

Dividend and Corporate Taxation in an Agency Model of the Firm*

Raj Chetty, Harvard University and NBER
and
Emmanuel Saez, UC Berkeley and NBER

August 12, 2009

Abstract

Recent empirical studies of dividend taxation have found that: (1) dividend tax cuts cause large, immediate increases in dividend payouts, and (2) the increases are driven by firms with high levels of share ownership among top executives or the board of directors. These findings are inconsistent with existing theories of dividend taxation. We show that an agency model in which managers and shareholders have conflicting interests explains the evidence. In this model, dividend taxation encourages managers to divert earnings into unproductive projects instead of paying out profits to shareholders and therefore creates a first-order deadweight cost. In contrast, corporate taxes do not distort the manager's choice between payouts and unproductive investment and may only create second-order efficiency costs. Corporate income taxation may therefore be a more efficient way to generate revenue than dividend taxation, challenging existing intuitions based on neoclassical models.

*Raj Chetty, chetty@fas.harvard.edu, Harvard University, Department of Economics, Cambridge, MA 02138. Emmanuel Saez, saez@econ.berkeley.edu, University of California - Berkeley, Department of Economics, 549 Evans Hall #3880, Berkeley, CA 94720. We thank Alan Auerbach, Martin Feldstein, Roger Gordon, Kevin Hassett, James Poterba, two anonymous referees, and numerous seminar participants for very helpful comments. Gregory Bruich, Joseph Rosenberg, and Ity Shurtz provided outstanding research assistance. Financial support from National Science Foundation grants SES-0134946 and SES-0452605 is gratefully acknowledged.

1 Introduction

The 2003 dividend tax reform in the U.S. has sparked a new wave of research on the effects of dividend and corporate taxation (Chetty and Saez (2005), Brown et al. (2007), Nam et al. (2004)). Chetty and Saez document four empirical results: (1) Regular dividends rose sharply after the 2003 tax cut, with an implied net-of-tax elasticity of dividend payments of 0.75. (2) The response was very rapid – total dividend payouts rose by 20% within one year of enactment – and was stronger among firms with high levels of accumulated assets. (3) The response was much larger among firms where top executives owned a larger fraction of outstanding shares (see also Brown et al. (2007) and Nam et al. (2004)). (4) The response was much larger among firms with large shareholders on the board of directors.

It is difficult to reconcile these four findings with either of the two leading theories of corporate taxation – the “old view” (Harberger 1962, 1966, Feldstein 1970, Poterba and Summers 1985) and the “new view” (Auerbach 1979, Bradford 1981, King 1977). The increase in dividends appears to support the old view because dividends should not respond to permanent dividend tax changes under the new view.¹ However, the increase in dividend payments is too rapid to be explained by the supply-side investment mechanism of the old view model.² The rapid dividend payout response could potentially be explained by incorporating a signaling value for dividends as in Poterba and Summers (1985) or Bernheim (1991).³ However, neither signaling models nor the standard old and new view models directly predict findings (3) and (4) on the cross-sectional heterogeneity in the dividend payout response by firm ownership structure.⁴

In this paper, we propose a simple alternative model of dividend and corporate income taxation that matches the four empirical findings based on the agency theory of the firm (Jensen and Meckling 1976). The critical feature of the model is a divergence between the preferences of

¹One way of reconciling the dividend increase with the new view is if the tax cut was perceived as temporary by firms. However, Auerbach and Hassett (2007) document that the share prices of immature firms that are predicted to pay dividends in the future rose when the reform was announced, suggesting that firms perceived the tax cut as fairly permanent.

²Poterba’s (2004) estimates using an old view model implied that the 2003 tax reform would increase dividend payments by 20 percent in the long run, but that only a quarter of the long-run effect would occur within three years after the tax cut.

³There is debate in the corporate finance literature about the signal content of dividends. Conditional on information available at time t , dividend increases have little predictive power for future earnings (see e.g., Bernartzi et al. 1997, Grullon et al. 2005).

⁴The empirical evidence is also not fully explained by Sinn’s (1991) “life cycle” model in which firms progress from the old view to the new view. In that model, the payout response should be smaller among firms with higher levels of accumulated assets, but the data exhibit the opposite pattern.

managers and shareholders. We model this divergence as arising from perks and pet projects, although the underlying source of the conflict between managers and shareholders does not matter for our analysis. Shareholders can provide incentives to managers to invest and pay out dividends through costly monitoring and pay-for-performance. Only the large shareholders of the firm choose to monitor the firm in equilibrium (Shleifer and Vishny 1986, 1997). In this model, a dividend tax cut leads to an immediate increase in dividend payments because it increases the manager's preference for dividends relative to the pet project and increases the amount of monitoring by large shareholders. Firms where managers place more weight on profit maximization – either because the manager owns a large number of shares or because there are more large shareholders – are more likely to increase dividends in response to a tax cut.

After showing that the positive predictions of the agency model fit the recent evidence on dividend taxation, we characterize its implications for the efficiency costs of dividend and corporate taxation by deriving empirically implementable formulas for excess burden. We obtain two results that challenge intuitions from existing neoclassical models. First, dividend taxes create a deadweight cost even if the marginal source of investment is retained earnings by distorting the tradeoff between pet project investment and dividend payouts. Second, if the contract between shareholders and the manager is second-best inefficient – as is the case in a model with diffuse shareholders – dividend taxation creates a first-order efficiency cost. In contrast, the corporate tax may generate only standard second-order efficiency costs because it does not amplify the manager's incentive to hoard cash for pet projects.⁵ This suggests that corporate taxes may be a more efficient way to generate revenue than dividend taxes. Indeed, our analysis suggest that a Pigouvian dividend *subsidy* would be desirable to correct the negative externality created by agency problems in firms.

The most important limitation of our analysis is that it does not explicitly model share repurchases, which give firms a way to return money to shareholders without paying dividend taxes. In the appendix, we extend the model to permit costly share repurchases, as in Poterba and Summers (1985). The formulas for excess burden remain the same, but the first-order agency-related term depends upon the elasticity of total payouts (share repurchases plus dividends) with respect to taxes. Intuitively, dividend taxes do not have first-order efficiency costs if they simply induce substitution between dividends and repurchases without changing pet

⁵The corporate tax does not always have second-order efficiency costs in our model; if it distorts the manager's contract, it too may generate first-order efficiency costs.

project investment. Note that this extension of our analysis relies on a reduced form model for share repurchases and hence does not explain the puzzle of why firms pay dividends even though dividends are tax-disadvantaged. Understanding the microeconomic foundations of the cost of share repurchases is an issue of great importance for future work, independent of its potential implications for taxation.

This paper is related to two contemporaneous theoretical studies motivated by evidence from the 2003 dividend tax cut. Gordon and Dietz (2006) contrast the effects of dividend taxation in new view, signaling, and agency models and conclude that the agency model is most likely to fit the empirical evidence. The central difference between our model and Gordon and Dietz's agency model is in the assumption about which agent sets the firm's dividend policy. Gordon and Dietz assume that dividend payout decisions are made by shareholders, whereas we assume that they are made by management. This leads to different results in both the positive and efficiency analysis. Gordon and Dietz's model does not directly predict a link between executive or board share ownership and behavioral responses to dividend taxation. Taxing dividends does not create a first-order distortion in their model, since dividends are always set at the second-best efficient level by shareholders. Their model does, however, generate the empirically validated prediction that dividend policies change rarely over time, which our model does not produce. Our model and Gordon and Dietz's analysis should therefore be viewed as complementary efforts to explain different aspects of dividend policies.

A second recent study is Korinek and Stiglitz (2009), who build on Sinn's (1991) model to analyze the effects of temporary changes in dividend tax rates. They incorporate financing constraints and establish new results on intertemporal tax arbitrage opportunities for firms. In contrast with our model, Korinek and Stiglitz assume that retained earnings are allocated efficiently by the manager. As a result, they obtain the new view neutrality result that permanent dividend tax policy changes have no effects on economic efficiency.

The remainder of the paper is organized as follows. In section 2, we present a neoclassical two period model that nests the old and new views as a benchmark. In section 3, we introduce agency problems into the model and characterize manager and shareholder behavior. In section 4, we characterize behavioral responses to dividend taxation and compare the agency model's predictions with the recent empirical evidence. In section 5, we analyze the efficiency consequences of dividend and corporate taxation. Section 6 concludes.

2 The Old and New Views in a Two Period Model

We begin with a neoclassical two period model that nests the old and new views and serves as a point of departure for our agency analysis. Consider a firm that has initial cash holdings of X at the beginning of period 0. These cash holdings represent profits from past operations.⁶ The firm can raise additional funds by issuing equity (E). The firm's manager can do two things with the firm's cash holdings: pay out dividends or invest the money in a project that yields revenue in the next period. Let I denote the level of investment and $D = X + E - I$ the firm's dividend payment in period 0. In period 1, the firm generates net profits of $f(I)$, where f is a strictly concave function.⁷ The firm then closes and returns its net-of-tax profits and principal to shareholders.

The firms' profits are subject to two types of taxes. First, the firm pays a corporate tax at rate t_c on its net profits in period 1, so that net-of-corporate-tax profits are $(1 - t_c)f(I)$. Second, it pays a dividend tax at rate t_d on distributed profits in all periods. However, the principal invested by shareholders is not subject to the dividend tax (E).⁸ Hence, the net-of-tax payout in period 0 is $(1 - t_d)D$ and the net-of-tax payout in period 1 is $(1 - t_d)[(1 - t_c)f(X + E - D) + X - D] + E$. Investors can also purchase a government bond that pays a fixed, untaxed interest rate of $r > 0$ (which is unaffected by the dividend tax rate).⁹

The manager's objective is to choose the level of equity issues and dividends (and investment) that maximize the value of the firm:

$$\max_{D,E} V = (1 - t_d)D - E + \frac{(1 - t_d)[(1 - t_c)f(X + E - D) + X - D] + E}{1 + r} \quad (1)$$

To characterize these choices, it is useful to distinguish between two cases: (1) A cash-rich firm, which has retained profits X such that its net-of-corporate-tax marginal return $(1 - t_c)f'(X) \leq r$ and (2) a cash-constrained firm, which has cash X such that $(1 - t_c)f'(X) > r$. The "new view" model considers firms of the first type, while the "old view" pertains to firms of the second-type.

⁶We can allow part of the existing cash holdings X to represent the principal of shareholders without any impact on the analysis as long as firms cannot return the principal before liquidation and firms do not choose to distribute all their past profits in period 0.

⁷The gross production function is $F(I) = f(I) + I$; $f(I)$ denotes profits net of the depreciation of capital used for production.

⁸In the United States, distributed profits are considered dividends for tax purposes, but returning shareholders' principal is not considered a dividend.

⁹Throughout this paper, we abstract from general-equilibrium effects through which changes in t_d may affect the equilibrium rate of return, r .

Cash-Rich Firms – The New View. First observe that the firm will never set $E > 0$ and $D > 0$ simultaneously. If a firm both issued equity and paid dividends, it could strictly increase its value V by reducing both E and D by \$1 and lowering its tax bill by $\$t_d r/(1+r)$. Now consider the marginal value of issuing equity when $D = 0$ for the cash-rich firm:

$$\frac{\partial V}{\partial E}(D = 0) = -1 + \frac{(1-t_d)(1-t_c)f'(X) + 1}{1+r} = \frac{(1-t_d)(1-t_c)f'(X) - r}{1+r} \leq 0$$

This expression implies that a cash rich firm optimally sets $E^* = 0$. The optimal choice of dividends satisfies the first order condition

$$(1-t_c)f'(X - D^*) = r$$

Cash rich firms invest to the point where the net-of-corporate-tax marginal product of investment $f'(I)$ equals the return on investment in the bond, r . Increases in the corporate tax rate reduce the level of investment, increase period 0 dividend payments, and reduce period 1 dividend payments. However, the dividend tax rate t_d has no impact on dividend payments and investment levels. This is the classic “new view” dividend tax neutrality result (Auerbach 1979, Bradford 1981, King 1977). The source of this result is transparent in the two period case: the $(1-t_d)$ term factors out of the value function in equation (1) when $E = 0$. Dividend taxation has no impact on the behavior of cash-rich firms because they must pay the dividend tax regardless of whether they pay out profits in the current or next period. In contrast, the corporate tax changes the relative price of paying out dividends immediately and investing to earn further profits, and therefore distorts behavior.

Cash-Constrained Firms – The Old View: Now consider a firm with X such that $(1-t_c)f'(X) > r$. The marginal value of paying dividends when $E = 0$ for this “cash constrained” firm is

$$\frac{\partial V}{\partial D}(E = 0) = 1 - t_d - \frac{1-t_d}{1+r}[(1-t_c)f'(X) + 1] = (1-t_d)\frac{r - (1-t_c)f'(X)}{1+r} < 0.$$

A cash-constrained firm does not pay dividends in the first period because its marginal product of investment exceeds the interest rate. This firm therefore invests all the cash it has: $I = X + E$. The optimal choice of equity issues is given by

$$E^* = 0 \text{ if } (1-t_d)(1-t_c)f'(X) < r \tag{2}$$

$$(1-t_d)(1-t_c)f'(X + E^*) = r \text{ if } (1-t_d)(1-t_c)f'(X) \geq r \tag{3}$$

These conditions show that firms which finance their marginal dollar of investment from new equity issues invest to the point where the marginal net-of-dividend and corporate tax return to investment equals the return on investment in the bond, r . Firms that have X sufficiently large so that $(1 - t_d)(1 - t_c)f'(X) < r$ have a net-of-tax return below the interest rate for the first dollar of equity. These “medium cash” firms choose the corner solution of no equity (and no dividends) because of the tax wedge.

Unlike for cash-rich firms, the dividend tax distorts the behavior of low cash firms. Implicit differentiation of (3) shows that increases in t_d reduce equity issues and investment ($\partial I^*/\partial t_d < 0$, $\partial E^*/\partial t_d < 0$). This is because the $(1 - t_d)$ term does *not* factor out of the value function in equation (1) when $D = 0$ and $E > 0$. Intuitively, a dividend tax increase lowers the marginal product of investment but does not affect the price of investment for cash-constrained firms. Firms therefore reduce investment, issue less equity, and pay fewer dividends in period 2, the classic “old view” predictions (Poterba and Summers 1985). Corporate taxes produce the same effects because they affect the value of cash-constrained firms in exactly the same way as dividend taxes. Note that dividend payments are *not* affected by tax changes in the short-run. Following a dividend or corporate tax change, investment and equity issues respond immediately (period 0), and dividends change only when the additional investment pays off (period 1).

Efficiency Costs. Finally, we characterize the efficiency cost of introducing a dividend and corporate tax for the two types of firms. Let

$$P_d = D + [(1 - t_c)f(I) + X - D]/(1 + r)$$

denote the dividend tax base, i.e. the total dividend payout over the two periods, and

$$P_c = f(I)/(1 + r)$$

denote the corporate tax base. Total surplus in the economy is $W = V + t_d P_d + t_c P_c$. Using the envelope conditions, differentiating (1) yields $dV/dt_d = -P_d$ and $dV/dt_c = -(1 - t_d)P_c$. Therefore, we obtain the standard Harberger-type formulas for marginal deadweight burden:

$$\frac{dW}{dt_d} = t_d \frac{dP_d}{dt_d} + t_c \frac{dP_c}{dt_d} \tag{4}$$

$$\frac{dW}{dt_c} = t_d \left(\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c} \right) + t_c \frac{dP_c}{dt_c} \tag{5}$$

where $\frac{\partial P_d}{\partial t_c} = -P_c$ denotes the mechanical effect of increasing t_c on the firm’s payout and $\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c}$ thus measures the distortion in dividend payments created by the corporate tax due to

behavioral responses. Equations (4) and (5) apply to both cash-rich and cash-constrained firms. To obtain further insight into the key determinants of excess burden, it is helpful to consider the old and new view firms separately.

For new view firms, dividend taxes do not distort behavior: $\frac{dP_d}{dt_d} = \frac{dP_c}{dt_d} = 0$. In addition, because these firms choose D and I to maximize P_d itself, $\frac{dP_d}{dt_c} = -P_c$. Hence, for new view firms, the formulas for excess burden simplify to:

$$\frac{dW}{dt_d} = 0 \quad (6)$$

$$\frac{dW}{dt_c} = t_c \frac{dP_c}{dt_c} \quad (7)$$

Intuitively, an increase in t_c does not distort total dividend payments because the marginal reduction in period 1 dividends is cancelled out by the marginal increase in period 0 dividends for a profit maximizing firm. As a result, the only distortionary effect of the corporate tax comes from its effect on the corporate tax base itself.

For old view firms, dividend and corporate taxes both distort the return to investment in the same way, implying $\frac{dP_c}{dt_d} = \frac{dP_c}{dt_c}$. Because old view firms pay dividends only in period 1, the effects of t_d and t_c on the dividend tax base are fully determined by their effects on profits, which equal the corporate tax base: $\frac{dP_d}{dt_d} = (1 - t_c) \frac{dP_c}{dt_d}$ and $\frac{dP_d}{dt_c} = -P_c + (1 - t_c) \frac{dP_c}{dt_c}$. Combining these results, we obtain

$$\frac{dW}{dt_d} = \frac{dW}{dt_c} = (t_c + t_d - t_c t_d) \frac{dP_c}{dt_c} \quad (8)$$

Intuitively, both dividend and corporate taxes reduce the profits earned by old view firms in the same way. The total revenue obtained from the two taxes is $(t_d(1 - t_c) + t_c)P_c$, leading to the formula in (8).

This analysis yields two general lessons about efficiency costs that we will revisit below. First, dividend taxation has an efficiency cost only for firms which finance investment from new equity issues, whereas corporate taxation has an efficiency cost for both types of firms. Because most investment is accounted for by firms with large amounts of retained earnings, this leads to the view that dividend taxes are a more efficient instrument for raising tax revenue than corporate taxes. Second, when one starts from a situation with no taxes, the introduction of a small corporate tax has a second-order (i.e., small) efficiency cost, as does the introduction of a small dividend tax for old view firms.

The main predictions of the old and new view models are summarized on the left side of Table 1. The central assumption underlying these results is that firms' managers choose policies solely to maximize firm value. This assumption contrasts with the modern corporate finance literature, which emphasizes the tension between executives' and shareholders' interests in explaining corporate behavior and payout policies. The next section incorporates these considerations into the model.

3 An Agency Model of Firm Behavior

In the remainder of the paper, we restrict attention to cash-rich firms, i.e. those with $(1 - t_c)f'(X) > r$. Firms with $(1 - t_c)f'(X) < r$ never pay dividends. Since our goal is to construct a model consistent with recent evidence on dividend payout behavior, it is the behavior of cash-rich firms that is of greatest interest.¹⁰ The predictions of the agency model are summarized in the right half of Table 1.

3.1 Setup

The source of agency problems in corporations is a divergence between the objectives of managers and shareholders. We model the source of the divergence as a "pet project" that generates no profits for shareholders but yields utility to the manager. In particular, the manager can now do three things with the firm's cash X : pay out dividends D , invest I in a "productive" project that yields net profits $f(I)$ for shareholders, or invest J in a pet project that gives the manager private benefits of $g(J)$.¹¹ Assume that both f and g are strictly concave.

The function g should be interpreted as a reduced-form means of capturing divergences between the managers' and shareholders' objectives. For example, the utility $g(J)$ may arise from allocation of funds to perks, tunneling, a taste for empire building, or a preference for projects that lead to a "quiet life".¹² While there is debate in corporate finance about which of

¹⁰The working paper version (Chetty and Saez 2007) extends the efficiency analysis to the case with equity issues.

¹¹The manager returns the capital used for investment in the pet project (J) back to the shareholders in period 1.

¹²There is a large literature in corporate finance providing evidence for agency models. Recent examples include Rajan et al. (2000), Scharfstein and Stein (2000), and Bertrand and Mullainathan (2003). Empirical studies have also provided support for the agency theory as an explanation of why firms pay dividends (see e.g., Lang and Litzenberger 1989, Christie and Nanda 1994, LaPorta et al. 2000, Fenn and Liang 2001, Desai, Foley, and Hines 2007).

these elements of $g(J)$ is most important, the underlying structure that determines $g(J)$ does not matter for our analysis.

Manager's Objective. The agency problem arises because shareholders cannot observe real investment opportunities and have to let the manager choose I , J , and D . Shareholders push managers toward profit maximization through two channels: incentive pay and monitoring. Incentive pay is achieved through features of the manager's compensation contract such as share grants and bonuses. We model such financial incentives by assuming that the shareholders compensate the manager with a fraction α of the shares of the company. Monitoring effectively reduces the manager's utility from the pet project because it increases the probability that pet projects are detected and penalized. We model monitoring by assuming that $\gamma \geq 0$ units of monitoring reduces the utility the manager derives from the pet project from $g(J)$ to $g(J)/(1 + \gamma)$.

Given the shareholders' choice of α and γ , the manager chooses I and D to maximize

$$V^M = \alpha(1 - t_d) \cdot \left[D + \frac{(1 - t_c)f(I) + X - D}{1 + r} \right] + \frac{1}{1 + r} \cdot \frac{g(J)}{1 + \gamma} \quad (9)$$

subject to the constraint $I + J + D = X$. Monitoring increases the weight managers put on profits relative to the pet project by a factor $1 + \gamma$. Let $\omega = \alpha(1 - t_d)(1 + \gamma)$ denote the relative weight that managers place on profits. When ω is low, the manager has little stake in the profits of the firm and is therefore tempted to retain excess earnings and invest in the pet project.¹³

Shareholders' Objectives. Next, we model how shareholders choose the level of monitoring (γ). Following Shleifer and Vishny (1986), each shareholder who chooses to monitor the firm incurs a cost of monitoring, whereas the benefits of better manager behavior accrue to all shareholders. There are N shareholders, each of whom owns a fraction α_i of the shares (so that $\sum_1^N \alpha_i = 1 - \alpha$). Each shareholder chooses a level of monitoring $\gamma_i \geq 0$. The total monitoring level is $\gamma = \sum \gamma_i$.

Shareholders incur a fixed cost k if they monitor the firm, i.e. if they set $\gamma_i > 0$. In addition, they pay a convex and increasing variable cost $c(\gamma_i)$ to do γ_i units of monitoring, where $c'(\gamma_i = 0) = 0$. Each shareholder chooses γ_i to maximize his net profits

$$V_i = (1 - t_d)\alpha_i \cdot \left[D + \frac{(1 - t_c)f(I) + X - D}{1 + r} \right] - k \cdot 1(\gamma_i > 0) - c(\gamma_i) \quad (10)$$

¹³The pet project $g(J)$ is presumably small relative to the firm's productive project $f(I)$. However, ω is also likely to be small in large publicly traded corporations, where executives own a small fraction of total shares and diffuse shareownership can lead to a low level of monitoring. Combining a small pet project $g(J)$ with a small ω can make the manager deviate substantially from the shareholders' optimal investment level.

where $1(\gamma_i > 0)$ is an indicator function. In the Nash equilibrium, γ is determined such that each shareholder's choice of γ_i is a best response to the others' behavior. It is well known from the public goods literature that monitoring will be below the social optimum (i.e., the level that would be chosen if one shareholder owned the entire firm) in equilibrium.¹⁴ There is a threshold level $\bar{\alpha}$ such that small shareholders with $\alpha_i < \bar{\alpha}$ will not monitor the firm, while large shareholders with $\alpha_i > \bar{\alpha}$ do monitor. Since the number of large shareholders is typically small, it is natural to assume that these individuals cooperatively choose the level of monitoring γ by forming a "board of directors" that is in charge of monitoring the manager. Let α_B denote the total fraction of shares held by the board of directors. The board chooses γ to maximize its joint profits net of monitoring costs:

$$V^B = (1 - t_d)\alpha_B \cdot \left[D(\omega) + \frac{(1 - t_c)f(I(\omega)) + X - D(\omega)}{1 + r} \right] - c(\gamma) \quad (11)$$

Ownership Structure. To close the model, we must specify how the firm's ownership structure (α and α_B) is determined. We draw a distinction between the short-run positive analysis and the long-run efficiency analysis in the specification of the firm's ownership structure. In the short run, ownership structures are relatively stable in practice.¹⁵ Since the evidence on dividend payout behavior we are attempting to explain concerns the effect of the 2003 dividend tax reform within a two year horizon, we take α and α_B as fixed in our positive analysis. In the longer run, and particularly when new firms are started, α and α_B are presumably endogenous to the tax regime. In the efficiency analysis in section 5, we model how α and α_B are determined. Allowing for endogenous ownership structure is particularly important in the efficiency analysis because the deadweight cost of taxation depends critically on how α and α_B are determined.

3.2 Manager Behavior

We now characterize the manager's behavior as a function of his weight on profits $\omega = \alpha(1 - t_d)(1 + \gamma)$. The manager chooses I and D to

$$\max_{I, D \geq 0} \omega \left[D + \frac{(1 - t_c)f(I) + X - D}{1 + r} \right] + \frac{g(X - I - D)}{1 + r}.$$

¹⁴The Coasian solution (Coase 1960) is unlikely to emerge in this setting because of transaction costs in coordinating many small shareholders.

¹⁵Chetty and Saez (2007) present evidence that managerial and board share ownership is much more stable than dividend payments in the three years after the 2003 dividend tax cut.

Assume that $g'(0) > \omega f'(X)$, which guarantees an interior optimum in investment behavior. Then I and D are determined by the following first-order conditions:

$$(1 - t_c)\omega f'(I) = g'(X - I - D) \quad (12)$$

$$\omega \cdot r \leq g'(X - I - D) \text{ with strict equality iff } D > 0 \quad (13)$$

Let $D(\omega)$ and $I(\omega)$ denote the dividend and investment choices of the manager as a function of ω . To characterize the properties of these functions, define the threshold

$$\bar{\omega} = \frac{g'(X - I^*)}{r} > 0,$$

where I^* denotes the optimal investment level from the shareholders' perspective: $(1 - t_c)f'(I^*) = r$. Note that $\bar{\omega}$ is a monotonic decreasing function of X . We therefore label firms with $\omega > \bar{\omega}$ as “very high cash” firms and those with $\omega < \bar{\omega}$ but $(1 - t_c)f'(X) > r$ as “high cash” in Table 1.

Lemma 1 $D(\omega)$ and $I(\omega)$ follow threshold rules:

- If $\omega \leq \bar{\omega}$ then $D(\omega) = 0$ and $I(\omega)$ is chosen such that $(1 - t_c)\omega f'(I) = g'(X - I)$.
- If $\omega > \bar{\omega}$ then $I(\omega) = I^*$ and $D(\omega) > 0$ is chosen such that $\omega \cdot r = g'(X - I^* - D)$.

Proof. Consider $\omega \leq \bar{\omega}$. Suppose the firm sets $D > 0$. Then the first order conditions (13) and (12) imply that $(1 - t_c)\omega f'(I) = r$ and hence $I = I^*$. This implies $\omega \cdot r = g'(X - I^* - D) > g'(X - I^*) = \bar{\omega} \cdot r$, contradicting the supposition. Hence $\omega \leq \bar{\omega} \Rightarrow D(\omega) = 0$.

Now consider $\omega > \bar{\omega}$. Suppose the firm sets $D = 0$. Then the first order conditions (12) and (13) imply that $(1 - t_c)\omega f'(I) \geq r$ and hence $I \leq I^*$. This implies $\omega \cdot r \leq g'(X - I) \leq g'(X - I^*) = \bar{\omega} \cdot r$, contradicting the supposition. Hence $\omega > \bar{\omega} \Rightarrow D(\omega) > 0$, and (13) yields the desired expression for $D(\omega)$. QED.

Figure 1 illustrates the threshold rules that the manager follows by plotting $D(\omega)$, $I(\omega)$, and $J(\omega)$ with quadratic production functions when $t_c = 0$. When ω is below the threshold value $\bar{\omega}$, the marginal value of the first dollar of dividends is negative in the manager's objective function. The optimal level of dividends is therefore zero, the corner solution. Intuitively, if managers have a sufficiently weak interest in profit maximization, they retain as much money as possible for pet projects and do not pay dividends. For ω above $\bar{\omega}$, further increases in

the weight on profits ω lead to increases in dividends and reductions in pet investment on the intensive margin:

$$D'(\omega) = -\frac{r}{g''(J(\omega))} > 0 \text{ for } \omega > \bar{\omega} \quad (14)$$

Now consider the manager's investment choice. When $\omega \leq \bar{\omega}$, the manager pays no dividends, and splits retained earnings between investment in the profit-generating project and the pet project. He chooses I to equate his private marginal returns of investing in the two projects, as in equation (12). An increase in ω increases productive investment I and reduces pet investment J :

$$I'(\omega) = -\frac{(1-t_c)f'(I(\omega))}{(1-t_c)\omega f''(I(\omega)) + g''(X-I(\omega))} > 0 \text{ for } \omega < \bar{\omega} \quad (15)$$

Once $\omega > \bar{\omega}$, the manager has enough cash to pay a dividend to shareholders. He sets the investment level such that $(1-t_c)f'(I) = r$, implying that I is fixed at I^* for $\omega > \bar{\omega}$. Intuitively, the manager would only pay a dividend if his private return to further investment in the profitable project was below the interest rate. Since the tradeoff between dividends and profitable investment is the same for managers and shareholders, the manager only begins to pay a dividend once he has reached the optimal level of investment from the shareholder's perspective, I^* .

3.3 Board Behavior

In the short run, the board's only decision is to choose the level of monitoring. The board takes α_B as fixed and chooses γ to maximize

$$V^B = (1-t_d)\alpha_B \cdot P_d(\omega) - c(\gamma) \quad (16)$$

where $P_d(\omega) = D(\omega) + [(1-t_c)f(I(\omega)) + X - D(\omega)]/(1+r)$ denotes the firm's total payout as a function of ω . Because both D and I are (weakly) increasing in ω , $P_d(\omega)$ is also (weakly) increasing in ω . We have $d\omega/d\gamma = \alpha(1-t_d)$. Hence, the first order condition with respect to γ is:

$$c'(\gamma) = (1-t_d)\alpha_B \cdot P'_d(\omega) \cdot \alpha(1-t_d). \quad (17)$$

Intuitively, the board chooses γ such that the marginal increase in the board's share of profits by raising ω is offset by the marginal cost of monitoring. The second-order condition for an interior maximum is:

$$(1-t_d)\alpha_B \cdot P''_d(\omega) \cdot [\alpha(1-t_d)]^2 - c''(\gamma) < 0. \quad (18)$$

Since $c'(\gamma = 0) = 0$ by assumption, the optimal γ is always in the interior, and hence (18) must be satisfied at the optimal level of monitoring $\gamma(t_d)$.¹⁶ This second-order condition turns out to be useful for the comparative statics analysis below.

4 Positive Analysis: Effects of Dividend Taxation

We now characterize the effect of dividend tax changes on firm behavior to show that the agency model explains the four empirical findings discussed in the introduction as well as other evidence.

For any variable $x \in \{D, I, J\}$,

$$\frac{dx}{dt_d} = \frac{dx}{d\omega} \frac{d\omega}{dt_d}$$

because t_d affects the manager's objective only through his weight on profits ω . We characterized $\frac{dx}{d\omega}$ in the previous section. As $\omega = (1 + \gamma)\alpha(1 - t_d)$, we have

$$\frac{d\omega}{dt_d} = -\alpha(1 + \gamma) + \alpha(1 - t_d) \cdot \frac{d\gamma}{dt_d} \quad (19)$$

To calculate $\frac{d\gamma}{dt_d}$, implicitly differentiate the board's first-order-condition for γ in (17) to obtain:

$$\frac{d\gamma}{dt_d} = -\frac{\alpha_B[2\alpha(1 - t_d)P'_d(\omega) + [\alpha(1 - t_d)]^2(1 + \gamma)P''_d(\omega)]}{c'' - P''_d \cdot \alpha_B(1 - t_d) \cdot [\alpha(1 - t_d)]^2}. \quad (20)$$

Combining (19) and (20) leads to:

$$\frac{d\omega}{dt_d} = -\frac{2\alpha_B\alpha^2(1 - t_d)^2P'_d(\omega) + \alpha(1 + \gamma)c''}{c'' - P''_d \cdot \alpha_B(1 - t_d) \cdot [\alpha(1 - t_d)]^2} < 0. \quad (21)$$

The board's second-order condition for γ in (18) implies that the denominator of this expression is positive. The numerator is positive because P_d is increasing in ω and c is convex. Equation (21) therefore shows that a reduction in the dividend tax rate leads to an increase in the weight ω that managers put on profits through two channels. First, a decrease in t_d mechanically increases the net stake $(1 - t_d)\alpha$ that the manager has in the firm, effectively by reducing the government's stake (t_d) in the firm's profits. Second, a decrease in t_d generally increases the level of monitoring γ by the board.¹⁷ Intuitively, monitoring rises because the return to monitoring is increased – since the external shareholders' net stake $(1 - t_d)\alpha_B$ also rises when t_d falls – while the cost of monitoring is unchanged.

¹⁶The second order condition could hold with equality, a knife-edge case that we rule out by assumption.

¹⁷It is possible that $\frac{d\gamma}{dt_d} > 0$ if the third derivatives $g'''(J)$, $f'''(I)$, $c'''(\gamma)$ are sufficiently large in magnitude. When f , g , and c are quadratic, $\frac{d\gamma}{dt_d}$ is unambiguously negative. Hence, barring sharp changes in the local curvature of the production functions, monitoring falls with the dividend tax rate.

Given that $\frac{d\omega}{dt_d} < 0$, it is straightforward to characterize the short-run effect of dividend taxation on firm behavior. Because the manager follows a threshold rule in ω , changes in t_d lead to both intensive and extensive margin responses. We therefore analyze the effects of a discrete dividend tax cut from $t_d = t_1$ to $t_d = t_2 < t_1$ on a firm's behavior. Let $\Delta x = x(t_2) - x(t_1)$ denote the change in a variable x caused by the tax cut, and note that $\Delta\omega = \omega(t_2) - \omega(t_1) > 0$ from (21).

Proposition 1 *A dividend tax cut ($t_2 < t_1$) has the following effects on behavior for a cash-rich firm:*

- (i) *If $\omega(t_2) \leq \bar{\omega}$: $\Delta D = 0$, $\Delta I > 0$, $\Delta J < 0$, and $\Delta I + \Delta J = 0$.*
- (ii) *If $\omega(t_1) < \bar{\omega} < \omega(t_2)$: $\Delta D > 0$, $\Delta I > 0$, $\Delta J < 0$, and $\Delta I + \Delta J < 0$.*
- (iii) *If $\bar{\omega} \leq \omega(t_1)$: $\Delta D > 0$, $\Delta I = 0$, and $\Delta J < 0$.*

Proof.

(i) When $\omega(t_2) \leq \bar{\omega}$, $D(t_2) = 0$ by Lemma 1. Since $\omega(t_2) > \omega(t_1)$, $D(t_1) = 0$ also. Therefore $\Delta D = 0$. Since $I + J + D = X$, and X is fixed, it follows that $\Delta I + \Delta J = 0$. Finally, (15) implies that $\frac{dI}{dt_d} = \frac{dI}{d\omega} \frac{d\omega}{dt_d} < 0$ when $\omega \leq \bar{\omega}$. Hence, $\Delta I > 0$ and $\Delta J = -\Delta I < 0$.

(ii) When $\omega(t_1) < \bar{\omega} < \omega(t_2)$, Lemma 1 implies $D(t_1) = 0$ while $D(t_2) > 0$. Hence $\Delta D > 0$. Since $\Delta D > 0$, $\Delta I + \Delta J = -\Delta D < 0$. By Lemma 1, $I(t_2) = I^*$ while $I(t_1)$ satisfies $(1 - t_c)\omega(t_1)f'(I(t_1)) = g'(X - I(t_1))$. Since $\omega(t_1) \cdot r < g'(X - I(t_1))$ by (13), it follows that $(1 - t_c)f'(I(t_1)) > r = (1 - t_c)f'(I^*)$, which implies $I(t_1) < I(t_2)$. Hence $\Delta I > 0$ and $\Delta J = -\Delta D - \Delta I < 0$.

(iii) When $\bar{\omega} \leq \omega(t_1)$, $I(t_1) = I(t_2) = I^*$ because $\omega(t_2) > \omega(t_1)$. Equation (14) implies that $\frac{dD}{dt_d} = \frac{dD}{d\omega} \frac{d\omega}{dt_d} < 0$ when $\omega > \bar{\omega}$. Hence $t_2 < t_1 \Rightarrow \Delta D > 0$. Finally, $\Delta J = -\Delta D < 0$. QED.

Proposition 1 shows that the dividend tax cut (weakly) increases dividend payments for all cash-rich firms because it raises the weight $\omega(t_d)$ that managers place on profits. The effect differs across three regions of ω . For managers who place a very low weight on profits ($\omega(t_2) < \bar{\omega}$), dividend payments remain undesirable after the tax cut and $\Delta D = 0$. The second region consists of firms who were non-payers prior to the tax cut ($\omega(t_1) < \bar{\omega}$), but cross the threshold for paying when the tax rate is lowered to t_2 . These firms initiate dividend payments after the tax cut. The third region consists of firms who had ω high enough that they were already paying dividends prior to the tax cut. The tax cut leads these firms to place greater

weight on net-of-tax profits relative to the pet project, and therefore leads to increases in the level of dividends. Note that these changes in dividend payout policies occur in period 0 itself. This is consistent with the evidence that many firms announced dividend increases in the weeks after the 2003 tax reform was enacted.

Now consider the effect of the dividend tax cut on investment behavior. The tax cut increases the net-of-tax return to the profit-generating project while leaving the return to pet investment unaffected. As a result, the manager substitutes from investing in perks to the profit-generating project, and I (weakly) increases while J falls. In the first region, where $\omega(t_2) < \bar{\omega}$, the manager shifts toward I from J but total investment ($I + J$) is unchanged. In the second region, where the firm initiates a dividend payment, investment in I rises to the shareholders' optimum I^* , while investment in J is reduced to finance the dividend payment and the increase in I . In this region, total investment falls when the tax rate is cut. Finally, when $\omega > \bar{\omega}(t_1)$, the manager maintains I at I^* and reduces investment in J to increase the dividend payment.

An interesting implication of these results is that a dividend tax cut weakly *lowers* total investment $I + J$ for cash-rich firms with an agency problem. Total investment, $I + J$, is the measure that is typically observed empirically since it is difficult to distinguish the components of investment in existing datasets. This prediction contrasts with the old view model, where a tax cut raises investment and with the new view model, where a tax cut has no effect on investment. Intuitively, a tax cut reduces the incentive for cash-rich firms to (inefficiently) over-invest in the pet project. It is important to note that the same result does not apply to cash-constrained firms in the agency model: A tax cut raises equity issues and productive (as well as unproductive) investment by such firms. Hence, a dividend tax cut leads to an (efficiency increasing) reallocation of capital and investment across firms, but its effect on aggregate investment is ambiguous. This result is potentially consistent with the large empirical literature on investment and the user cost of capital, which has failed to identify a robust relationship between tax rates and aggregate investment (see e.g., Chirinko 1993, Desai and Goolsbee 2004).

Next, we examine how the effect of the tax cut on dividend payments varies across firms with different ownership structures. We again distinguish between extensive and intensive margin responses.

Proposition 2 *Heterogeneity of Dividend Response to Tax Cut ($t_2 < t_1$) by Ownership Structure:*

(i) *Extensive Margin: Likelihood of Initiation.* If $\omega(t_1) < \bar{\omega}$, initiation likelihood increases with α and α_B :

- If $\Delta D > 0$ for α then $\Delta D > 0$ for $\alpha' > \alpha$
- If $\Delta D > 0$ for α_B then $\Delta D > 0$ for $\alpha'_B > \alpha_B$

(ii) *Extensive Margin: Size of Initiation.* If $\omega(t_1) < \bar{\omega} < \omega(t_2)$: $\frac{\partial \Delta D}{\partial \alpha} > 0$, $\frac{\partial \Delta D}{\partial \alpha_B} > 0$.

(iii) *Intensive Margin.* If $\bar{\omega} \leq \omega(t_1)$ and g and c are quadratic: $\frac{\partial \Delta D}{\partial \alpha} > 0$, $\frac{\partial \Delta D}{\partial \alpha_B} > 0$.

Proof.

(i) The result follows directly from the effect of α and α_B on ω . Observe that

$$\frac{\partial \omega}{\partial \alpha} = (1 - t_d)(1 + \gamma) + \alpha(1 - t_d) \frac{\partial \gamma}{\partial \alpha} = \frac{(1 + \gamma)(1 - t_d)c'' + P'_d \alpha_B \alpha (1 - t_d)^3}{c'' - P''_d \cdot \alpha_B (1 - t_d) \cdot [\alpha(1 - t_d)]^2} > 0.$$

using the second-order condition for γ in (18). Similarly,

$$\frac{\partial \omega}{\partial \alpha_B} = \alpha(1 - t_d) \frac{\partial \gamma}{\partial \alpha_B} = \frac{\alpha^2(1 - t_d)^3 P'_d(\omega)}{c'' - P''_d \cdot \alpha_B (1 - t_d) \cdot [\alpha(1 - t_d)]^2} > 0$$

Note that $\Delta D > 0$ at a given $\alpha \Rightarrow D(\omega(t_2, \alpha)) > 0$. Since $\frac{\partial \omega}{\partial \alpha} > 0$, we know that $\omega(t_2, \alpha') > \omega(t_2, \alpha)$. From (14), we have $\frac{\partial D}{\partial \omega} > 0$, which in turn implies $D(\omega(t_2, \alpha')) > D(\omega(t_2, \alpha)) > 0 \Rightarrow \Delta D > 0$ for α' . Exploiting the result that $\frac{\partial \omega}{\partial \alpha_B} > 0$ yields the analogous result for α_B .

(ii) When $\omega(t_1) < \bar{\omega} < \omega(t_2)$, $D(t_1) = 0$ and hence $\Delta D = D(t_2)$. It follows that $\frac{\partial \Delta D}{\partial x} = \frac{\partial D(t_2)}{\partial x} = \frac{\partial D}{\partial \omega} \frac{\partial \omega}{\partial x}$ for $x \in \{\alpha, \alpha_B\}$. We know that $\frac{\partial D}{\partial \omega} > 0$ from (14). Since $\frac{\partial \omega}{\partial \alpha} > 0$ and $\frac{\partial \omega}{\partial \alpha_B} > 0$ from (i), it follows that $\frac{\partial \Delta D}{\partial \alpha} > 0$ and $\frac{\partial \Delta D}{\partial \alpha_B} > 0$, which proves the claim.

(iii) When $\bar{\omega} < \omega(t_1)$, the dividend level is positive both at the initial and new tax rate and hence there is an intensive-margin response. Using equation (21), we have

$$\frac{dD}{dt_d} = \frac{dD}{d\omega} \cdot \frac{d\omega}{dt_d} = \frac{r}{g''(J(\omega))} \cdot \frac{2\alpha_B \alpha^2 (1 - t_d)^2 P'_d(\omega) + \alpha(1 + \gamma)c''}{c'' - P''_d \cdot \alpha_B (1 - t_d) \cdot [\alpha(1 - t_d)]^2} \quad (22)$$

When $\bar{\omega} < \omega$, $P_d(\omega) = D(\omega) + \frac{(1-t_c)f(I^*)+X-D}{1+r}$. Since $g''(J(\omega))$ is constant when g is quadratic, we have $D'(\omega) = -r/g''$ constant and hence $D''(\omega) = 0$ and hence $P'_d(\omega)$ is also constant and $P''_d(\omega) = 0$. Equation (22) therefore simplifies to

$$\frac{dD}{dt_d} = \frac{r}{g''} \left(\alpha(1 + \gamma) - 2\alpha_B \alpha^2 (1 - t_d)^2 \frac{r}{g'' \cdot c''} \right)$$

Recognizing that $c'' > 0$ and $g'' < 0$ are constant, we have:

$$\Delta D = \frac{r}{g''} \left\{ \alpha \int_{t_1}^{t_2} (1 + \gamma(t_d)) dt_d + \frac{2}{3} \alpha_B \alpha^2 [(1 - t_2)^3 - (1 - t_1)^3] \frac{r}{g'' \cdot c''} \right\}.$$

Because $t_1 > t_2$, the first term inside the curly brackets is negative. The second term inside the curly brackets is also negative because $t_1 > t_2$ and $g'' < 0$. Because, the multiplicative factor r/g'' outside the curly brackets is negative, we have: $\frac{\partial \Delta D}{\partial \alpha} > 0$ and $\frac{\partial \Delta D}{\partial \alpha_B} > 0$. QED.

Figure 2a plots D against α in two dividend tax regimes, with $t_1 = 35\%$ and $t_2 = 20\%$ and the corporate tax $t_c = 0$. The figure illustrates the three results in Proposition 2. First, among the set of firms who were non-payers prior to the tax cut, those with large executive shareholding (high α) are more likely to initiate dividend payments after the tax cut. This is because managers with higher α are closer to the threshold ($\bar{\omega}$) of paying dividends to begin with, and are therefore more likely to cross that threshold. Second, conditional on initiating, firms with higher α initiate larger dividends. Because $D(t_2)$, the optimal dividend conditional on paying, is rising in α , the size of the dividend increase, $\Delta D = D(t_2)$, is larger for firms with higher values of α in this region. Third, among the firms who were already paying dividends prior to the tax cut, the intensive-margin increase in the level of dividends is generally larger for firms with higher α .¹⁸ Intuitively, the manager's incentives are more sensitive to the tax rate when he owns a larger fraction of the firm. These three results apply analogously to the board's shareholding (α_B), as shown in Figure 2b. A change in t_d has a greater effect on γ when α_B is large, leading to a larger dividend response.

Auxiliary Predictions. The agency model predicts that firms with more assets and cash holdings (higher X) are more likely to initiate dividend payments following a tax cut.¹⁹ In contrast, neoclassical models that nest the old and new views (Sinn 1991) predict that firms with higher assets will respond *less* to a tax cut. Chetty and Saez (2005) document that firms with higher assets or cash holdings were more likely to initiate dividends after the 2003 tax reform, consistent with the agency model.

The importance of the interests of “key players” (executives and large external shareholders) is underscored by Chetty and Saez's finding that firms with large *non-taxable* shareholders (such as pension funds) were much less likely to change dividend payout behavior in response to the 2003 tax reform. Although we have not allowed for heterogeneity in tax rates across

¹⁸This result holds as long as there are no sharp changes in the local curvature of the production functions. If $g'''(J)$ and $c'''(\gamma)$ are sufficiently large in magnitude, it is possible to have $\frac{\partial^2 D}{\partial t_d \partial \alpha_B} > 0$.

¹⁹Firms with higher X are closer to the threshold of paying dividends because $\bar{\omega}$ is falling in X and γ is rising in X . A tax cut is therefore more likely to make firms with higher X cross the threshold and initiate dividend payments.

shareholders in our stylized model, it is straightforward to show that the introduction of non-taxable shareholders would generate this prediction. If the board includes non-taxable large shareholders, a change in t_d has a smaller impact on the board's incentive to increase monitoring. Hence, a tax cut causes a smaller increase in γ and generates smaller ΔD .

5 Efficiency Costs of Dividend and Corporate Taxation

We divide our analysis of the efficiency costs of dividend and corporate taxes into two parts. We first build intuition using a special case where ownership structure (α and α_B) is fixed and monitoring (γ) is fixed at 0. We then relax these assumptions and characterize efficiency costs when the manager's contract is endogenously determined. The lessons obtained from the special case carry over to the general model with some qualifications.

5.1 Fixed Contracts

When γ is fixed at 0, total surplus in the economy (W) is simply the sum of the shareholders' payoff, the manager's payoff, and government revenue from the dividend and corporate taxes:

$$\begin{aligned} W &= V^M + V^S + t_d P_d(\omega) + t_c P_c(\omega) \\ &= \alpha(1 - t_d) \left(D + \frac{(1 - t_c)f(I) + X - D}{1 + r} \right) + \frac{g(J)}{1 + r} \\ &\quad + (1 - \alpha)(1 - t_d)P_d(\omega) + t_d P_d(\omega) + t_c P_c(\omega) \end{aligned}$$

where $P_d = D + \frac{(1-t_c)f(I)+X-D}{1+r}$ and $P_c = f(I)/(1+r)$ denote the dividend and corporate tax bases as above.

Recognizing that D and I are chosen by the manager to maximize his own surplus, we exploit envelope conditions and obtain the following expressions for the marginal excess burden of raising the two tax rates:

$$\frac{dW}{dt_d} = t_c \frac{dP_c}{dt_d} + t_d \frac{dP_d}{dt_d} + (1 - t_d)(1 - \alpha) \frac{dP_d}{dt_d} \quad (23)$$

$$\frac{dW}{dt_c} = t_c \frac{dP_c}{dt_c} + t_d \left(\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c} \right) + (1 - t_d)(1 - \alpha) \left(\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c} \right) \quad (24)$$

where $\frac{\partial P_d}{\partial t_c} = -P_c$ denotes the mechanical effect of increasing t_c on the firm's payout as above. The first two terms in each of these formulas correspond exactly to those in the equations for deadweight loss in the neoclassical model in (4) and (5). These terms reflect the traditional

Harberger-type distortions created by taxes because the firm under-invests relative to the social optimum. Although these terms are identical to those in the neoclassical models, the elasticities themselves may differ: even cash-rich firms have $\frac{dP_d}{dt_d} < 0$, in contrast with the new view model.

The third term in the two formulas arises from the agency problem ($\alpha < 1$). This term reflects the externality that the manager imposes on other shareholders by under-providing dividends and investing in the pet project. An increase in tax rates exacerbates this pre-existing distortion. Note that unlike the Harberger terms, which are second-order (proportional to t_d and t_c), the agency term is first-order. This first-order term disappears if t_d is set at t_d^* such that $\alpha(1 - t_d^*) = 1$, as $t_d^* + (1 - t_d^*)(1 - \alpha) = 0$. The dividend subsidy $t_d^* < 0$ exactly corrects the externality due to the misalignment between managers' and shareholders' objectives. Absent revenue requirements, setting $t_d = t_d^* < 0$ and $t_c = 0$ maximizes social welfare. Rather than taxing dividends, it would be desirable to implement a Pigouvian dividend subsidy to correct the externality that arises from the misalignment between managers' and shareholders' objectives. In contrast, there is no such rationale for subsidizing corporate profits in the agency model.

As in the neoclassical model, it is helpful to distinguish firms that pay dividends in period 0 from those that do not to gain more insight into the excess burden formulas. First consider "very high cash" firms that have X large enough so that $\omega > \bar{\omega}$. By Lemma 1, these firms pay dividends in period 0, set $I = I^*(t_c)$, and set J such that $\alpha(1 - t_d)r = g'(J)$. For such firms, profitable investment is unaffected by the dividend tax ($\frac{\partial I}{\partial t_d} = 0$), implying $\frac{dP_d}{dt_d} = \frac{r}{1+r} \frac{dD}{dt_d}$. Conversely, the corporate tax does not affect pet project investment ($\frac{\partial J}{\partial t_c} = 0$) because t_c does not affect the tradeoff between D and J . Because the manager sets I to maximize P_d , the only effect of a change in the corporate tax on total dividend payouts is the mechanical effect:

$$\frac{dP_d}{dt_c} = -P_c + \left\{ \frac{(1 - t_c)f'(I)}{1 + r} \frac{\partial I}{\partial t_c} + \frac{r}{1 + r} \frac{dD}{dt_c} \right\} = -P_c - \frac{r}{1 + r} \frac{\partial J}{\partial t_c} = -P_c = \frac{\partial P_d}{\partial t_c}.$$

Combining these results, we obtain the following expressions for marginal excess burden for dividend-paying (very high cash) firms:

$$\frac{dW}{dt_d} = [t_d + (1 - t_d)(1 - \alpha)] \frac{dP_d}{dt_d} \quad (25)$$

$$\frac{dW}{dt_c} = t_c \frac{dP_c}{dt_c} \quad (26)$$

The dividend tax has a first-order deadweight cost whereas the corporate tax has a second-order deadweight burden that coincides with that in the neoclassical new view model. Intuitively, for

firms that have sufficient cash holdings to pay dividends, investment is set at the optimal level from the shareholders' perspective. The agency problem only distorts the tradeoff between period 0 dividends and pet project investment. Dividend taxes encourage managers to increase pet project investment, exacerbating this pre-existing agency problem. In contrast, corporate taxes do not affect the tradeoff between pet investment and period 0 dividends.

Now consider high cash firms that do not issue equity but also do not pay dividends ($\omega < \bar{\omega}$). Such firms set I such that $(1 - t_c)(1 - t_d)\alpha f'(I) = g'(X - I)$. For these firms, dividend and corporate taxes both distort the return to investment in the same way, implying $\frac{dP_c}{dt_d} = \frac{dP_c}{dt_c}$. The effects of t_d and t_c on the dividend tax base are fully determined by their effects on profits, implying $\frac{dP_d}{dt_d} = (1 - t_c)\frac{dP_c}{dt_d}$ and $\frac{dP_d}{dt_c} = -P_c + (1 - t_c)\frac{dP_c}{dt_c}$. Combining these results, we obtain

$$\frac{dW}{dt_d} = \frac{dW}{dt_c} = (t_c + t_d - t_c t_d)\frac{dP_c}{dt_c} + (1 - t_d)(1 - t_c)(1 - \alpha)\frac{dP_c}{dt_c} \quad (27)$$

The first term in this formula coincides with that in equation (8) for excess burden for firms that do not pay dividends in the neoclassical (old view) model. The second term is due to the agency problem, which increases the excess burden of *both* the corporate and dividend tax for firms with $X < \bar{X}$. For managers choosing between untaxed pet project investment and taxed profitable investment at the margin, both the dividend and corporate taxes distort investment behavior. Because these managers are already under-investing in I from the shareholders' perspective, both taxes exacerbate this pre-existing distortion to the same degree.

How are the two lessons about efficiency costs obtained from the neoclassical analysis in section 2 affected by agency problems? First, dividend taxation always generates deadweight loss, even for cash-rich firms. Second, the dividend tax creates first-order deadweight costs by distorting dividend payout decisions, whereas the corporate tax generates second-order efficiency costs for firms that pay dividends. To see the importance of the distinction between the first-order and second-order terms, consider the marginal excess burden of raising the dividend tax from the current rate of $t_d = 15\%$. In the Execucomp data used in Chetty and Saez (2005), total executive share ownership averages less than $\alpha = 0.03$ in all years.²⁰ In equation (25) for dividend-paying firms, the first-order agency term $(1 - t_d)(1 - \alpha)$ therefore accounts for $\frac{(1 - t_d)(1 - \alpha)}{t_d + (1 - t_d)(1 - \alpha)} = 84\%$ of the marginal excess burden of a dividend tax increase. Hence, agency

²⁰Although this calculation focuses solely on stock ownership, accounting for other forms of incentive-based pay is unlikely to raise α significantly. Existing studies have measured α more broadly by computing the change in the wealth of a CEO when his firm's value increases by \$1. These studies estimate that α is less than 1% on average for CEOs of publicly traded corporations in the U.S. (see Murphy 1999 for a survey).

effects are likely to be the primary driver of any efficiency costs of dividend taxes.

A useful feature of the formulas for excess burden in (25) and (26) is that they are functions of a small set of parameters that can in principle be estimated empirically, such as the elasticities of dividend payments and corporate profits with respect to tax rates. The primitives of the model, such as the pet project payoff $g(J)$, affect efficiency costs only through the high-level elasticities that enter the formula. Estimating these structural parameters would be difficult as they represent reduced forms of complex contracts and payoffs for shareholders and management.

The formulas for excess burden we have derived above ignore the possibility that the firm may return profits to shareholders through share repurchases instead of dividends. In the appendix, we extend the model to allow for costly share repurchases, as in Poterba and Summers (1985). We obtain the same excess burden formulas as those above, except that the first-order agency term depends upon the effect of tax changes on *total* payout. Intuitively, the cash left over for pet project investment is determined by total payout, and not just dividends. Therefore, an increase in t_d has first-order deadweight costs if it reduces total payout and does not simply induce substitution between share repurchases and dividends.²¹ An increase in t_c continues to have second-order deadweight costs for very high cash firms. The main limitation of this approach to incorporating share repurchases is that it relies on an ad-hoc cost to explain why firms pay dividends despite the tax advantage of repurchases. Micro-founded models of share repurchases may have different welfare implications.

5.2 Endogenous Contracts

We now show that the formulas derived above generalize to a model with endogenous contracts and monitoring. We begin by modeling how the manager's contract (α) is determined and then turn to the efficiency analysis, which takes into account the impact of taxes on this contract.

Determination of Manager's Contract. We model the determination of the manager's contract using the standard principal-agent framework in the corporate finance literature with a risk-neutral principal and risk-averse manager. The critical assumption we make is that this contract is chosen by the board of directors, who initially own a fraction α_B of the firm's shares and whose objective is to maximize their own profits net of monitoring costs. The remaining shares $1 - \alpha_B$ are owned by small shareholders whose interests are not directly represented on

²¹In Chetty and Saez (2006), we present suggestive evidence that companies did not substitute dividends for repurchases. However, further empirical work is needed to estimate this substitution elasticity precisely.

the board of directors. This captures the fundamental conflict between ownership and management: small minority shareholders are passive investors who do not participate in management decisions.

For reasons we describe below, it is important to ensure that the board has a set of tools that spans the set of tax instruments available to the government. We therefore expand the manager's compensation contract to include three components. First, the board can compensate the manager by giving him a fraction α of the company's shares. Second, the manager receives a fixed salary S independent of profits and dividends. The salary S is paid in period 2 before the firm is liquidated. Third, the manager receives a bonus equal to a share b of after-tax corporate profits $(1 - t_c)f(I)$ generated in period 2. In addition to these three choice variables, the board continues to choose the level of monitoring γ as above.

The board faces a tradeoff in setting the manager's contract because he is averse to risk; if the manager were risk neutral, he would buy the entire firm to resolve the agency problem and maximize total surplus. For tractability, we use a standard CARA-Normal framework to model the risk the manager faces. In particular, assume that the firm's profits are given by $f(I) + \varepsilon$ where $\varepsilon \sim N(0, \sigma^2)$. In this generalized model, the manager's total consumption is

$$V^M = (1 - t_d)\alpha \left[X - I - J + \frac{(1 - t_c)(1 - b)(f(I) + \varepsilon) + I + J - S}{1 + r} \right] + \frac{S}{1 + r} + \frac{(1 - t_c)b(f(I) + \varepsilon)}{1 + r} + \frac{g(J)}{(1 + r)(1 + \gamma)}.$$

It is convenient to rewrite this expression as

$$V^M = \tilde{\alpha} \left[X - I - J + \frac{\tilde{\delta}(f(I) + \varepsilon) + I + J - S}{1 + r} \right] + \frac{S}{1 + r} + \frac{g(J)}{(1 + r)(1 + \gamma)},$$

where $\tilde{\alpha} = (1 - t_d)\alpha$ and $\tilde{\delta} = (1 - t_c)[(1 - b) + b/\tilde{\alpha}]$. Introducing $\tilde{\alpha}$ and $\tilde{\delta}$ allows us to eliminate t_c and t_d from the manager's objective, which is another way to see that the government and private sector have equivalent tools. Any change managerial incentives caused by changes in government policies can in principle be fully undone by changes in the manager's contract.

The manager's utility function is $u(V^M) = -\frac{1}{\mu}e^{-\mu V^M}$, where μ denotes the level of absolute risk aversion. Exploiting the CARA-Normal properties, the expected value received by the manager can be written as

$$EV^M = \tilde{\alpha} \left[X - I - J + \frac{\tilde{\delta}f(I) + I + J - S}{1 + r} \right] + \frac{S}{1 + r} + \frac{g(J)}{(1 + r)(1 + \gamma)} - \mu \frac{\sigma^2}{2} \frac{\tilde{\alpha}^2 \cdot \tilde{\delta}^2}{(1 + r)^2}.$$

Note that the maximization problem of the manager who chooses I and J to maximize EV^M is identical to the problem in the deterministic model solved in Lemma 1. Hence I and J depend upon $\tilde{\alpha}$, $\tilde{\delta}$, and γ .

The board chooses S , b , α , and γ to maximize the board's share value, taking into account the manager's incentive constraints and participation constraint $EV^M \geq 0$. As above, denote by $P_c = f(I)/(1+r)$ the corporate tax base and $P_d = X - I - J + [(1-t_c)(1-b)f(I) + I + J - S]/(1+r)$ the dividend tax base. Note that we can rewrite P_d as

$$P_d(\tilde{\alpha}, \tilde{\delta}, \gamma, t_c) = X - I - J + \frac{\frac{1-t_c-\tilde{\alpha}\tilde{\delta}}{1-\tilde{\alpha}}f(I) + I + J - S}{1+r},$$

where the dependence on t_c captures the mechanical change in P_d holding fixed the manager's contract $(\tilde{\alpha}, \tilde{\delta}, \gamma)$. Note that $\frac{\partial P_d}{\partial t_c} = -\frac{P_c}{1-\tilde{\alpha}}$. With this notation, the board chooses $(\tilde{\alpha}, \tilde{\delta}, \gamma)$ to maximize:

$$W_S = (\alpha_B(1-t_d) - \tilde{\alpha})P_d(\tilde{\alpha}, \tilde{\delta}, \gamma, t_c) - c(\gamma). \quad (28)$$

The minority shareholders surplus is:

$$W_M = (1 - \alpha_B)(1 - t_d)P_d(\tilde{\alpha}, \tilde{\delta}, \gamma, t_c). \quad (29)$$

Since the manager's surplus is pinned at zero by his participation constraint, total surplus in the economy (W) is the sum of the shareholders' welfare and government revenue:

$$W = t_d P_d + t_c P_c + W_S + W_M.$$

Efficiency Cost. Using the envelope theorem, we have $dW_S/dt_d = -\alpha_B \cdot P_d$ and $dW_S/dt_c = (\alpha_B(1-t_d) - \tilde{\alpha})\frac{\partial P_d}{\partial t_c} = -P_c\frac{\alpha_B(1-t_d)-\tilde{\alpha}}{1-\tilde{\alpha}}$. We therefore have $dW_M/dt_d = -(1 - \alpha_B) \cdot P_d + (1 - \alpha_B)(1 - t_d)dP_d/dt_d$ and $dW_M/dt_c = (1 - \alpha_B)(1 - t_d)dP_d/dt_c$. Combining these results yields:

$$\frac{dW}{dt_d} = t_c \frac{dP_c}{dt_d} + [t_d + (1-t_d)(1-\alpha_B)] \frac{dP_d}{dt_d} \quad (30)$$

$$\frac{dW}{dt_c} = t_c \frac{dP_c}{dt_c} + [t_d + (1-t_d)(1-\alpha_B)] \cdot \left(\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c} \right) \quad (31)$$

Equations (30) and (31) coincide with those in the special case above, replacing α with α_B . To understand these equations, it is useful to distinguish between two cases: $\alpha_B = 1$ and $\alpha_B < 1$.

Case 1: $\alpha_B = 1$. When there are no minority shareholders, the first-order terms in (30) and (31) disappear and deadweight burden becomes a second-order function of the tax rate

as in the neoclassical old view model. The marginal deadweight cost of taxation is small at low tax rates even though the contract between the manager and board has $\alpha < 1$, leading to inefficient pet project investment and under-provision of dividends by the manager. This result contrasts with the intuition developed in the previous section that taxing a market with a pre-existing distortion leads to a first-order efficiency cost, which is a classic result in public finance (Auerbach 1985, Hines 1999, Auerbach and Hines 2003, Goulder and Williams 2003, Kaplow 2008).

There are two reasons that our result differs from that of other studies in the tax literature. First, we have designed the model so that the government does not have an intrinsic technological advantage in fixing the agency problem relative to the private sector. Any change in incentives for the manager that can be achieved by changing the tax system (t_d, t_c) can be achieved by changing the private contract (α, b, S) . Second, the contract between the manager and the shareholders is constrained efficient when $\alpha_B = 1$: absent taxes, the compensation of the manager is designed to maximize surplus subject to the technological constraint that only managers can make the investment and payout decisions for the firm. Hence, the size of the pre-existing distortion due to agency problems is endogenously minimized by the private sector when $\alpha_B = 1$ in this model. In contrast, the pre-existing distortions analyzed in the previous section and in the studies cited above are exogenously fixed. The government has a technological advantage in fixing these distortions – it can use a dividend subsidy whereas the private sector cannot – and thus dividend taxes have first-order costs.

The general lesson, which is of relevance beyond dividend taxation, is that identifying a pre-existing distortion is *not* sufficient to infer that government taxes or subsidies will have first-order effects on welfare. It is critical to understand the private sector’s ability to alter the size of the distortion, in particular whether the private sector has the same tools as the government and whether the private sector reaches the second-best efficient outcome. In the context of dividend taxation, there is no obvious reason that government intervention is a superior method of resolving agency problems than the tools available to shareholders.²²

Case 2: $\alpha_B < 1$. When $\alpha_B < 1$, the interests of diffuse shareholders are ignored by the

²²Governments may be able to affect the contracting technology in a way that the private sector itself cannot achieve through regulation (Shleifer and Vishny 1997). For example, if shareholders rights are protected in courts, shareholders may have more control over managers, reducing $c(\gamma)$ and leading to a first-order efficiency gain. The key point is that dividend taxes do not affect contracting technology directly holding fixed the regulatory structure embodied by in the function $c(\gamma)$.

board and the private contract no longer maximizes total private surplus. As a result, taxes have first-order effects, as shown by the $(1 - t_d)(1 - \alpha_B)$ terms in (30) and (31).²³ As in the fixed contract case, setting $t_d = t_d^*$ (where $\alpha_B(1 - t_d^*) = 1$) corrects the externality and eliminates these first-order terms. Setting $t_d = t_d^*$ and $t_c = 0$ thus maximizes social welfare absent revenue constraints. With endogenous contracts, the size of the first-order term in the excess burden formulas is determined by α_B instead of α . This is because the ultimate source of the externality is that the large shareholders under-provide monitoring and pay-for-performance incentives to the manager when $\alpha_B < 1$.

The model can be further generalized to permit endogenous determination of the fraction of large shareholders α_B , as shown in Chetty and Saez (2007). Large shareholders often buy a large block of shares through tender offers (Shleifer and Vishny 1986). Such tender offers are made in the self-interest of the acquirer and do not take into account the interests of the remaining diffuse shareholders. This creates an agency problem because α_B is determined in a way that does not maximize total private surplus. As a result, dividend taxation continues to generate first order efficiency costs and a dividend subsidy can be used to correct the externality.

The results with endogenous contracts explain why our formulas for excess burden differ from that obtained in Gordon and Dietz's (2006) agency model. Gordon and Dietz assume that the board of directors set the level of dividends on behalf of *all* shareholders, which is analogous to assuming $\alpha_B = 1$ in our model. This is the reason that the efficiency cost of dividend taxation takes the standard second-order Harberger form in their model.

6 Conclusion

The public finance literature on corporate taxation has focused primarily on models of profit-maximizing firms. In contrast, since Jensen and Meckling (1976), the corporate finance literature has emphasized deviations from profit maximization by managers as a central determinant of firm behavior. This paper has taken a step toward bridging this gap. We analyzed the effects of dividend taxation in an agency model, and showed that it can explain many aspects of the empirical evidence on firms' responses to taxation that pose problems for existing neoclassical models.

²³In the case with endogenous contracts, the corporate tax can have first-order effects even for very high cash firms because it distorts the manager's contract, which in turn affects payout decisions.

We used this model to characterize the efficiency cost of dividend taxation. Dividend taxation has first-order efficiency costs when managers' interests differ from shareholders and companies are owned by diffuse shareholders – which is perhaps the most plausible description of modern corporations (Shleifer and Vishny 1997). Our analysis suggests that the main source of inefficiency from increasing the dividend tax rate is the misallocation of capital by managers because of reduced monitoring, and not the distortion to the overall level of investment emphasized in the “old view” model. From a policy perspective, if agency problems are prevalent, dividend taxation should be used relatively little if the government has other tools – e.g., progressive income taxation integrated with corporate taxation – that have similar distributional effects but do not create first-order distortions.

We see two important directions for future research. First, while our model explains evidence on the effects of dividend taxation, it does not directly explain other stylized facts about dividends such as the smoothness of dividends, payment of dividends while issuing equity, and the use of dividends despite the tax advantage of share repurchases. It is critical to build a micro-founded model that explains this evidence without appealing to ad hoc costs to fully understand the effects of dividend and corporate taxation. Second, our analysis calls for further empirical work related to agency issues in corporate taxation. In our model, a dividend tax cut raises efficiency by improving the allocation of capital: firms with excess cash holdings invest *less* following a tax cut, while cash-constrained firms invest more. Testing whether tax reforms generate such heterogeneous investment responses across firms would shed light on the empirical importance of this allocation efficiency mechanism.

Appendix: Incorporating Share Repurchases

Suppose the firm can return money to shareholders through untaxed share repurchases in period 0, which we denote by R . Returning R to shareholders through repurchases has a cost $c(R)$ that is distributed across all shareholders and is increasing and convex.²⁴

The neoclassical model in Section 2 can be extended to allow such share repurchases by replacing equation (1) with

$$\max_{D,E} V = R - c(R) + (1 - t_d)D - E + \frac{(1 - t_d)[(1 - t_c)f(X + E - D - R) + X - D - R] + E}{1 + r}. \quad (32)$$

Cash-rich firms paying dividends $D > 0$ set R such that $c'(R) = t_d$ so that an increase in t_d increases R , creating partial substitution between dividends and share repurchases. Cash-constrained firms that raise equity $E > 0$ do not repurchase shares. Intermediate firms may repurchase shares. The efficiency formulas (4) and (5) are unchanged.

In the agency model of Section 5.1 with exogenous α and no monitoring, let us focus on very high cash firms that pay dividends $D > 0$ for simplicity. For such firms, the resource constraint is $I + J = X - D - R$, and we can write the manager's value as

$$V^M = \alpha[R - c(R)] + \alpha(1 - t_d) \left(D + \frac{(1 - t_c)f(I) + I + J}{1 + r} \right) + \frac{g(J)}{1 + r}.$$

Denoting by $D' = D + R$ the total period 0 payout, we have $I + J = X - D'$ and

$$V^M = \alpha[t_d R - c(R)] + \alpha(1 - t_d) \left(D' + \frac{(1 - t_c)f(I) + I + J}{1 + r} \right) + \frac{g(J)}{1 + r}.$$

This is the sum of the problem in the baseline agency model with D' replacing D plus a separable repurchase problem involving R that is equivalent to the repurchase problem in the neoclassical model. The first order condition for R is therefore $c'(R) = t_d$ as in the neoclassical model above. The first order conditions for the other variables are identical to those in baseline model without repurchases: $(1 - t_c)f'(I) = r$ and $\frac{\alpha(1 - t_d)r}{1 + r} = \frac{g'(J)}{1 + r}$. Hence, the key comparative static results for the agency model in Section 3 and 4 hold with repurchases.

²⁴In practice, share repurchases are taxed at a lower rate than dividends. It is straightforward to introduce a tax rate t_s on share repurchases without changing the analysis as the results do not depend upon the specification of $c(R)$.

Now consider the efficiency analysis. Social welfare is

$$\begin{aligned}
W &= V^M + V^S + t_d P_d + t_c P_c \\
&= \alpha[R - c(R)] + \alpha(1 - t_d) \left(D + \frac{(1 - t_c)f(I) + I + J}{1 + r} \right) + \frac{g(J)}{1 + r} \\
&\quad + (1 - \alpha)(R - c(R)) + (1 - \alpha)(1 - t_d)P_d + t_d P_d + t_c P_c
\end{aligned}$$

where $P_d = D + \frac{(1-t_c)f(I)+I+J}{1+r}$ and $P_c = f(I)/(1+r)$ denote the dividend and corporate tax bases as above. The marginal excess burden of raising the dividend tax is

$$\frac{dW}{dt_d} = t_c \frac{dP_c}{dt_d} + t_d \frac{dP_d}{dt_d} + (1 - t_d)(1 - \alpha) \frac{dP_d}{dt_d} + (1 - \alpha)(1 - c'(R)) \frac{dR}{dt_d} \quad (33)$$

$$= t_c \frac{dP_c}{dt_d} + t_d \frac{dP_d}{dt_d} + (1 - t_d)(1 - \alpha) \frac{d(P_d + R)}{dt_d} \quad (34)$$

where $\frac{d(P_d+R)}{dt_d}$ is the effect of the dividend tax on total payout. This formula coincides with (23) except that the first-order term has $\frac{d(P_d+R)}{dt_d}$ instead of $\frac{dP_d}{dt_d}$. Intuitively, R is chosen optimally by the manager from the shareholders' perspective, so the first-order agency related term in the excess burden formula depends only on the distortion in pet project investment. Pet project investment is determined by total payout, not just dividend payments, and thus total payout is what matters for the agency problem. In contrast, the standard Harberger terms are related to distortions in the dividend tax base itself and therefore continue to have the same form as in the model without repurchases.

Similarly, the excess burden of raising the corporate tax rate is:

$$\frac{dW}{dt_c} = t_d \left(\frac{dP_d}{dt_c} - \frac{\partial P_d}{\partial t_c} \right) + t_c \frac{dP_c}{dt_c} + (1 - t_d)(1 - \alpha) \left(\frac{d(P_d + R)}{dt_c} - \frac{\partial P_d}{\partial t_c} \right).$$

As $c'(R) = t_d$, R is unaffected by t_c and therefore $\frac{d(P_d+R)}{dt_c} = \frac{dP_d}{dt_c} = -P_c = \frac{\partial P_d}{\partial t_c}$. Hence, even with share repurchases, corporate taxes have second-order deadweight burden for very high cash firms:

$$\frac{dW}{dt_c} = t_c \frac{dP_c}{dt_c}.$$

References

- Auerbach, Alan, 1979, "Wealth maximization and the cost of capital," *Quarterly Journal of Economics*, 93, 433-46.
- Auerbach, Alan, 1985, "The Theory of Excess Burden and Optimal Taxation," in Alan Auerbach and Martin Feldstein eds. *Handbook of Public Economics*, Volume 1, 61-127, Amsterdam: North-Holland.
- Auerbach, Alan and Kevin Hassett, 2007, "The 2003 Dividend Tax Cuts and the Value of the Firm: An Event Study," in A. Auerbach, J. Hines, and J. Slemrod, eds., *Taxing Corporate Income in the 21st Century*, Cambridge: Cambridge University Press, 93-126.
- Auerbach, Alan and James Hines, 2003, "Taxation and Economic Efficiency," in Alan Auerbach and Martin Feldstein eds. *Handbook of Public Economics*, Volume 3, Amsterdam: North-Holland.
- Benartzi, Shlomo, Roni Michaely, and Richard Thaler, 1997, "Do Dividends Signal the Future or the Past?" *Journal of Finance*, 52(3), 1007-1034.
- Bernheim, B. Douglas, 1991, "Tax Policy and the Dividend Puzzle," *RAND Journal of Economics* 22, 455-76.
- Bertrand, Marianne and Sendhil Mullainathan, 2003, "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences," *Journal of Political Economy* October.
- Bradford, David F., 1981, "The incidence and allocation effects of a tax on corporate distributions," *Journal of Public Economics* 15, 1-22.
- Brown, Jeffrey, Nellie Liang, and Scott Weisbenner, 2007, "Executive Financial Incentives and Payout Policy: Firm Responses to the 2003 Dividend Tax Cut," *Journal of Finance*, 62(4), 1935-1965.
- 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 Papers and Proceedings* 96(2), 124-129.
- Chetty, Raj and Emmanuel Saez, 2007, "An Agency Theory of Dividend Taxation," NBER Working Paper No. 13538.
- Chirinko, Robert S., 1993, "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications," *Journal of Economic Literature*, 31(4), 1875-1911.
- Christie, William and Vikram Nanda, 1994, "Free Cash Flow, Shareholder Value, and the Undistributed Profits Tax of 1936 and 1937," *Journal of Finance*, 49 (5) 1727-1754.
- Coase, Ronald, 1960, "The Problem of Social Cost," *Journal of Law and Economics* 3(1), 1-44.
- Desai, Mihir A., C. Fritz Foley, and James Hines, 2007, "Dividend Policy inside the Multinational Firm," NBER Working Paper No. 8698, forthcoming in *Financial Management*.
- Desai, Mihir A. and Austan D. Goolsbee, 2004 "Investment, Overhang and Tax Policy," *Brookings Papers on Economic Activity* 2, 285-355.

- 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, forthcoming in eds. Alan Auerbach and Daniel Shaviro *Institutional Foundations of Public Finance: Economic and Legal Perspectives*, Harvard University Press, Cambridge, MA.
- 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.
- Grullon, Gustavo, Roni Michaely, Shlomo Benartzi, and Richard H. Thaler, 2005, "Dividend Changes Do Not Signal Changes in Future Profitability," *Journal of Business*, 78, 1659-1682.
- 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.
- Hines, James R., Jr., 1999, "Three Sides of Harberger Triangles," *Journal of Economic Perspectives* 13(2), 167-188.
- 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.
- Kaplow, Louis, 2008, *The Theory of Taxation and Public Economics*, Princeton University Press: New Jersey.
- King, Mervin A., 1977, *Public Policy and the Corporation*, Chapman and Hall, London.
- Korinek, Anton and Joseph Stiglitz, 2009, "Dividend Taxation and Intertemporal Tax Arbitrage," *Journal of Public Economics*, 93(1-2), 142-159.
- Lang, Larry H. P., and Robert H. Litztenberger, 1989, "Dividend Announcements: Cash Flow Signalling versus Free Cash Flow Hypothesis?" *Journal of Financial Economics* 24, 181-192.
- 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.
- Murphy, Kevin J., 1999, "Executive Compensation," in Ashenfelter, Orley and David Card eds. *Handbook of Labor Economics*, Volume 3, Amsterdam: North-Holland.
- 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.
- Poterba, James, 2004, "Corporate Payout Policy," *American Economic Review Papers and Proceedings*, 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.

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, 1986, "Large shareholders and Corporate Control," *Journal of Political Economy* 94 (3), 461-488.

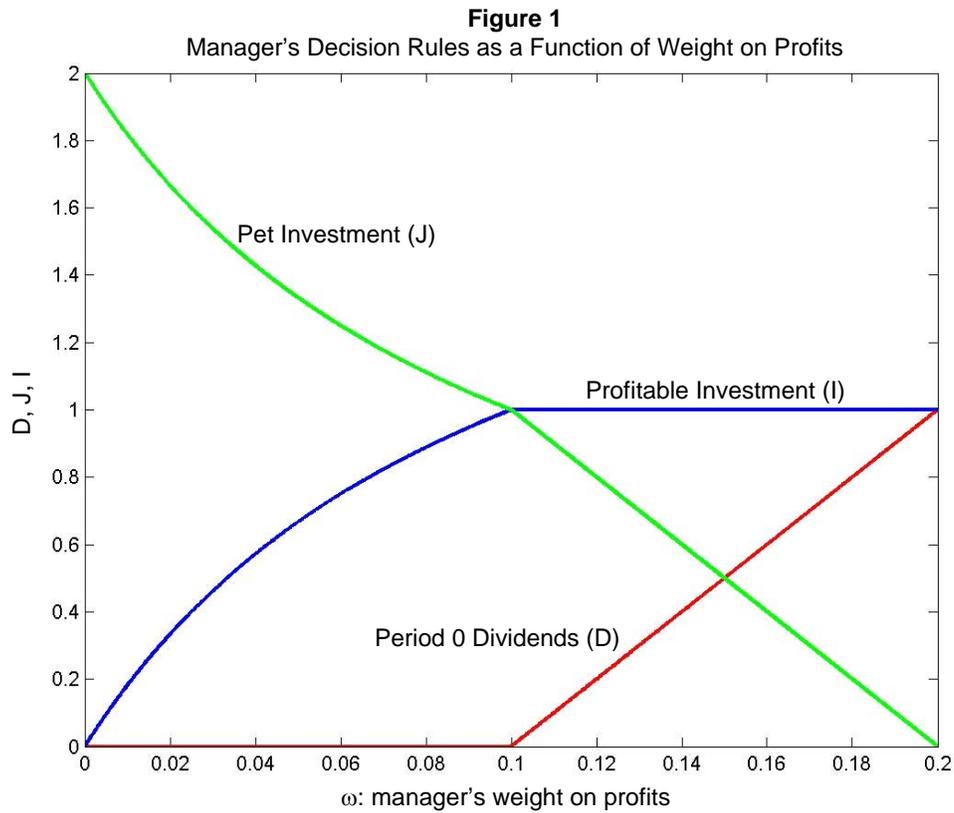
Shleifer, Andrei and Robert Vishny, 1997, "A Survey of Corporate Governance," Nobel Symposium on Law and Finance, August 1995, *Journal of Finance*.

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

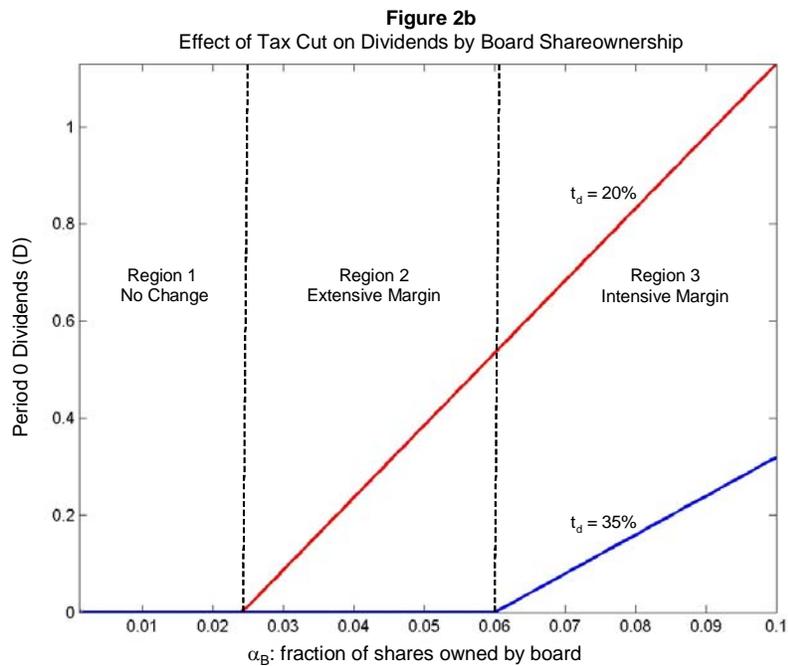
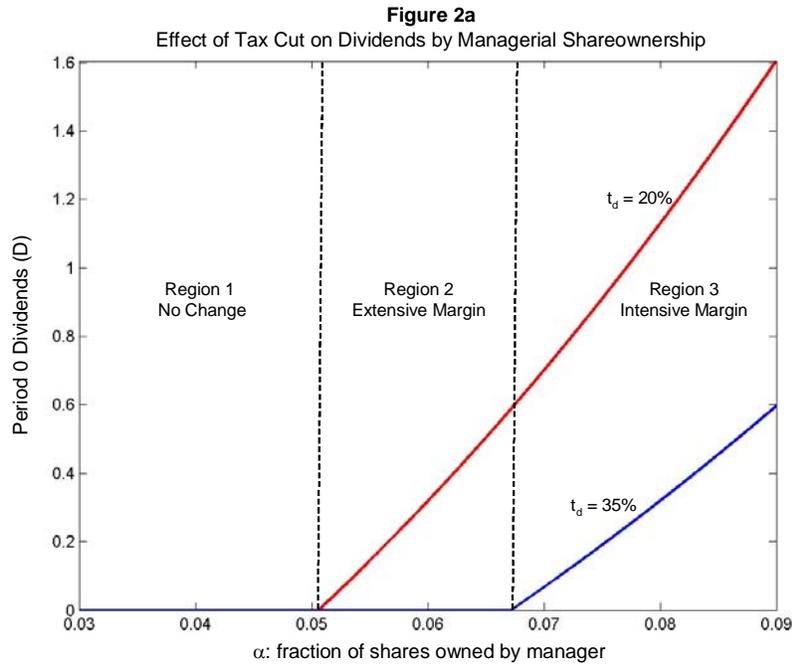
Table 1
Summary of Key Predictions: Neoclassical vs. Agency Models

	Neoclassical Model			Agency Model	
	Old View		New View		
Initial Cash X	Low: $f(X) > r / [(1-t_d)(1-t_c)]$	Medium: $r / [(1-t_d)(1-t_c)] \geq f(X) \geq r / (1-t_c)$	High: $r / (1-t_c) > f(X)$	High: $r / (1-t_c) > f(X)$ and $g'(X-I^*) \geq \omega r$	Very High: $r / (1-t_c) > f(X)$ and $g'(X-I^*) < \omega r$
Dividends D	D=0	D=0	D>0, $f(X-D) = r / (1-t_c)$	D=0	D>0, $g'(X-I^*-D) = \omega(1+r)$
Equity Issues E	E>0, $f(X+E) = r / [(1-t_d)(1-t_c)]$	E=0	E=0	E=0	E=0
Productive Investment I	I>X, $f(I) = r / [(1-t_d)(1-t_c)]$	I=X, $f(I) = f(X)$	I<X, $f(I) = r / (1-t_c)$	$g'(X-I) = (1-t_c)\omega f(I)$, $(1-t_c)f(I) > r$	$(1-t_c)f(I) = r$ (i.e., $I=I^*$), $g'(J) = \omega r$
Effects of reducing dividend tax t_d	No effect on D I increases, E increases	Intensive margin: No effect on D, E, I Extensive margin: Some firms shift to low cash regime, start issuing E and increase I	No effect on D, E, I	Intensive margin: No effect on D and E, I increases, J decreases Extensive margin: Some firms shift to very high cash regime, start paying dividends	D increases, J decreases No effect on I and E
Heterogeneity of D response to tax cut by ownership structure	none	none	none	Extensive margin: higher likelihood and larger D initiations if exec. or board share high	Larger increase in D if exec. or board share high (if third derivatives of g, c small)
Efficiency cost of t_d	second-order $dW/dt_d = (t_c + t_d - t_c t_d) dP_d/dt_c$	none $dW/dt_d = 0$	none $dW/dt_d = 0$	first-order if $\alpha < 1$ $dW/dt_d = (t_c + t_d - t_c t_d) dP_d/dt_c + (1-t_d)(1-t_c)(1-\alpha) dP_d/dt_c$	first-order if $\alpha < 1$ $dW/dt_d = [t_d + (1-t_d)(1-\alpha)] dP_d/dt_d$
Efficiency cost of t_c	second-order $dW/dt_c = (t_c + t_d - t_c t_d) dP_d/dt_c$	none $dW/dt_c = 0$	second-order $dW/dt_c = t_c dP_d/dt_c$	first-order if $\alpha < 1$ $dW/dt_c = (t_c + t_d - t_c t_d) dP_d/dt_c + (1-t_d)(1-t_c)(1-\alpha) dP_d/dt_c$	second-order $dW/dt_c = t_c dP_d/dt_c$

NOTES-- This table summarizes the firm's choice of dividends (D), equity issues (E), and investment (I) in the neoclassical and agency models. Behavior depends on the level of initial cash holding X, which varies across the columns. I^* denotes the optimal investment level from the shareholders' perspective given the corporate tax t_c , which satisfies $f(I^*) = r / (1-t_c)$. In the agency model, we only consider the case where initial cash is high enough so that the firm does not issue equity. Positive predictions reported are for the model in Sections 3 and 4 with an exogenous manager share α and endogenous monitoring γ so that $\omega = \alpha(1-t_d)(1+\gamma)$. The efficiency costs are reported for the special case in Section 5.1 with exogenous α and no monitoring. Section 5.2 shows that the formulas extend with endogenous α and monitoring by substituting α for α_B (share ownership of large shareholders). Note that the efficiency cost formulas ignore changes in the thresholds that define the low vs. high cash categories and therefore apply only to firms in the interior of these categories.



NOTE—This figure plots the manager's optimal choice of dividends, profitable investment, and pet project investment as a function of his weight on profits, ω . The simulation assumes a total cash holding of $X = 2$, profitable investment production function $f(I) = \frac{1}{10} (2I - \frac{I^2}{2})$, pet production function $g(J) = \frac{1}{100} [2J - \frac{J^2}{2}]$, and interest rate $r = 10\%$.



NOTE—These figures show how the effect of a dividend tax cut on dividends varies across firms with different ownership structures. In Figure 2a, the lower curve plots dividends versus the fraction of shares owned by the manager (α) when the tax rate is 35%. The upper curve plots the same when the tax rate is 20%. Figure 2b plots dividends versus the fraction of shares owned by the board of directors in the two tax regimes. Simulations use the same parametric assumptions as in Figure 1 along with $c(\gamma) = \frac{1}{1000} \gamma^2$.