividend payments and investment were higher when the tax rate on dividends was lower. More recently, Poterba (2004) uses U.S. time series data from 1929 to 2002 and finds a negative association between dividend payments and the dividend tax rates relative to taxes on capital gains. However, using U.S. data from 1981 to 1998, Auerbach and Hassett (2003) show that consistent with the new view, dividend payments are sensitive to changes in investment at the firm level, suggesting that retained earnings are indeed the marginal source of investment funds.

More recently, several studies (Chetty and Saez (2005), Brown et al. (2006), Nam et al. (2005)) have analyzed the effects of the 2003 dividend tax cut in the United States. In particular, Chetty and Saez (2005) document four patterns around this large reform. First, regular dividends rose sharply after the 2003 tax cut, with an implied tax elasticity of -0.5. Second, the response was very rapid (total dividend payouts rose by 20% within one year of enactment) and was strong among firms with high retained earnings. Third, the response was much larger among firms where top executives owned a larger fraction of outstanding shares. Fourth, the response was much larger among firms with stronger principals (large shareholders).

Unfortunately, it is difficult to reconcile these four findings with either the old view or the

new view. The fact that dividends rose after the tax cut appears to support the old view because dividends should not respond to permanent dividend tax changes under the new view.[1] However, the speed of the response is too large for a supply-side mechanism where dividend payouts rise because of increased savings, leading to increased investment, and eventually to higher profits and dividend payouts. Indeed, Poterba's [2004] estimates imply that the 2003 tax reform should increase dividend payments by 20 percent in the long run, but that the adjustment process will be slow, with only a quarter of the long-run effect taking place within three years. The empirical evidence is also not fully explained by the model of Sinn (1991), who synthesizes the old view and new view by building a life cycle model of firms where firms start as old view firms and become new view firms when they reach maturity and start paying dividends. In this synthesized model, the payout response should be very small among firms with high levels of retained earnings, but the data reveal no such heterogeneity. The rapid dividend payout response could be explained by building in a signalling value for dividends as in Poterba and Summers (1985) or Bernheim (1992). However, neither the signalling model nor the standard old and new view models would predict findings (3) and (4) on the cross-sectional heterogeneity in the dividend payout response along agency variables.

In this paper, we propose a simple alternative model of dividend taxation, motivated by agency models of firm behavior that have been emphasized in the corporate finance literature. Our stylized model matches the four empirical findings described above. The key difference between the model proposed here and the existing old and new view models is that we allow the manager's objective function to depart from pure profit maximization. In particular, we consider a model where top managers are empire builders and have an interest in retaining earnings because they enjoy the perks and recognition that come from running a cash-rich company. Managers make dividend payout decisions to maximize a weighted average of their utility and the utility of large shareholders, who monitor the firm and can directly exert influence on corporate behavior at some cost. As in the new view model, we assume that the marginal dollar of investment is assumed to be financed from retained earnings, which as argued by Auerbach and Hassett (2003) describes many firms in the economy.

[1] One way of reconciling the dividend response with the new view is if the tax cut was perceived as temporary by firms. Although we cannot definitively rule out this possibility, the fact that firms chose to raise regular dividends (rather than make a few special dividend payments) and that the dividend increase has persisted (see Chetty and Saez, 2006) suggests that the payout changes are permanent. More importantly, the basic new view model would not explain findings (3) and (4) even for a temporary tax cut.

In our model, managers have a higher preference for retained earnings than shareholders because of their empire-building motive. As a result, they pay too few dividends and overinvest. In the short run, a dividend tax cut leads to an immediate increase in dividend payments and dividend initiations by changing the relative price of retaining earnings and distributions for the manager. Firms where managers place more weight on profit maximization (e.g. because the manager owns a large number of shares or because of monitoring by large shareholders) are more likely to raise dividends in response to a tax cut. Hence, the model offers a simple explanation of the empirical findings from the 2003 dividend tax cut that is consistent with marginal investment being funded out of retained earnings.

After setting up this model, we proceed to analyze its implications for the efficiency costs of taxation. The efficiency implications differ in the short run and long run. In the short run, when the contracts signed by owners and managers are fixed, a dividend tax cut may generate a large, first-order efficiency gain because it helps mitigate the overinvestment due to agency issues. In the long run however, contracts are presumably endogenous to the tax regime, and a dividend tax cut will lead shareholders to reduce costly monitoring of managers and sell less of the firm to the manager in the form of shares. Intuitively, the owners need to provide less incentives to the manager to maximize profits when tax rates are lower. This change in contracts attenuates the effect of the dividend tax cut on dividend payments in the long run, reducing the efficiency gain from cutting dividend taxes in the long run.

A key feature of our analysis is that it yields relatively simple reduced-form formulas for the short-run and long-run deadweight costs of taxation that are not sensitive to functional form. The parameters in these formulas include the short-run elasticity of dividend payouts with respect to the dividend tax rate and other elasticities that could at least in principle be estimated empirically. In this sense, our model provides a tractable framework for computing the efficiency costs of dividend taxation in a model consistent with empirical evidence.

Our study is related to a large literature in corporate finance analyzing agency issues within the firm. Following the seminal contribution by Jensen and Meckling (1976), several authors have analyzed models of corporate governance where the preferences of shareholders (principals) and managers (agents) are not perfectly aligned (see e.g. Shleifer and Vishny 1995 for a survey). While the agency theory of corporate behavior is very well developed in the corporate finance literature, it has had less influence in the public finance literature

on taxation. Our contribution is to connect these two literatures to derive new measures of the efficiency costs of capital income taxation, and identify key empirical parameters to be estimated in order to calibrate these measures in future work.

In a recent study independent of our work, Gordon and Dietz (2006) contrast the efficiency effects of dividend taxation in new view, signalling, and agency models. While our analysis shares some aspects with the model they develop, there are several important differences between our studies. First, Gordon and Dietz build an agency model that includes many more features than our stylized model in order to contrast it with other models on several dimensions. Our framework offers somewhat simpler, empirically implementable expressions for the efficiency costs of taxation. Second, Gordon and Dietz assume that dividend payout decisions are always made by the board (and not managers, as in our model), a distinction which makes their short-run analysis quite different from ours. In particular, our results on the higher efficiency costs of dividend taxation in the short run are not obtained in their framework. The two models provide complementary insights, and testing between them would be an interesting direction for future research.

The remainder of the paper is organized as follows. Section 2 describes the basic setup of the model. Section 3 explores comparative statics and efficiency costs of dividend taxation in the short run. Section 4 considers the same issues in the long-run, when contracts and ownership structure are endogenous. Section 5 offers concluding remarks.

2 Model

Consider a firm that has total cash holdings of X at $t = 0$. The firm can either invest $\$I$ in a project that has a payoff of $f(I)$ at $t = 1$ or pay out its cash holdings as dividends. At time 1, the firm closes and pays out $\$f(I)$ as dividends. Normalize the interest rate at 0. Let $D = X - I$ denote the firm's dividend payment at $t = 0$. A tax of t_d is levied on dividend payments in all periods.

Assume that the firm's investment and dividend payout policies are decided by the "key players" in corporate management: top executives and large shareholders. The objective of the large shareholders is to maximize the total amount of dividends that they get from the firm over the two periods. The top executives have an interest in maximizing dividend income

as well, but also have an empire-building tendency: they get a marginal private benefit worth $\$\beta$ of income for running a firm that is \$1 larger. The benefit β from retaining earnings can arise from either pure utility of empire building or benefits from spending the firm's free cash flow on perks.

The firm's policies are determined by maximizing a weighted average of the objectives of the key players. Let $\alpha < 1$ denote the weight placed on the objective of maximizing total dividend payouts. The firm's policies are determined by maximizing the following objective function:

$$\alpha\{D(1-t_d) + (1-t_d)f(X-D)\} + \beta(X-D) \tag{1}$$
$$\text{s.t. } D \geq 0$$

This objective function reflects the fact that the firm's decision makers (1) place some weight on total net-of-tax monetary payoffs, given by the first term and (2) the managers get benefits from empire-building, measured by the parameter the β. The weight on profits α is an increasing function of the fraction of shares owned by the top executives and the number of large stakeholders who are influential in corporate decisions. For instance, the presence of a large shareholder on the board of directors may lead to a higher value of α in the firm's objective function.

Note that when $\beta = 0$, this model reduces to a standard new-view model where dividends and investment are financed out of retained earnings and firms maximize average net-of-tax profits. Thus the model nests the new view. The old view requires that the marginal dollar of investment is funded by equity issues, and is not nested within this framework.

The key parameter in determining the firm's behavior is the ratio $\gamma = \frac{\alpha}{\beta}$, which captures the importance of neoclassical profit-maximization objectives relative to the empire-building motive. The parameter γ is a measure of the extent to which the interests of the principals (shareholders) and agents (top executives) diverge. Higher γ firms have less of an agency problem. In the short run, it is reasonable to assume that γ is exogenously determined, as firms' ownership and contracting structure is fairly stable over time and may not adjust quickly in response to a tax change.

In the long run, and particularly when firms are started, the parameter γ itself is presum-

ably endogenously chosen to maximize the principal's objective. In the long run, the principal chooses γ to maximize net-of-tax profits, which corresponds to the first term in (1). If there were no additional constraints or costs in the model, the principal would achieve his objective by making γ arbitrarily large, so that the firm focuses exclusively on profit maximization. In practice, there are two natural channels through which principals can raise γ: (1) by making executive shareownership large ("sell the firm to the manager") or (2) monitoring the firm closely as a large shareholder on the board of directors. However, both of these approaches create costs. A high level of executive shareownership improves incentives but reduces expected utility for a risk-averse manager, effectively forcing the principal to raise executive pay, thereby reducing net profits. A high level of monitoring could also involve significant direct costs. We model the costs of raising γ in a reduced-form manner by assuming that the principal must pay a cost of $c(\gamma)$ to generate an ownership structure that results in a given level of γ. Assume that c is increasing and convex. In the long run, γ is chosen to:

$$\max_{\gamma}\{D(\gamma)(1-t_d)+(1-t_d)f(X-D(\gamma))\}+\beta(X-D(\gamma))-c(\gamma)$$

where $D(\gamma)$ denotes the value of D that maximizes (1) for a given γ.

3 Dividend Taxation: Short Run Analysis

In this section, we study the effects of dividend taxation on firm behavior and economic efficiency when γ is fixed. Define $\overline{\gamma}(t_d)=\frac{1}{(1-t_d)(1-f'(X))}$. The firm's dividend payout level in period 0, D^*, is characterized by the following expression:

$$\begin{aligned} D(t_d,\gamma) &= \begin{cases} 0 \text{ if } \gamma \le \overline{\gamma}(t_d) \\ D^*(t_d) \text{ if } \gamma > \overline{\gamma}(t_d) \end{cases} \\ &\quad \text{where } D^*(t_d) \text{ is s.t. } f'(X-D^*(t_d)) = 1-\frac{1}{\gamma(1-t_d)} \end{aligned}$$

When γ is below the threshold value $\overline{\gamma}(t_d)$, the marginal value of the first dollar of dividends is negative in the firm's objective function. The optimal level of dividends is therefore zero, the corner solution. Intuitively, if managers have a sufficiently strong interest in empire building, they wish to retain as much money as possible within the firm, and do not choose to pay out dividends. For γ above this threshold value, the managers choose a level of dividends that balances the marginal benefit of empire building with the marginal benefit of paying out money and generating dividend income for themselves and the large shareholders.

Once γ exceeds the critical threshold $\overline{\gamma}(t_d)$, further increases in γ lead to increases in dividend payments:

$$\frac{\partial D^*}{\partial \gamma} = -\frac{1}{f''}\frac{1}{\gamma^2(1-t_d)} > 0.$$

Intuitively, firms that place less weight on empire building relative to profits have less interest in retaining earnings, and therefore pay out more dividends.

Now consider the effect of a tax cut on dividend payout behavior. Suppose the initial dividend tax rate is $t_1 > 0$ and the new tax rate is $t_2 < t_1$. This implies that $\overline{\gamma}(t_2) < \overline{\gamma}(t_1)$. The following result characterizes the effect of this tax cut on change in dividend payout, $\Delta D = D(t_2) - D(t_1)$:

Proposition 1 *A dividend tax cut has the following effects on dividend payout behavior:*
(i) if $\gamma < \overline{\gamma}(t_2)$, $\Delta D = 0$
(ii) if $\gamma \in [\overline{\gamma}(t_2), \overline{\gamma}(t_1)]$, $\Delta D > 0$ *and* $\frac{\partial \Delta D}{\partial \gamma} > 0$
(iii) if $\gamma > \overline{\gamma}(t_1)$, $\Delta D > 0$ *and* $lim_{\gamma\to\infty}\Delta D = 0$

Proof. (i) When $\gamma < \overline{\gamma}(t_2)$, $D(t_2) = 0$ by equation (XX). Since $\overline{\gamma}(t_2) < \overline{\gamma}(t_1)$, $D(t_1) = 0$ also. Therefore $\Delta D = 0$.

(ii) For $\gamma \in [\overline{\gamma}(t_2), \overline{\gamma}(t_1)]$, $D(t_1) = 0$ while $D(t_2) = D^*(t_2) > 0$ where $D^*(t_2)$ is such that $f'(X - D^*) = 1 - \frac{1}{\gamma(1-t_2)}$. Hence $\Delta D = D^*(t_2) > 0$ and $\frac{\partial \Delta D}{\partial \gamma} = \frac{\partial D^*}{\partial \gamma} > 0$ as shown in equation (XX).

(iii) When $\gamma > \overline{\gamma}(t_1)$, $\Delta D = D^*(t_2) - D^*(t_1)$. Note that $\frac{\partial D^*}{\partial t_d} = \frac{1}{f''}\frac{1}{\gamma(1-t_d)^2} < 0$. Hence $t_2 < t_1 \Rightarrow \Delta D > 0$. To compute $lim_{\gamma\to\infty}\Delta D = 0$, let g denote the inverse of $f'(\cdot)$. Then $D^*(t_d) = X - g(1 - \frac{1}{\gamma(1-t_d)})$ and

$$\Delta D = g(1 - \frac{1}{\gamma(1-t_1)}) - g(1 - \frac{1}{\gamma(1-t_2)})$$

Hence $lim_{\gamma\to\infty}\Delta D = g(1) - g(1) = 0$. ■

The results of Proposition 1 are illustrated in Figure 1, which plots D against the γ parameter in the two tax regimes, with $t_1 = 0.4$ and $t_2 = 0.2$. These simulations assume $f(x) = x - \frac{x^2}{2}$ with initial cash holding $X = 1$. The effect of the tax cut on dividend payout behavior differs across three regions of the γ parameter. When $\gamma < \overline{\gamma}(t_2)$, paying any dividends is suboptimal after the tax cut, and hence before the tax cut as well. Firms in this region do not pay dividends before the tax cut, and do not initiate dividend payments after

the tax cut; hence, $\Delta D = 0$ for such firms. The second region consists of firms who were non-payers prior to the tax cut ($\gamma < \overline{\gamma}(t_1)$), but cross the threshold for paying when the tax rate is lowered to t_2. These firms initiate dividend payments after the tax cut. Since D^*, the optimal dividend conditional on paying, is rising in γ, the size of the dividend increase, ΔD, is larger for firms with higher values of γ in this region.

The third region consists of firms who had sufficiently high levels of γ that they were already paying dividends at the higher tax rate t_1 prior to the tax cut. The tax cut leads these firms to place greater weight on net-of-tax profits relative to empire-building benefits and therefore causes increases in dividend payments. In general, the relationship between the size of the dividend increase and the value of γ in this region is indeterminate, and depends on the third derivative of the production function f. However, as γ approaches ∞, the size of the dividend increase ΔD always converges to zero, regardless of the shape of f. Intuitively, for large values of γ, the firm effectively puts little weight on empire building and essentially maximizes total profits:

$$D(1 - t_d) + (1 - t_d)f(X - D)$$

In this expression, the $(1 - t_d)$ term factors out and changes in the dividend tax rate have no impact on investment or dividend payout behavior because the relative price of paying dividends in periods 1 and 2 is unchanged. This result for the limiting case is not surprising: when $\gamma \to \infty$, the model collapses to the standard new view model, where permanent dividend tax reductions have no effect on dividend payouts. Hence, for firms with very high values of γ that have no divergence of interests between principals and agents, dividend tax cuts have no effect on real behavior.

3.1 Relationship to empirical evidence

The short-run comparative statics implied by this agency model of dividend taxation match the four motivating empirical findings outlined in the introduction:

(1) The model predicts that a tax cut induces an increase in dividend payments on both the intensive and extensive margins, consistent with the evidence.

(2) Unlike in Sinn's synthesized new and old view model, and consistent with the evidence, the dividend increases are not predicted to vary systematically with the size of retained earnings (denoted by X). The model also predicts an immediate adjustment (in period 0) of dividend

payout policies rather than a slow adjustment process. In the agency model, dividend payouts change immediately because the managers incentives to retain earnings are reduced after a tax cut, and do not involve the lags inherent in raising additional capital, investing, etc..

(3, 4) In addition to matching the observed pattern of dividend changes in the time series, the model also fits the cross-sectional heterogeneity of the dividend response following the 2003 tax cut. At low levels of γ, the model predicts that higher γ firms are more likely to initiate dividend payments after the tax cut. Higher γ firms are those where manager's interests are more closely aligned with shareholders in the sense that greater weight is placed on profit maximization. Firms where top executives own a large fraction of shares and where large shareholders are influential in corporate decision making are more likely to focus on profit maximization than perks or empire building. Hence, consistent with the empirical evidence, the model predicts that such firms are more likely to initiate dividends following a tax cut.

The key feature of the model that generates these comparative statics is that corporate decisions are driven by the interests of the "key players" rather than the shareholders at large. Further evidence for this assumption comes from the empirical finding in Chetty and Saez (2005) that firms with large non-taxable shareholders such as pension funds were much less likely to change dividend payout behavior in response to the tax cut. Although we have not explicitly allowed heterogeneity in tax rates across shareholders in our stylized model, the introduction of non-taxable shareholders would generate this prediction. In particular, if non-taxable shareholders are influential, more weight is placed on pre-tax profits, in which case the change in t_d has a smaller impact on the firm's objective function and therefore generates smaller ΔD.

One caveat in connecting the model to the empirical evidence is that for firms with very high levels of γ, the model predicts little or no change in dividend payout behavior. However, the evidence from the 2003 tax cut does not suggest such a non-monotonic relationship between ΔD and executive shareownership or large shareholder's presence. One way to reconcile the theoretical prediction with the evidence is that the range of γ's observed in data on publicly traded corporations is likely to be confined to the lower two regions in Figure 1. In practice, the largest executive and outside blockholder's shareholdings rarely exceed 10% of outstanding shares among publicly traded corporations. This suggests that γ is never very high within this group; there are non-trivial divergences of interests between corporate managers and

shareholders in most large firms. Therefore, the downward-sloping region of the ΔD vs. γ curve may not be observed in datasets on publicly traded corporations (such as CRSP or COMPUSTAT) that are typically used in empirical work. However, it is possible that among closely held firms, γ may be much higher, and the dividend tax cut could have had smaller effects on total payout. Empirical analysis of the effects of the 2003 dividend tax cut on the behavior of these smaller firms is an interesting direction for future research. In particular, it would be very interesting to compare the behavior of subchapter C versus subchapter S closely-held corporations, as subchapter S corporations profits constitute a natural "control group" because they are taxed only at the individual level and hence were not affected by the 2003 tax cut.

3.2 Efficiency analysis

Let $D_{TOT}(t_d) = D + f(X - D)$ denote the total (pre-tax) amount of dividends paid by the firm over two periods given the dividend tax rate. The government's revenue from dividend taxation is then $R = D_{TOT}t_d$. Total social welfare is given by the money metric

$$W(t_d) = (1 - t_d)D_{TOT}(t_d) + \beta(X - D(t_d))$$

The social welfare function sums dividend income that accrues to *all* shareholders, ignoring the weight α placed on profits by top executives and large shareholders in the firm's objective. The social welfare and the firm's objective function do not coincide because of the agency problem in running the firm.

The marginal deadweight cost of raising the dividend tax can be defined as the marginal revenue from increasing t_d plus the change in welfare from this increase:

$$\begin{aligned} MDWL(t_d) &= \frac{\partial R}{\partial t_d} + \frac{\partial W}{\partial t_d} \\ &= D_{TOT} + t_d\frac{\partial D_{TOT}}{\partial t_d} + [-D_{TOT}(t_d) + (1-\alpha)(1-t_d)\frac{\partial D_{TOT}}{\partial t_d}] \\ &= [t_d + (1-\alpha)(1-t_d)]\frac{\partial D_{TOT}}{\partial t_d} \end{aligned}$$

To understand the key features of this expression, it is useful to consider the case where there is no tax in place to begin with: $t_d = 0$. In this case, the marginal efficiency cost of raising t_d is

$$MDWL(t_d = 0) = (1 - \alpha)\frac{\partial D_{TOT}}{\partial t_d} < 0$$

This expression shows that imposing even a small dividend tax has a first-order efficiency cost in the agency model. Intuitively, the principal-agent problem leads to under-provision of dividends in the private market equilibrium because managers effectively ignore the negative externality they are having on shareholders by focusing too much attention on empire building. By taxing dividends, the government amplifies the cost of this externality, by exacerbating a pre-existing market distortion (see Auerbach and Hines 2003 and Goulder and Williams, 2003). This creates a first-order deadweight loss from dividend taxation, the size of which is determined by α. If α is small, the firm's objective is less aligned with social welfare, and the efficiency costs of dividend taxation are greater.

To understand the policy implications of this point more concretely, suppose the government has access to a second tax, such as a labor income tax, and that the labor market does not have any pre-existing distortions. Letting t_l denote the tax on labor income and l denote labor supply earnings, it is easy to show that the marginal deadweight loss of taxing labor income is given by

$$MDWL(t_l) = t_l \frac{\partial l}{\partial t_l}$$

This expression shows that the marginal efficiency cost of taxing labor income is 0 when $t_l = 0$. Hence, a benevolent government that needs to raise a small amount of revenue should rely on only labor income taxation and should not tax dividends at all. Indeed, the government would improve welfare by *subsidizing* dividend payments ($t_d < 0$) and financing that subsidy through the labor income tax.

More generally, one can solve for the set of optimal Ramsey tax rates on dividend and labor earnings which maximize welfare (or equivalently, minimize deadweight burden) while raising a given amount of revenue for the government. Denoting by $\lambda > 1$ the marginal cost of public funds ($\lambda - 1$ is a measure of marginal deadweight burden of taxation), the optimal tax rates are given by:

$$\frac{t_l}{1 - t_l} = \frac{\lambda - 1}{\lambda} \cdot \frac{1}{\varepsilon_l} \tag{2}$$

$$\frac{t_d}{1 - t_d} = \frac{\lambda - 1}{\lambda} \cdot \frac{1}{\varepsilon_d} - \frac{1 - \alpha}{\lambda} \tag{3}$$

where ε_l and ε_d denote the elasticities of labor earnings and dividends with respect to their tax

rates. The second term in equation (3) is negative, and is the corrective term for the negative externality of dividend taxes on agency issues within the firm. This result parallels that of the standard theory of optimal taxation with externalities (Sandmo 1975).

A useful feature of the expressions derived here for optimal tax rates and the deadweight loss of taxation is that they depend only on the elasticity of dividends with respect to the dividend tax rate and the parameter α. The details of the remaining structure of the model are irrelevant conditional on knowing these values. The parameter α can be interpreted as the answer to the question, "What salary increase would make a manager indifferent between a \$1 increase in profits and that salary change?"

4 Dividend Taxation: Long Run Analysis

We now consider the long run analysis, where the owners of the firm re-optimize γ given the new dividend tax rate. The maximization program

$$\max_{\gamma}\{D(\gamma)(1-t_d)+(1-t_d)f(X-D(\gamma))\}+\beta(X-D(\gamma))-c(\gamma)$$

defines an implicit equation $\gamma(t_d)$ which captures the fact that, in the long-run, a change in t_d will generate a change in γ. As a result, the change in dividend due to the tax change can be decomposed into a direct effect of t_d (as in the short-run analysis), and an indirect effect via the change in γ. Using the first order condition for the manager's choice of dividends D^*

$$f'(X-D^*)=1-\frac{1}{\gamma(1-t_d)},$$

we obtain:

$$\frac{\partial D^*}{\partial t_d}=\frac{1}{f''}\frac{1}{\gamma(1-t_d)^2}-\frac{1}{f''}\frac{1}{\gamma^2(1-t_d)}\frac{\partial\gamma}{\partial t_d},$$

Denote by

$$\varepsilon_d^S=\frac{1-t_d}{D^*}\frac{\partial D^*}{\partial(1-t_d)}|_\gamma$$

the short-run elasticity of dividends with respect to the net-of-tax rate $(1-t_d)$ where γ is held constant. Denote by

$$\varepsilon_d^L = \frac{1-t_d}{D^*}\frac{dD^*}{d(1-t_d)}$$

the long-run elasticity of dividends with respect to the net-of-tax rate $(1-t_d)$, where γ is endogenous. Then

$$\varepsilon_d^L = \varepsilon_d^S \cdot \left(1 - \frac{1-t_d}{\gamma}\frac{\partial \gamma}{\partial t_d}\right).$$

Hence, if an increase in t_d leads to an increase in monitoring (to help align incentives and alleviate the over-investment problem), the long-run elasticity is smaller than the short-run elasticity.

For example, if we consider the simple parametric functional form:

$$f(I) = I - \frac{I^{1+\frac{1}{e}}}{1+\frac{1}{e}}$$

and let $c(\gamma) = \gamma$, we obtain:

$$\frac{1-t_d}{\gamma}\frac{\partial \gamma}{\partial t_d} = \frac{e}{2+e+\alpha/(1-\alpha)} > 0.$$

When e is small, the long-term offset is small and when e is large, the long-run offset is close to 100%. Intuitively, when the government changes the dividend tax rate, it changes its stake of ownership in the firm. This change in the amount of third party ownership will induce re-contracting between the owners and agents to reoptimize incentives, mitigating the effect of the government tax change on firm behavior. The degree of re-contracting is in general determined by the shape of the production function, costs of monitoring, etc.

4.1 Efficiency analysis

In the long-run, because γ is optimized, the envelope condition implies that a small dividend tax no longer generates a first-order deadweight burden. Therefore, the usual deadweight burden formulas with no externalities apply, as in the labor income tax example given above. The long-run optimal Ramsey tax on dividends becomes:

$$\frac{t_d^L}{1-t_d^L} = \frac{\lambda - 1}{\lambda} \cdot \frac{1}{\varepsilon_d^L}. \tag{4}$$

Therefore, in the long run, the optimal taxation of dividends follows the usual Ramsey inverse elasticity rule using the long-run elasticity. This elasticity might be substantially smaller than the short-run elasticity for reasons described above. The key parameter that determines the difference in the elasticities is the effect of a change in the dividend tax rate on γ. Empirically, this parameter could be estimated by examining the effect of dividend tax changes on the structure of executive contracts (e.g. pay-for-performance measures) and degree of monitoring by corporate boards.

5 Conclusion

This paper has proposed a simple agency model of dividend taxation, in which managers have a higher preference for retained earnings than shareholders, and hence pay too few dividends and overinvest. In this setting, a dividend tax cut leads to an increase in dividend payments and dividend initiations in the short run. Firms where top executives place more weight on profit maximization are more likely to raise dividends in response to a tax cut. In the short run, revenue-neutral dividend tax cuts can potentially yield substantial efficiency gains by extracting money from firms that are inefficiently hoarding funds.

In the long run, however, a dividend tax cut could lead shareholders to reoptimize their contracts with managers. The dividend tax cut therefore has smaller effects on dividend payments in the long-run, and the negative externality effects of dividend taxation are internalized by the shareholders. As a result, the efficiency costs of dividend taxation in the long-run are smaller than in the short-run. In future work, it would be interesting to test empirically whether there was a weakening in monitoring of managers or in incentives provided to managers following the 2003 dividend tax cut.

References

Allen, Franklin and Roni Michaely, 2003, "Payout Policy," in Constantinides, George, Milton Harris and Rene Stulz eds. *Handbook of the Economics of Finance* (Amsterdam: North-Holland)

Auerbach, Alan, 1979, "Wealth maximization and the cost of capital," *Quarterly Journal of Economics* 93, 433-46.

Auerbach, Alan and James Hines (2003), "Taxation and Economic Efficiency," in Auerbach, Alan and Martin Feldstein eds. *Handbook of Public Economics*, Volume 3, (Amsterdam: North-Holland)

Auerbach, Alan, 2003, "Taxation and Corporate Financial Policy," in Auerbach, Alan and Martin Feldstein eds. *Handbook of Public Economics*, Volume 3, (Amsterdam: North-Holland)

Auerbach, Alan and Kevin Hassett, 2003, "On the Marginal Source of Investment Funds," *Journal of Public Economics* 87(1), 205-232.

Auerbach, Alan and Kevin Hassett, 2004, "The Effects of the 2003 Dividend Tax Cut on Stock Prices: An Event Study," unpublished working paper, University of California at Berkeley.

Bernhardt, Dan, J. Fiona Robertson and Ray Farrow, 1994, "Testing Dividend Signaling Models," Working Paper, Queen's University.

Bernheim, Douglas and Adam Wantz (1995), "A Tax-Based Test of the Dividend Signalling Hypothesis", *American Economic Review* 85(3), 532-551.

Bradford, David F., 1981, "The incidence and allocation effects of a tax on corporate distributions," *Journal of Public Economics* 15, 1-22.

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni Michaely (2004), "Payout policy in the 21st century", forthcoming *Journal of Financial Economics*.

Brown, Jeffrey, Nellie Liang, and Scott Weisbenner (2006) "Executive Financial Incentives and Payout Policy: Firm Responses to the 2003 Dividend Tax Cut", forthcoming in *Journal of Financial Economics*.

Chetty, Raj and Emmanuel Saez (2005), "Dividend Taxes and Corporate Behavior: Evidence from the 2003 Dividend Tax Cut", *Quarterly Journal of Economics* 120(3), 791-833.

Chetty, Raj and Emmanuel Saez (2006), "The Effects of the 2003 Dividend Tax Cut on Corporate Behavior: Interpreting the Evidence", *American Economic Review* 96(2), 124-129.

Feldstein, Martin, (1970) "Corporate taxation and Dividend Behavior," *Review of Economic Studies* 37, 57-72.

Fenn, George W. and Nellie Liang, 2001, "Corporate payout policy and managerial stock incentives", *Journal of Financial Economics*, 60(1), 45-72.

Gordon, Roger and Martin Dietz, 2006, "Dividends and Taxes", NBER Working Paper No. 12292.

Goulder, Lawrence and Roberton Williams (2003) "The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution," *Journal of Political Economy*, 111, 898-927.

Harberger, A.C., 1962, "The incidence of the corporation income tax", *Journal of Political Economy* 70, 215-240.

Harberger, A.C., 1966, Efficiency effects of taxes on income from capital, in: M. Krzyzaniak, ed., *Effects of corporation income tax* (Wayne State University Press, Detroit).

Jensen, Michael C, 1986, "Agency costs of free cash flow, corporate finance and takeovers," *American Economic Review* 76, 323-329.

Jensen, Michael C. and William H. Meckling , 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3(4), 305-360.

King, Mervin A., 1977, *Public Policy and the Corporation,* (Chapman and Hall, London).

Lambert, Richard, William Lanen and David Larcker, 1989, "Executive stock option plans and corporate dividend policy," *Journal of Financial and Quantitative Analysis* 24, 409- 425.

LaPorta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 2000, "Agency problems and dividend policies around the world," *Journal of Finance* 55(1), 1-33.

Nam, Jouahn, Jun Wang, and Ge Zhang (2004), "The Impact of Dividend Tax Cut and Managerial Stock Holdings on Firm's Dividend Policy", unpublished mimeograph, Baruch College.

Perez-Gonzalez, Francisco, 2003, "Large shareholders and dividends: Evidence from U.S. tax reforms", Working paper, Columbia University.

Poterba, James, 2004, "Corporate Payout Policy", *American Economic Review*, 94(2), 171-175.

Poterba, James and Lawrence Summers, 1985 "The Economic Effects of Dividend Taxation", in E. Altman and M. Subrahmanyam, eds., *Recent Advances in Corporation Finance* (Homewood, IL: Dow Jones-Irwin), 227-284.

Rajan, Raghuram, Henri Servaes, and Luigi Zingales, 2000, "The cost of diversity; The diversification discount and inefficient investment," *Journal of Finance* 55, 35-80.

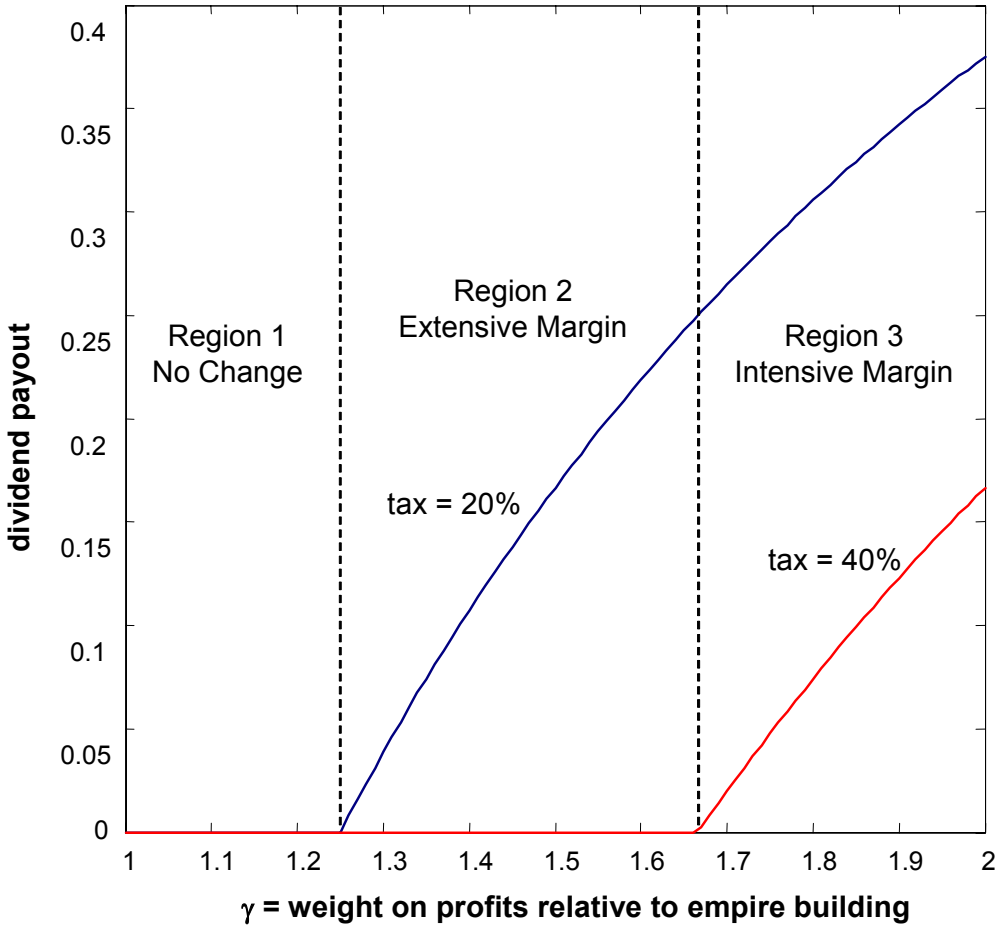
Sandmo, Agnar (1975), "Optimal Taxation in the Presence of Externalities", *Swedish Journal of Economics*, 77, 86-98.

Scharfstein, David and Jeremy Stein, 2000, "The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment," *Journal of Finance* 55 (6).

Shleifer, Andrei and Robert Vishny, "A Survey of Corporate Governance," Nobel Symposium on Law and Finance, August, 1995, *Journal of Finance*, June 1997.

Sinn, Hans-Werner, (1991) "The Vanishing Harberger Triangle," *Journal of Public Economics*, 45(3), 271-300.

Figure 1: Effect of Dividend Tax Cut on Dividend Payouts



Notes: This figure shows dividend payouts in the first period as a function of γ, the weight on profits relative to empire building in the firm's objective function. Dividend payouts are shown under two tax regimes to illustrate how the effect of a dividend tax cut varies across firms with different levels of γ.