Taxable Cash Dividends – A Useful Way of Burning Money.

by

Ken L. Bechmann og Johannes Raaballe
Taxable Cash Dividends –
A Useful Way of Burning Money

Ken L. Bechmann
Department of Finance
Copenhagen Business School

and

Johannes Raaballe
Department of Management
University of Aarhus

Current version: September 13, 2005

Key words: Dividends, Share Repurchases, Signaling, Single-Crossing Property, Money Burning
JEL Classification: G35, D82

Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark
Phone: +45 38152953 Fax: +45 38153600 e-mail: kb.fi@cbs.dk

University Park 322, DK-8000, Aarhus C, Denmark
Phone: +45 89421564 Fax: +45 86135132 e-mail: jraaballe@econ.au.dk

The authors owe thanks to Elazar Berkovitch, Bruce D. Grundy, Kristian Rydqvist, Clas Wihlborg, and Josef Zechner, as well as to the participants in the European Finance Association Conference in London, the European Financial Management Association Conference in Helsinki, the Nordic Workshop on Corporate Finance in Copenhagen, and to the participants in seminars at the Center for Financial Studies, Melbourne University, the University of Technology, Sydney, Copenhagen Business School, University of Aarhus, the Aarhus School of Business, and the Centre for Analytical Finance.
Taxable Cash Dividends –
A Useful Way of Burning Money

Key words: Dividends, Share Repurchases, Signaling, Single-Crossing Property, Money Burning

JEL Classification: G35, D82

Abstract

Firms pay out cash using both dividends and share repurchases. In many aspects these two means are similar, but one important difference is that dividends are generally taxed more heavily than share repurchases. Nevertheless firms persist in paying out large amounts in dividends. This paper provides an explanation for this dividend puzzle by developing a class of signaling models violating the “single-crossing” property in which information about the quality of the firm is asymmetric between the management and the shareholders. In these models a high-quality firm can always signal its quality by using share repurchases only. However, in certain cases share repurchases become costlier on the margin for a high-quality firm than for a low-quality imitator. In such cases, the high-quality firm signals most cost efficiently by means of a combination of share repurchases and taxable cash dividends financed by the issuance of new shares. Taxable cash dividends financed by the issuance of new shares then can be considered a positive kind of money burning whose role is to signal a firm’s high quality. The implications of the models are consistent with several important empirical facts about dividends and share repurchases. Thus, this paper’s main contribution is to examine a range of new signaling models that provides a role for taxable cash dividends and share repurchases and to derive their empirical implications.
1. Introduction

Cash dividends and share repurchases are two means by which a firm can pay out cash to shareholders and they have the same consequences in many respects. However, cash dividends are normally taxed more heavily than share repurchases. Not only is the tax rate on dividends generally higher than the tax rate on capital gains, but also, even if the two tax rates are equal, dividends are taxed more heavily than share repurchases in the sense that share repurchases offer the possibility of a tax deferral in contrast to cash dividends. Hence, dividends are generally a more costly payout source than share repurchases, but nevertheless firms persist in paying out huge amounts of cash in taxable dividends, which gives rise to the so-called “dividend puzzle” (Black, 1976).

In recent years, a large amount of literature has investigated the dividend puzzle empirically by considering the use of cash dividends relative to share repurchases, whereas surprisingly few papers have considered the dividend puzzle based on new models and/or new theoretical explanations. In particular, the number of theoretical contributions is difficult to comprehend given that until now theoretical models have not been satisfactory either because of their assumptions or their implications, leading Brealey and Myers (2005, p. 964) to conclude that the dividend puzzle is still one of the ten unsolved problems in finance.

The main contribution of this paper is that it will provide a general role for taxable cash dividends within signaling models that allows for share repurchases and dividends as signaling devices, thus suggesting a solution to the dividend puzzle. This is done by extending one of the most classical signaling models as well as by providing several other new models that build on assumptions generally accepted within the corporate finance literature. Finally, the models provide a range of testable implications which are generally either consistent with the existing literature or new implications that seem unique to these models. Therefore, the implications also provide new findings that should be tested empirically in order to test these models and to increase knowledge about pay out policies.

More precisely, this paper provides a new explanation for the dividend puzzle by developing a class of signaling models in which information about the quality of the firm is asymmetric.

---

1 Some recent empirical papers that examine the dividend puzzle are Fama and French (2001), Gruillon and Michaely (2002), DeAngelo, DeAngelo and Skinner (2004), Trojanowski (2004), and Brav et al. (2005), just to mention a few. Trojanowski (2004) is especially interesting because, in contrast to many of the abovementioned studies of the US, he does not find that companies in the UK are substituting repurchases for dividends. As discussed later in this paper, any new theoretical explanations within the last several years basically amount to those of Allen, Bernardo and Welch (2000).
between the management and the shareholders. In the literature there is an abundance of signaling models based on the “single-crossing” property (Spence, 1974 and Riley, 1979), including many of the classical signaling models discussed below. Therefore, another important contribution of our paper is to provide a signaling explanation for the dividend puzzle that is based on models where the “single-crossing” property does not hold. In these models, management can use taxable cash dividends as well as share repurchases in order to signal firm quality to uninformed shareholders and violation of the “single-crossing” property is a necessary condition for the existence of positive taxable cash dividends.

More specifically, the results show that a high-quality firm can always signal its quality using share repurchases only. However, this turns out not to be the most efficient way of signaling in cases when signaling with share repurchases becomes more costly on the margin for the high-quality firm than for a low-quality imitator. In these cases, the high-quality firm will also find it optimal to signal using taxable cash dividends financed by the issuance of new shares. Although paying out cash dividends financed by a share issuance is tantamount to burning money from the shareholder’s point of view, money burning turns out to be the cheapest way of signaling “high quality”.

Based on two specific fully analyzed models, it is argued that the implications of the model are consistent with at least four important empirical facts about dividends, which are as follows: 1) Some firms will not pay out using cash dividends; 2) Higher cash dividends imply a higher market value; 3) Cash dividends and share issuance can be observed simultaneously, and 4) Cash dividends turn out to be less volatile than earnings, i.e. there is dividend smoothing.

Several other papers explain a firm’s dividend policy using asymmetric information models and most of these papers assume asymmetric information between the management and the shareholders. Two such classical papers are Bhattacharya (1979) and Miller and Rock (1985). In Bhattacharya (1979) taxable dividends signal firm quality but the driving force is actually the asymmetric cost of obtaining liquidity. In Miller and Rock (1985) dividends are not taxed and could therefore be more precisely denoted as total payout. Both papers share the feature that if both taxable cash dividends and non-taxable share repurchases are available as a payout source, only share repurchases will be used. Hence, both papers are able to explain why cash paid out to shareholders may function as a signal of firm quality, but not why taxable dividends are used in some cases instead of share repurchases.

---

2 Allen and Michaely (2002) provide an extensive survey of the literature on payout policy. Therefore, the following focuses only on the papers most relevant for this paper.
Several other models deal with taxable dividends as a signaling device. Examples based on asymmetric information between the management and the shareholders include John and Williams (1985), Ambarish, John and Williams (1987), and Williams (1988). All these models provide explanations for the use of taxable cash dividends, but build on assumptions that lead cash dividends to be different from share repurchases, even in the absence of taxes.

Somewhat different approaches are taken by Brennan and Thakor (1990), Bernheim (1991), and Allen, Bernardo and Welch (2000). In Brennan and Thakor (1990), a shareholder can become informed by paying a fixed cost, and in the case of a share repurchase, the informed shareholders will sell shares if and only if the repurchase price is higher than the true value of the shares. Therefore, share repurchases lead to an expropriation of uninformed shareholders, whereas all shareholders will have to pay the tax on dividends. Shareholders decide the payout policy by voting. If the amounts to be paid out are small, it does not pay off for most of the shareholders to become informed. Instead, the majority of shareholders will vote for cash dividends as they find it optimal to pay dividend taxes rather than being expropriated by large shareholders. For larger amounts, share repurchases are used because enough shareholders become informed due to a higher expropriation incentive and hence, the majority will vote for share repurchases.

In Bernheim (1991), taxable cash dividends function as a signal about firm quality and can be seen simultaneously with share repurchases. However, it turns out that the financing of the total payout induces internal problems within the model. Furthermore, the model has problems in explaining, for example, dividend smoothing. Allen, Bernardo and Welch (2000) explain taxable cash dividends as a way of attracting institutional investors, who are not tax-disadvantaged, and thus attract better-informed monitoring investors.

Bernheim and Redding (2001) suggest – as do Bernheim (1991) and this paper – that cash dividends act as a money burning device because of a violation of the “single-crossing” property. However, their focus and their results are quite different from ours. First, the main contribution of Bernheim and Redding (2001) is a general theoretical analysis of how to choose between several different ways of burning money. Second, when applying theoretical results to corporate dividend policies, the focus is on the choice of taxable cash dividends compared to other money burning mechanisms and not on the choice between share repurchases and taxable cash dividends. Finally, only very little is said about the empirical predictions of the model.

While several pieces of the dividend puzzle have been put into place, a generally accepted model remains to be developed, and as mentioned, the dividend puzzle is still deemed to be one of the ten unsolved problems in finance (Brealey and Myers, 2005, p. 964). Some models are based
on unpalatable assumptions, others are in conflict with empirical facts, and still others are not rich enough with respect to predictions. A detailed review of these models can be found in Allen and Michaely (2002) and in Raaballe and Bechmann (2001); the latter focusing on the unpalatable assumptions and models in conflict with empirical facts.

In summary, the main contribution of the present paper is that it shows how dividends can be used as a positive kind of money burning in several different signaling models where the “single-crossing” property does not hold true. Bernheim and Redding (2001), for example, show that the violation of the “single-crossing” property complicates the analysis of the model considerably. Therefore, in order to keep the model and the analysis as simple as possible, most of the analysis here considers one simple two-type model and will in all cases be restricted to a separating equilibrium. This allows us to derive the main results and to explain the main intuition behind these results as simply as possible. However, some of the implications of the model, such as monotonicity, turn out to depend crucially on the two-type assumption. Unfortunately, the analysis becomes quite technical in the case of a continuum of types. Therefore, the following is limited to a non-technical description of the results and the related empirical implications in a specific continuum-of-types model. A detailed analysis of the continuum-of-types model can be found in Raaballe and Bechmann (2001).

The rest of the paper is organized as follows. In section 2 we focus upon one specific model that explains why firms want to pay taxable cash dividends even though it amounts to burning shareholders’ money. The main results in the simplest case are derived with only two different types of firms. Section 3 describes how these results extend to the case with a continuum of types and discusses the empirical predictions. Section 4 presents other models that cause dividends to become a money-burning mechanism and discusses whether cash dividends financed by a share issuance are the only attractive money-burning strategy. Conclusions are given in section 5.

2. The model

2a. Setup

Development of our basic idea is based on two extensions of the Miller and Rock (1985) model, probably the most well known dividend-signaling model. Just as in Miller and Rock (1985), the assumption is made that, measured by its level of earnings, $x$, at time 0, a firm financed by equity
alone can have different types (qualities). The earnings can be used for real investments, \( I \), or paid out to shareholders. However, as the first extension of the Miller and Rock (1985) model we specifically allow the firm to pay out using taxable cash dividends, \( D \), and share repurchases, \( B \), and thus, the cash flow equation at time 0 is given by:

\[
(1) \quad x = I + D + B \Leftrightarrow x - I = D + B.
\]

Cash dividends, \( D \), are assumed to be non-negative and taxed at the rate \( 1 - \alpha \), which means that shareholders receive \( \alpha D \) after tax. Without loss of generality, the assumption is made that share repurchases, \( B \), are not taxed. Share repurchases can be positive, corresponding to shares being repurchased or negative, corresponding to shares being issued.\(^3\)

The model assumes asymmetric information between the management of the firm and the shareholders. Management knows the level of earnings, \( x \), and decides on the distribution between investments, \( I \), cash dividends, \( D \), and share repurchases, \( B \). The shareholders observe only the cash dividends and share repurchases being paid out. Hence, \( x \) and \( I \) are both unobservable for the shareholders. Using the observed values of \( D \) and \( B \), the shareholders can only calculate \( x - I \) using equation (1). However, the shareholders can also study the accounting reports in order to obtain additional information about earnings and investments. Since accounting reports permit flexibility as to how expenses and investments are categorized, the reports will not perfectly reveal the level of earnings, \( x \), but only a number of possible values of \( x \) (an interval).\(^4\)

Illustrating this important point is the example of a firm which pays out a total of \( D + B = 150 \) and reports a level of earnings of \( x = 250 \) and a level of investments of \( I = 100 \). We note that equation (1) holds, but such an accounting report may be misleading. If the level of earnings is due to an invoice of 1000 and total costs of 750, the accounting report reveals in this case the true type of the firm. However, the manner in which total costs are accounted for in the report, make it possible to manipulate the level of earnings. For example, assume that the total costs were actually 800 and that the firm reported 50 of these costs as investments. By doing so, the firm can still report a level of earnings of 250 and a level of investments of 100 and pay out \( D + B = 150 \) even though the true type is now only \( x = 200 \) and real investments are only \( I = 50 \). Hence, the accounting report will only reveal a set of possible values of \( x \). The number of possible values of \( x \) depends on the degree of flexibility in the categorization of expenses.

\(^3\) It is assumed that investments in financial assets are observable and are therefore included in \( B \). This discussion is looked at again in section 3b.

\(^4\) A survey of such accounting manipulation is found in, for example, Healy and Wahlen (1999).
The real investments return an expected value of $F(I)$ in the subsequent period (time 1 – the final period). In the second extension, in contrast to Miller and Rock (1985), it is assumed that the production function, $F(I)$, is S-shaped, i.e. that it exhibits increasing returns to scale first and decreasing returns to scale later, as shown in Figure 1a. Traditional S-shaped production functions have been widely discussed and used in the economics literature (see Knight (1921), Cassels (1936), and Frisch (1965) for some of the classical references).

Finally, the objective function of management needs to be described and discussed. First, consider a firm of type $x$, i.e. having true earnings $x$. The shareholders in the firm are assumed to be risk neutral and to consist of two groups, e.g. the fraction $k$, which sells its shares at time 0 (short-term shareholders) and the fraction $1-k$, which does not sell its shares until time 1 (long-term shareholders). If the firm pays out $D$ in cash dividends and $B$ in share repurchases, the shareholders will receive $\alpha \cdot D + B$ after taxes at time 0. Furthermore, from equation (1) the level of real investments will be $I = x - D - B$. Hence, assuming that the interest rate is zero means that the present value of the selling price to the long-term shareholders is $F(I) = F(x - D - B)$. Similarly, if the firm paying out $D$ and $B$ is taken to be of quality $x$ by the stock market, the selling price at time 0 for the short-term shareholders will also be $F(I) = F(x - D - B)$. Thus, the total value of the firm to the present shareholders is:

$$\Pi = \alpha \cdot D + B + k \cdot F(x - D - B) + (1-k) \cdot F(x - D - B)$$

Management’s objective is to maximize the total value of the firm for the shareholders which may provide an incentive for the low-quality firms to try imitating the high-quality firms by imitating their payout i.e. to break the “tentative separating equilibrium” described above. To illustrate this incentive, consider a type $y$ firm with a quality (earnings) lower than $x$, i.e. $y < x$. This firm can attempt to imitate the high-quality firm by acting in two ways. First, the firm of type $y$ increases its $D$ and $B$ to the same levels as those chosen by type $x$, and second, type $y$ simultaneously manipulates its earnings as described above so that the accounting report appears to reveal the same level of investments and earnings as type $x$. However, as seen from equation (1), the drawback is that this imitation results in a decrease in the level of real investments $I$ for
firm \(y\). If the firm is successful in its imitation, the total value of the type \(y\) firm for shareholders will be:

\[
\Pi = \alpha \cdot D + B + k \cdot F(x - D - B) + (1 - k) \cdot F(y - D - B)
\]

As will be shown later, there are cases in which a low-quality firm has a strict incentive to imitate a high-quality firm. The reason is that the positive effects on total firm value from the increased payout and higher selling price for the short-term shareholders can more than outweigh the negative effect on the selling price for long-term shareholders.

In order to prevent such an imitation and hence, to ensure a high selling price to the shareholders selling at time 0, a high-quality firm has an incentive to signal. As discussed in the introduction, management can signal the quality to the stock market using taxable cash dividends and share repurchases. This leads to what are denoted as the two pure signaling strategies available to management. One pure strategy is to pay out cash using only share repurchases (and hence causing investment disturbances). The other pure strategy is to pay out taxable cash dividends financed entirely by the issuance of new shares or reduced share repurchases (which will only lead to the payment of taxes).

If share repurchases are chosen, the high-quality firm then incurs costs associated with a change in the level of investments and thereby in the production returned at time 1. This decreases the firm’s value for the shareholders and is thus costly to the firm. Similarly, a low-quality firm also incurs costs when the level of investments changes if it tries to imitate the high-quality firm by paying out the same amount of cash using share repurchases. The relative size of these marginal costs to the two different types of firms will depend on where the firms are situated on the production function in Figure 1a. If the second strategy with taxable cash dividends financed by issuing shares is instead chosen, shareholders in both firms incur taxes on the dividends received, hence making this strategy equally costly to the two types of firms. The following analyzes this in further detail.

2b. Analysis based on two types of firms

In this section the cases where the earnings of the company are either low, \(\bar{x}\), or high, \(x\) are analyzed. Since it turns out that the form of the optimal signaling equilibrium depends on the size of \(x\) relative to \(\bar{x}\), \(x\) is varied parametrically in the interval \([\bar{x}, \infty]\).
As a starting point for the analysis, it is assumed that independently of type, the company chooses the first best level of investment (Fisherian optimum) and of payout (zero taxable cash dividends). The implication is that the share repurchases of the high type must be $x - \bar{x}$ higher than the share repurchases of the low type. It is then shown that in this case the low type will find it profitable to imitate the first best payout of the high type, i.e. the tentative equilibrium is broken. Based on the imitation profit of the low type, the following question can then be asked: What is the cheapest way the high type can eliminate the imitation profit of the low type (and hence deter imitation)? The answer turns out to be:

1. If the high type’s earnings are not much higher than the low type’s earnings, the cheapest way to deter imitation from the low type is to increase share repurchases/under-investments.
2. If the high type’s earnings are much higher than the low type’s earnings, the cheapest way to deter imitation from the low type is by means of a combination of increased share repurchases/under-investments, as mentioned above, and money-burning in the form of a payout of taxable cash dividends financed by share issues.

a. The first best level of investments and payout

If investors were fully aware of the true earnings of the company ($\bar{x}$ or $x$), the first best choices of the level of investments, share repurchases, and taxable cash dividends would be given by:

$$D_{FB}(x) = D_{FB}(\bar{x}) = 0$$

$$F'(I_{FB}) \equiv F'(x - B_{FB}(x)) = F'(\bar{x} - B_{FB}(\bar{x})) = 1.$$  

(Insert Figure 1b here)

The explanation is as follows (see also Figure 1b). If share repurchases are increased by 1, the investors receive an additional unit in cash. However, an additional share repurchase of 1 reduces the level of investments by 1, reducing the value of the company by $F'(I)$. Optimality requires $F'(I_{FB}) = 1$ (point A in Figure 1b). Suppose the company, e.g. type $x$ for illustrative
purposes, has chosen the first best level of investments, $I_{FB}$, by means of share repurchases as well as some positive taxable cash dividends, i.e. $I_{FB} = x - B(x) - D(x)$. This is not an optimal solution since the only effect of decreasing $D(x)$ by 1 and increasing $B(x)$ by 1 is that the shareholders receive an additional after-tax liquidity of $1 - \alpha > 0$.

Hence, under symmetric information both types find it optimal to choose the Fisherian level of investments $I_{FB} = x - B_{FB}(x) = x - B_{FB}(\bar{x})$ and not to pay out any taxable cash dividends. For future reference, we note that this implies that $B_{FB}(x) - B_{FB}(\bar{x}) = x - \bar{x}$.

b. Imitation occurs if the high type holds on to the first best solution

In case the low type, $\bar{x}$, imitates by choosing the same level of first best share repurchases as the high type, $x$, the low type gets the following imitation gain:

$$\pi(x, \bar{x}) = B_{FB}(x) + kF(x - B_{FB}(x)) + (1-k)F(\bar{x} - B_{FB}(\bar{x})) - B_{FB}(\bar{x}) - F(\bar{x} - B_{FB}(\bar{x}))$$

$$= B_{FB}(x) - B_{FB}(\bar{x}) - (1-k)[F(\bar{x} - B_{FB}(\bar{x})) - F(\bar{x} - B_{FB}(x))]$$

$$= x - \bar{x} - (1-k) \int_{\bar{x} - x}^{0} F'(\bar{x} - B_{FB}(\bar{x}) + \tau) d\tau$$

$$= \int_{0}^{\bar{x} - x} d\tau - (1-k) \int_{0}^{\bar{x} - x} F'(\bar{x} - B_{FB}(\bar{x}) - \tau) d\tau,$$

where $B_{FB}(x) - B_{FB}(\bar{x}) = x - \bar{x}$ from above has been used to obtain the third expression.

The decomposition of the low type’s imitation gain is as follows: Both short-term and long-term shareholders get a higher payout. While short-term shareholders get an unchanged sale price, long-term shareholders get a lower sale price due to reduced production.

This imitation gain has been depicted for two different types, $x' < x''$, in Figure 2.

(Insert Figure 2 here)
If the low type, \( x \), imitates a type \( x' \), the imitation gain corresponds to area \( A \) in Figure 2. If the low type imitates a type \( x'' \), the imitation gain corresponds to area \( A + B - C + D \) in Figure 2.

In order to have an imitation problem for all possible values of \( x \), it is assumed in this section that

\[
\pi(x, x) > 0 \quad \text{for all } x \in [x, \infty[.
\]

Note that this assumption will always hold if, for example, the fraction of short-term shareholders, \( k \), is sufficiently large. Later, it is observed that such an assumption is not needed in the model based on a continuum of types.

c1. How the high type deters imitation – the case where \( x \) is only slightly higher than \( \bar{x} \)

In Figure 3a, point A corresponds to the high type when he holds on to the first best solution. Point B corresponds to the low type when he imitates the first best payout of the high type. The low type earns an imitation profit \( \pi(x, x) \), e.g. 5, when he imitates the high type. In order to deter imitation, the high type has to inflict costs of 5 on the low type if the low type imitates. This can be achieved by increased share repurchases or by money burning.

(Insert Figure 3a here)

*Increased share repurchases/under-investments.* If the high type increases share repurchases by one, it follows from equation (2) that his short-term shareholders incur a total loss of \( k(F'(x - B^{FB}(x)) - 1) \) and his long-term shareholders incur a total loss of \( (1-k)(F'(x - B^{FB}(x)) - 1) \). When the low type imitates the high type, it follows from equation (3) that his short-term shareholders also incur a total loss of \( k(F'(x - B^{FB}(x)) - 1) \), but his long-term shareholders incur a total loss of \( (1-k)(F'(x - B^{FB}(x)) - 1) \). Since \( F'(x - B^{FB}(x)) > F'(x - B^{FB}(x)) \) in Figure 3a, long-term shareholders of the imitator experience a greater loss than do long-term shareholders of the high type. In this way, points A and B shift to the left in Figure 3a. At point D, which corresponds to the low type when he imitates the share repurchase level of

---

5 Since the production function is convex/concave, it is necessary, strictly speaking, to check whether the “trivial solution”, \( B(x)=x, I(x)=0, \) and \( D(x)=0 \) is a better solution. In the following it is assumed that this is not the case.
$B^*(x)$, it is assumed that enough costs (corresponding to the 5 mentioned above) have been inflicted upon the low type. Since $F'(x - B) > F'(x - B)$ for all $B \in [B^{FB}(x), B^*(x)]$, lower costs, e.g. only 2, have been incurred by shareholders of the high type company, i.e. the imitator is hurt more than the high type hurts himself.

*Money burning.* Again, the high type has to inflict costs of 5 upon the low type if the low type imitates. If the dividend tax rate is 25%, this can be achieved by paying out taxable dividends of 20 financed by share issues of 20. This hurts the shareholders of the imitating company by 5. However, the shareholders of the high type are hurt by the same amount.

Based on Figure 3a and the arguments above, increased share repurchases/under-investments are chosen by the high type. Quite simply, the imitator has to be hurt by a fixed amount to prevent imitation while the high type chooses the strategy that inflicts the least signaling costs on himself. In Figure 3a, the productivity of the low type when the share repurchase strategy is imitated is always higher on the margin than the productivity of the high type, and hence, increased share repurchases/under-investments offer the high type a better trade-off than money burning does.

The solution so far in Figure 3a corresponds to the Miller and Rock solution where deterrence is achieved solely by increased share repurchases/under-investments. The low type chooses the first best solution, whereas the high type chooses increased share repurchases ($B^*(x) > B^{FB}(x)$)/under-investments ($I^*(x) < I^{FB}(x)$) and no cash dividends.

c2. How the high type deters imitation – the case where $x$ is much higher than $x$

Consider the case depicted in Figure 3b.

(Insert Figure 3b here)

Again, it is assumed that the low type earns an imitation profit of 5 when he imitates the first best share repurchases of the high type. Since the productivity of the low type when imitating (point B) is higher than the productivity of the high type (point A), deterrence by means of increased share repurchases/under-investments, in the beginning, offers the best trade-off on the margin for the high type. In this way points A and B shift to the left in Figure 3b. At points C and
D, corresponding to a repurchase level of \( B'(x) \), it is assumed that costs of only 4 have been inflicted upon the low type (and e.g. a cost of 1 on the high type). If the high type holds onto the repurchase strategy, he will be hurt, on the margin, by more than the low type. Hence, it is better to shift to the money-burning strategy and eliminate the low type’s remaining profit of 1 by paying out 4 in taxable cash dividends (again assuming a dividend tax rate of 25%) financed by reduced share repurchases (or increased share issues). In this way, the low type will in all be hurt by 5 (4+1), as required, and the high type by 2 (1+1). Therefore, the optimal strategy is a combination of a strictly favorable strategy, i.e. share repurchases, which is applied until it becomes unfavorable on the margin, and a one-to-one strategy, i.e. payout of taxable cash dividends financed by share issues.6

Based on the last example the deterrence strategy can be formulated as: The increased share repurchases/under-investments strategy is used as long as it is favorable (hurts the low type more than the high type). If the share repurchases/under-investments strategy becomes unfavorable on the margin, the money-burning strategy, i.e. payout of taxable cash dividends financed by share issues, is used in addition to the share repurchases/under-investments strategy.

d. Discussion

The second part of the deterrence strategy, which gives a role for taxable cash dividends, hinges on the S-shaped production function.7 If, like Miller and Rock, a strictly concave production function is assumed, the productivity of the low type when he imitates the high type will always be higher than the productivity of the high type. Hence, under this assumption share repurchases/under-investments will always be the most favorable deterrence strategy and there will not be any role for taxable cash dividends.

In more general terms, the S-shaped production function implies that the “single-crossing” property (Spence, 1974 and Riley, 1979) is violated, providing a positive role for taxable cash dividends. As mentioned in the introduction, section 4 presents other models where a positive role for taxable cash dividends is obtained, but where the S-shaped production function has been

---

6 If the productivity of the low type when he imitates the first best level of share repurchases of the high type turns out to be below the productivity of the high type – i.e. point B is far to the left in Figure 3b and below 1 – the optimal deterrence strategy will be a combination of over-investments and a payout of taxable cash dividends financed by share issues.

7 It is noted that the optimal levels of investments for both types are situated at the decreasing returns to scale part of the production function. The investment level of the low type may be situated at the increasing returns to scale part of the production function only if the low type imitates the payout pattern of the high type. However, by construction this imitation is successfully deterred.
replaced by other “characteristics”, that also gives rise to a violation of the “single-crossing property”.

The two-type model presented above is a useful device in explaining and understanding the intuition behind the role of taxable cash dividends. However, we pay a price for this simplification. For example, the imitation gain by imitating the high type $x$ is not necessarily a monotonic increasing function of $x$ (see Figure 2). This can be shown to imply that payout of taxable cash dividends may not be a uniform signal of quality in the two-type model. Similarly, it is possible that a higher total payout to shareholders is not a uniform signal of quality in the two-type model. Therefore, in order to derive richer empirical implications of the model, it is necessary to look at the model with a continuum of types. In particular, a model with a continuum of types will ensure that the payout of taxable cash dividends is a monotonic function of quality. The next section gives a heuristic presentation of the basic intuition and results with a continuum of types.

3. Intuition and results based on a continuum of types

3a. Intuition

In this subsection it is argued – in a non-technical manner – that the basic intuition is maintained when the two-type model is replaced with a model based on a continuum of types, i.e. $x \in [x, \bar{x}]$. A detailed and quite technical analysis of the continuum-of-types model can be found in Raaballe and Bechmann (2001). In the following, low types are referred to as types that have low earnings.

Results 1 and 2, characterizing the signaling equilibrium for the continuum-of-types model, are stated and derived below.

*Result 1.* Low types deter imitation solely by means of share repurchases/under-investments.

For all types having earnings $x$ below a critical level $x_*$, the following holds:

i. If type $x$ deter imitation from the type having marginally lower earnings $x - dx$, type $x$ also automatically deters imitation from all lower types.

ii. Deterrence is achieved most favorably by means of increased share repurchases/under-investments, which is to say: Higher share repurchases signal higher earnings ($B'(x) > 1$).
The explanation is as follows. Consider a type with earnings $x$ below the critical level, $x_c$. Suppose this type has deterred imitation from the type of marginal lower earnings $x - dx$ as well as all lower types. Arguing inductively, suppose type $x + dx$ chooses the same level of investments as type $x$, which is to say, type $x + dx$ repurchases $dx$ units more than does type $x$. Thus, in this case, $B(x + dx) = B^*(x) + dx$ (where a $^*$ denotes an optimal value). If type $x + dx$ holds on to this level of share repurchases, it can be shown that type $x$ will earn an imitation profit when he imitates the share repurchases of type $x + dx$. In Figure 4, point A corresponds to type $x + dx$ when he chooses the above level of share repurchases and point B corresponds to type $x$ when he profitably imitates this level of share repurchases.

(Insert Figure 4 here)

Since $F'(x - B(x + dx)) > F'(x + dx - B(x + dx))$, in line with the analysis in the preceding section, deterrence by means of increased share repurchases is more favorable than the money-burning strategy, that is: Imitation by type $x$ is prevented by means of share repurchases at a level of $B^*(x + dx) > B^*(x) + dx$. Then the question can be asked as to whether lower types are also deterred. Take, for example, the company with earnings $x - dx$ whose net gain by imitating the company of type $x$ is by construction equal to zero. If it tries to imitate the company of type $x + dx$, its marginal gross gain is equal to the gross gain of company $x$ when company $x$ imitates company $x + dx$. However, the marginal imitation costs of company $x - dx$ are higher than the marginal imitation costs of company $x$ (see point C in Figure 4). Hence, when type $x$ is deterred, type $x - dx$ is also deterred. This is the argument leading to Result 1.

The only thing that can disrupt this reasoning is the following: Suppose we have a type with very low earnings, $y$, compared to $x$. When this type imitates type $x + dx$, it will have a very low level of investment, $y - B^*(x + dx)$, which is far less than $I_{max}$ (see Figure 4). In this manner, type $y$ can have really low marginal costs of imitation, which is where the critical level of earnings, $x_c$, comes in. $x_c$ is defined as the smallest level of earnings where type $x_c - dx$ as well as a type $y$ ($x \leq y < x_c - dx$) are both on the verge of imitating type $x_c$ when all lower types optimally deter imitation solely by means of the share repurchases/under-investment strategy. If the production function is sufficiently S-shaped and $\overline{\alpha}$ is high enough, such an $x_c < \overline{\alpha}$ will always exist. This is the point at which taxable cash dividends/money burning come into play.
Before discussing the case of \( x > x_c \), two technical implications of the critical level of earnings, \( x_c \), will be stated:

a) From Result 1ii) the level of investments under the share repurchase/under-investment strategy is a monotonically decreasing function of earnings \( x \); that is: In Figure 5 a shift is made from point A and to the left under the share repurchase strategy. However, it can be shown that \( x_c \) will be reached before the level of investments under the share repurchase strategy has reached \( I_{\text{max}} \) (see point B, Figure 5).

b) When \( x_c \) has been reached, two types \((x_c - dx\) and \( y\), where \( x \leq y < x_c - dx \)) are on the verge of imitating type \( x_c \). It can be proven that this will only occur for a type \( y \) \((x \leq y < x_c - dx)\) with \( F'(y - B^*(x_c)) \leq F'(x_c - B^*(x_c))\), i.e. the productivity of the very low type when he imitates is not higher than the productivity of the critical type, \( x_c \).

(Insert Figure 5 here)

These characteristics of the critical level, \( x_c \), allow Result 2 to be stated and explained as:

**Result 2.** High types deter by means of a combination of share repurchases and taxable cash dividends financed by share issues.

*For all types having earnings \( x \) above the critical level, \( x_c \), the following holds:*

i. Such a type \( x \) is on the verge of being imitated by two types that are the type that has marginally lower earnings \( x - dx \) and a type which has significantly lower earnings \( y \). The type with significantly lower earnings has a marginal productivity of \( F'(y - D^*(x) - B^*(x)) \leq F'(x - D^*(x) - B^*(x))\) when he imitates type \( x \).

ii. Deterrence is achieved most favorably by means of share repurchases/under-investments combined with a payout of taxable cash dividends financed by share issues (money burning). \( D'(x) > 0 \) and \( 0 \leq D'(x) + B'(x) < 1 \) implying that investments are an increasing function of earnings.
Inductive argumentation is used again in the following. Consider a type with earnings \( x \) above the critical level, and let \( x \) be signaling by means of \( D^*(x) \) and \( B^*(x) \). Based on the inductive assumption, type \( x \) is on the verge of being imitated by type \( x - dx \) and by a type \( y \) \((x \leq y < x_c - dx)\). The following examines this in two cases dependent on the marginal productivity of type \( y \) when \( y \) is on the verge of imitating type \( x \).

**Case a:** \( F'(y - D^*(x) - B^*(x)) < F'(x - D^*(x) - B^*(x)) \)

Point A in Figure 6a corresponds to the optimal strategy of type \( x \) and point B corresponds to the strategy of type \( y \) if \( y \) is imitating type \( x \).

(Insert Figure 6a here)

Now consider a type \( x + dx \) and assume that he follows the optimal strategy of type \( x \) such that \( D(x + dx) = D^*(x) \) and \( B(x + dx) = B^*(x) \). Based on this strategy, type \( x + dx \) will be situated at point C in Figure 6a and hence, type \( x + dx \) will be imitated by type \( x \) as well as type \( y \) (since type \( y \) is on the verge of imitating type \( x \)). Both of these imitators will earn the same imitation profit, for example, 1. In order to deter imitation, type \( x + dx \) has to inflict costs of at least 1 on type \( x \) as well as on type \( y \). The repurchase strategy offers a favorable trade-off with respect to type \( x \) (see Figure 6a), but an unfavorable trade-off with respect to type \( y \). Hence, the least costly way to eliminate the imitation gain of both imitators is by means of the money-burning strategy, i.e. payout of taxable cash dividends financed by share issues. Accordingly, \( D'(x) = -B'(x) > 0 \) (dosed such that the imitation profit of 1 is eliminated). The optimal strategy of type \( x + dx \) is thus equal to the optimal strategy of type \( x \) plus a dose of the money-burning strategy. For higher types \((> x + dx)\) the implication of this strategy is that type \( y \) (when he is on the verge of imitating) will stay at point B in Figure 6a (since \( D'(x) + B'(x) = 0 \)) and that the higher types will shift to the right in Figure 6a (since \( D'(x) + B'(x) = 0 \) and \( x \) increases). Eventually there will be an \( x \) where \( F'(y - D^*(x) - B^*(x)) = F'(x - D^*(x) - B^*(x)) \) or \( \bar{x} \) will be reached. Case b, discussed below, occurs if equality of the marginal productivities happens first.
Case b: \( F'(y - D^*(x) - B^*(x)) = F'(x - D^*(x) - B^*(x)) \)

Point A in Figure 6b corresponds to the optimal strategy of type \( x \), and point B corresponds to the strategy of type \( y \) when type \( y \) is imitating type \( x \).

(Insert Figure 6b here)

Again, we let the type \( x + dx \) of marginally higher earnings adopt the optimal strategy of type \( x \), which means that type \( x + dx \) is at point C and will be imitated by type \( x \) as well as by type \( y \). As a starting point, the share repurchase strategy offers a favorable trade-off with respect to deterrence of both types. This strategy implies that the points A, B and C all shift to the left in Figure 6b.

Suppose we let \( B(x + dx) = B^*(x) + dx \) and \( D(x + dx) = D^*(x) \). First, it can be shown that this strategy eliminates neither the entire gross imitation gain of type \( x \) nor type \( y \). But more importantly, this is not an optimal strategy since it implies \( F'(x + dx - D(x + dx) - B(x + dx)) = F'(x - D^*(x) - B^*(x)) > F'(y - D^*(x) - B^*(x) - dx) \) (see Figure 6b), i.e. the share repurchase strategy has been dosed too heavily on type \( y \). The optimal strategy is instead as follows. First, choose share repurchases \( B(x + dx) \) such that \( F'(y - D^*(x) - B(x + dx)) = F'(x + dx - D^*(x) - B(x + dx)) \). The implication is that \( 0 < B(x + dx) - B^*(x) < dx \). Part of the imitation profit of types \( x \) and \( y \) has thus been eliminated. Since \( F'(x - D^*(x) - B(x + dx)) > F'(y - D^*(x) - B(x + dx)) \), type \( y \) is now the toughest imitator. The remaining profit of type \( y \) is simply eliminated by means of money burning, i.e. payout of taxable cash dividends financed by share issues. Hence, the optimal strategy of type \( x + dx \) is equal to the optimal strategy of type \( x \) supplemented with a dose of the share repurchase strategy (less than \( dx \)) and a dose of the money-burning strategy. Accordingly, \( 0 < D'(x) + B'(x) < 1 \) and \( D'(x) > 0 \).

The sign of \( B'(x) \) depends on the exact parameter values.\(^8\)

---

\(^8\) In case b the solution is actually somewhat more complicated. However, this does not change any of the qualitative results. The complications are as follows. First, when type \( x \) is on the verge of being imitated by type \( x - dx \) and type \( y \), the toughest imitators of type \( x + dx \) turn out to be type \( x \) and \( y - dy \) (provided type \( y \) is not the lowest type). The implication is that type \( x + dx \) should use the share repurchase strategy a little less and the money burning strategy a little more. Second, when type \( x \) is on the verge of being imitated by type \( x - dx \) and type \( y \), the toughest imitators of type \( x + dx \) may turn out to be type \( x \) and a type with significantly lower earnings than \( y \), i.e. \( y \) may jump discontinuously. The implication is that type \( x + dx \) deters by means of the solution in Figure 6a.
The basic intuition in the case of a continuum of types can be summarized as follows. For low levels of earnings, the share repurchase/under-investment strategy offers a favorable remedy against imitation and the intuition is basically unchanged compared to the two-type model. In Figure 5 the solution moves from point A toward point B for higher levels of earnings. Higher share repurchases (total payout) constitute a uniform signal of quality.

For higher types, the share repurchase/under-investment strategy is not a uniformly favorable remedy against imitation. The basic reason for this is that (much) lower types are now potential imitators. If deterrence is achieved solely by share repurchases/under-investments, a potential imitator will have low productivity when he imitates. Therefore, the optimal means of deterrence is a combination of share repurchases and a payout of taxable cash dividends financed by share issues. Although the mechanics are more complicated, the basic intuition is unchanged compared to the two-type model. Deterrence by means of a combination of share repurchases and money burning implies that the level of investments is an increasing function of earnings, i.e. in Figure 5 movement is from point B and to the right and may eventually end up giving an over-investment problem. Total payouts and taxable cash dividends are increasing functions of earnings and hence, together are a uniform signal of quality.

3b. Empirical implications

In the extended Miller and Rock model, the empirical implications that proceed from the results of this analysis of taxable dividends are as follows:

1. A whole group of firms exists that will not use dividends at all. This is the group of firms with qualities (earnings) below a certain critical level. Instead, cash will be paid out using only share repurchases. The share repurchases can be either positive or negative. For this group of firms, higher payout implies a higher quality firm, and hence also a higher market value of the shares.

2. A contrasting group of firms exists that will also pay out money using dividends. This is the group of firms with qualities (earnings) above a certain critical level.

   a) For these firms, higher dividends imply a higher quality firm, and hence also a higher market value of the shares. Therefore, a positive announcement effect will be associated with an increase in dividends.
b) This group will also use share repurchases. Since share repurchases can be either positive or negative, it is possible to observe dividends being paid out at the same time as shares are being issued.

c) The dividends paid can be smooth relative to earnings for the group of firms paying dividends. This will be the case, for example, if the tax rate on dividends is sufficiently high. Because the first group does not pay out dividends, the dividends will have a tendency to be smooth relative to earnings across all firms.

d) The dividends paid out are a decreasing function of the tax rate on dividends. However, consistent with dividends being a money-burning mechanism, the proceeds from the taxation of dividends turn out to be independent of the tax rate. Furthermore, share repurchases are an increasing function of the tax rate on dividends.

e) With respect to total payout, the implications of the model are that total payout increases with the quality of the firm. Therefore, a higher total payout implies a higher market value of the shares. However, share repurchases may decrease with the quality of the firm.

Finally, two characteristics of the firms influence the extent to which dividends are used: The fraction of short-term shareholders and the degree of asymmetric information. A higher fraction of short-term shareholders leads firms to use cash dividends more often. Similarly, a higher degree of asymmetric information between the management and the shareholders also leads the firms to use cash dividends more often. Here, the degree of asymmetric information should be measured by the range of the set of possible types, i.e. the difference between $\pi$ and $\bar{\pi}$.

Most of these implications are consistent with – or at least not in conflict with – well-known empirical findings on firms’ payout policies as surveyed in, for example, Allen and Michaely (2002). This particularly concerns the implications in 2a, 2b, and 2c, which are all important and well-documented facts about cash dividends. Furthermore, several studies have looked at different aspects of taxes and found evidence consistent with the predictions above. First, there is a negative relation between the relative taxation of dividends and the total level of dividends (Poterba and Summers, 1985, for the UK, and Poterba, 1987, for the US). Second, there is a positive relation
between the dividend tax rate and the share price response per dollar of dividend (Bernheim and Wantz, 1995). Finally, Siddiqi (1995) provide evidence that firms are responding to changes in dividend taxation by appropriately changing the dividends paid.

However, some of the implications are also new and need to be investigated in further detail. This is particularly the case for the two implications with respect to the fraction of short-term shareholders and the degree of asymmetric information. To the best of our knowledge, it has not been studied if there is a relationship between the fraction of short-term shareholders and the extent to which firms pay out dividends. The fraction of short-term shareholders could be proxied by different measures of turnover or free-float or calculated more directly from information on large shareholders or changes in ownership over time. Given the results in this paper, as well as the results from the more classical signaling models, this seems to be an important issue to investigate in future empirical research on dividends.

Evidence of a relationship between the degree of asymmetric information and dividends can be found in Eddy and Seifert (1988), who find that the dividends announcement effects are larger for smaller firms, where the problem of asymmetric information is likely to be larger. Pugh and Jahera (1990) find similar results for share repurchases.

As a final remark, it should be noted that some of the implications must be investigated with care. For example, when measuring $B$ empirically, all payments to claimholders, except dividends, should be included. This means that in addition to share repurchases or issues, repayment of loans should be added to $B$ and new loans subtracted. Furthermore, it may be difficult to control for differences across firms in the production function as well as difficult to obtain proxies for the degree of asymmetric information. Finally, as suggested by Olson and McCann (1994), it may be necessary to distinguish between different types of dividend “policies” when testing signaling hypotheses. Nevertheless, it is empirical investigations that will help to distinguish not only between different models, but also increase our general knowledge about firms’ payout policies.

4. What is the justification for taxable cash dividends?

In general, when a deterrence strategy becomes more costly on the margin for the high type than for the imitator, and thus becomes an unfavorable strategy, it is optimal to change to a money-

---

9 Howe, Vogt and He (2003) show that the stock market reaction to changes in cash dividends is larger for firms with high insider ownership. Therefore, if insiders can be taken to be long-term shareholders, these findings are at least consistent with the implications of the model.

10 This issue is discussed in further detail in Allen and Michaely (2002).
burning strategy. Payout of taxable cash dividends financed by share issues is just such a money-burning strategy.

This is what happens in the extended Miller and Rock model when the level of earnings becomes high enough for the high-quality firm, as analyzed above. This general result is illustrated in this section by three other models providing a role for taxable cash dividends because of a violation of the single crossing property. This shows that other models are able to explain dividends as a way of signaling by burning money which points to the generality of the results in the earlier sections. Finally, a short discussion of the payout of taxable cash dividends financed by share issues compared to other possible money-burning strategies is presented.

4a. Asymmetric information concerning firm value

In the Miller and Rock model, earnings as well as investments are unobservable. This is an assumption that suggests that instead of reporting true earnings and investments, a manager is in a position to cheat and report higher earnings as well as higher investments (e.g. by accounting for less expenditures on the cost accounts and more expenditures on the investment accounts). This increase in investments, however, is merely bookkeeping and does not lead to any increase in production.

Below, it is assumed that earnings as well as investments are observable, i.e. that all parts in the cash flow equation, \( x - I = B + D \), are observable. However, the ability of the manager to invest is now unobservable to investors. More precisely, the manager knows his own ability or type \( \theta \), which can be bad (\( \theta = \theta^b \)) or good (\( \theta = \theta^g \)), but this ability is unknown to investors. If the manager is of type \( \theta \), the expected outcome of the investments in the next period is given by the production function \( F(I, \theta) \), about which it is assumed that:

\[
F_I(I, \theta) \geq 0 \quad F_{II}(I, \theta) \leq 0 \quad \text{for } I \geq 0 \\
F(I, \bar{\theta}) > F(I, \theta) \quad \text{for } I \geq 0 \\
F_I(I, \bar{\theta}) > F_I(I, \theta) \quad \text{for } 0 \leq I < I^* \\
F_I(I, \bar{\theta}) < F_I(I, \theta) \quad \text{for } I^* < I.
\]

The intuition behind this production function is that the skilled manager, \( \theta = \bar{\theta} \), can implement the project with fewer resources than the unskilled manager. In Figure 7, a simple example of a production function fulfilling these assumptions is given.

21
With respect to the objective function it is still assumed that a weight of \( k \) is put on short-term investors and a weight of \( 1 - k \) is put on long-term investors.

Figure 8 illustrates the marginal productivities of the two types based on a “smooth” production function fulfilling the assumptions of this subsection. \( I^* \) is the unique level of investments where \( F_I(I^*, \theta) = F_I(I^*, \overline{\theta}) \).

If one assumes that, independent of type, the manager chooses the first best solution, then his first best level of investments is given by:

\[
F_I(I^{FB} (\theta), \theta) = F_I(x - D^{FB} (\theta) - B^{FB} (\theta), \theta) = 1 \quad \theta = \theta, \overline{\theta}.
\]

Noting that \( D(\theta) > 0 \) is not a first best solution, since the only effect of decreasing \( D(\theta) \) by 1 and increasing \( B(\theta) \) by 1 is that the shareholders receive an additional after-tax liquidity of \( (1 - \alpha) > 0 \), means that the first best solution is given by:

\[
D^{FB} (\theta) = 0 \quad \theta = \theta, \overline{\theta}
\]

\[
F_I(I^{FB} (\theta), \theta) = F_I(x - B^{FB} (\theta), \theta) = 1 \quad \theta = \theta, \overline{\theta}.
\]

With the first best solution, the high type chooses a lower level of investments and hence a higher level of share repurchases than the low type does (the high type is situated at point A and the low at point B in Figure 8). Given the assumption of asymmetric information, it is therefore natural to consider the construction of a separating equilibrium.

Assume that the low type can earn a positive profit, say 5, by imitating the first best level of share repurchases of the high type (\( B^{FB} (\overline{\theta}) \)). If the low type imitates the first best solution of the high type, the high type will be situated at point A and the low type at point C in Figure 8. In order to deter imitation by the low type, the high type has to inflict costs of 5 on the low type. Since the marginal productivity of the low type when he imitates is higher than the marginal productivity of
the high type at $I^{FB}(\bar{\theta})$ (see Figure 8), the most favorable means of deterrence is the share repurchase/under-investment strategy. As a consequence, points A and C in Figure 8 shift to the left. If the costs of 5 have been inflicted upon the imitator before the share repurchase strategy has reduced the level of investments to $I^*$, the share repurchase/under-investments strategy will be the favorable/optimal strategy throughout. However, if costs of only, for example, 3 have been inflicted upon the imitator when $I^*$ is reached, it will no longer be optimal to continue the under-investment strategy since this strategy has now become unfavorable. It is simply cheaper to inflict the remaining costs of 2 upon the imitator by means of the money-burning strategy: Payout of taxable cash dividends financed by share issues.

In the appendix a simple numerical example of this model is offered. First, the example verifies that a production function exists such that the low type can earn a positive profit by imitating the first best level of share repurchases of the high type. Second, the example shows that in this case deterrence is achieved most efficiently by a combination of share repurchases and money burning.

4b. When management likes investments in addition to total firm value

Up to this point, it has been assumed that the only objective of the manager has been to maximize total firm value for the shareholders where a weight of $k$ has been put on the current share price and a weight of $1-k$ has been put on the true value of the shares.\(^{11}\) In addition to this objective, it is now assumed that the manager also has a desire for investments (or size) measured by $\int IV$ where $0<\int IV$. This is consistent with much of the literature on management incentives – see for example Jensen (1986). To model this incentive, it is assumed that a weight of $h$ is put on the share price objective and a weight of $1-h$ is put on the size objective. Hence, in a signaling equilibrium where $D(x)$ and $B(x)$ reveal $x$, the value of the objective function is given by:

$$h[B(x) + \alpha D(x) + kF(x - D(x) - B(x)) + (1-k)F(x - D(x) - B(x))] + (1-h)V(x - D(x) - B(x))$$

It is now assumed that the production function $F(I)$ is concave for all levels of investments (like the original Miller and Rock assumption). If the manager’s desire for investments is also a concave function for all levels of investments, it can be proven that the share repurchase strategy

\(^{11}\) In the signaling equilibrium, the current share price is of course equal to the true value of the share.
will always be the most favorable deterrence strategy, i.e. there is no role for taxable cash dividends.

On the other hand, if the manager’s desire for investments is a S-shaped function of the level of investments and the weight put on this desire, \(1-h\), is at an intermediary level, it can be shown that the share repurchase strategy is not a favorable deterrence strategy on the margin for companies with high earnings.

An S-shaped \( V(I) \) roughly implies that the manager does not find much pleasure in a marginal increase in investments if the level of investment is currently at a very low and unsatisfactory level. However, this also implies that the manager’s loss of pleasure in firm size and production for the staying shareholders of the low type company can be smaller than the similar costs of the high type when he imitates the share repurchase strategy of the high type. Hence, payout of taxable cash dividends financed by share issues will be the preferred strategy on the margin.

This model has been analyzed thoroughly based on a continuum of types \( x \in [\underline{x}, \bar{x}] \) (see Raaballe and Bechmann, 2000). The empirical implications are almost the same as described in subsection 3b, i.e. as obtained in the extended Miller and Rock model (S-shaped production function and no managerial incentives).\(^{12}\)

4c. Are under-investments a credible means of deterrence?

In the original Miller and Rock model, high types under-invest in order to deter imitation from the low types. This ensures that the selling shareholders obtain a fair share price. However, when the shares have been sold, the new and the staying shareholders suffer from under-investment, i.e. they want to raise the level of investments to the first best level of investments by a second injection of funds, e.g. financed by a new issue of shares. If that was possible at no cost, there would be no deterrence associated with the original investment cut and hence no separating equilibrium. Hence, an implicit assumption in the Miller and Rock model is that it is too costly to undo the investment choice through a second injection of funds. However, if the investment choice can be redone by a second injection of funds at strictly positive costs (increasing in the absolute magnitude of the correction and neither too low nor too high), it can be proven that this

\(^{12}\) One exception is with respect to the level of investments. Because of the managerial incentives for size, the lowest type over invests compared to the Fisher optimum. Higher types having earnings lower than a critical level deter by means of excessive share repurchases. The implication is that the investments decrease as a function of earnings and eventually turn into under-investments. For companies having earnings above the critical level, the investments increase as a function of earnings and may again turn into over-investments.
two-stage investment process turns the original concave production function into an S-shaped production function. Based on this, the analysis of subsection 2b can be applied almost unchanged.

4d. Is the payout of taxable cash dividends the only attractive money-burning strategy?

Above, several models were provided wherein the best choice of deterrence strategy (on the margin) can be a money-burning strategy such as the payout of taxable cash dividends financed by share issues. Fortunately, this provides a sound explanation of the dividend puzzle: Why are cash dividends paid out despite the tax disadvantage? The answer is simply that they burn money. What remains is the question: Why are taxable cash dividends financed by share issues the preferred way of money-burning? We have no firm answer to that question.

Over time and across countries, huge amounts of tax disadvantaged cash dividends have been paid out. No other attractive money-burning strategy seems to be available on a large scale. It would seem that the payout of taxable cash dividends is an especially credible and efficient way to burn money. For one thing, there is no way that the company or its shareholders will get their money back from the government; for another, the shareholders’ payment of dividend taxes does not promote sales or influence the productivity of the company. Therefore, the payout of taxable cash dividends is close to literally burning money.

It may be argued that a stock split, for example, also fulfills the two conditions above and hence, might serve as a money-burning alternative to taxable cash dividends. However, from a market microstructure point of view, a stock split may also have liquidity effects. Furthermore, and more importantly, most of the money is seen to be wasted by means of taxable cash dividends and not by stock splits or similar actions. In a similar vein, it can be argued that the use of excessively costly investment bankers could also be a way of burning money. An issue of under-priced shares to new shareholders may also be considered a money-burning strategy. However, in most countries this is simply prohibited by law, the exception being initial public offerings.

Charity, uninformative advertising, excessive pay rolls and monuments could also all be examples of money burning. However, we doubt that these money-burning methods fulfill the two conditions above.

The arguments and reasoning above are also consistent with the results in Bernheim and Redding (2001), showing that, given the existence of several possible ways of burning money, the firm should use the money burning mechanism that allows for the “cleanest” inference. Therefore,

13 Details are available from the authors upon request.
given that cash dividends arguably are a more precise signal compared to the other ways of money burning, this is an explanation as to why cash dividends are so heavily used.

5. Conclusions

This paper has focused on explaining the use of share repurchases and taxable cash dividends as signaling devices. This has been done by considering a series of two-type models that are generally based on palatable assumptions. In these models, it has been shown that it is not always possible for a really high type to deter imitation from a low type by means of a strictly favorable deterrence strategy, a strategy that hurts the low type strictly more than it hurts the high type. The explanation is that if the high type adheres to such a strategy, the marginal productivity of the low type when he imitates will be lower than the marginal productivity of the high type (a violation of the “single-crossing” property) before enough costs have been inflicted upon the imitator. Thus, before this happens it will be more profitable for the high type to shift to a money-burning strategy such as paying out taxable cash dividends. This will hurt both types equally hard and shows the justification for the use of taxable cash dividends.

The two-type models are useful for explaining the role of taxable cash dividends and helping us to understand that role. However, it is only rarely possible to tell if the quality of a company is either low or high. More often, it is possible to tell that the quality of the company is in a certain interval. A (parametrically) two-type model gives results that are different from results provided by a similar model based on a continuum of types. Therefore, from a theoretical as well as an empirical point of view, it is important to analyze the predictions based on a continuum of types. This has been done for two of the models, and the predictions turn out to be rich and are generally consistent with the empirical facts. Furthermore, the empirical predictions also include new predictions that will be interesting to examine in the future.

The good news from this paper is that it provides a sound theoretical explanation for the dividend puzzle, i.e. why firms pay out taxable dividends even though it amounts to burning money. It has been found that money burning on the margin is the most effective way for a firm to signal that its quality is really high. The explanation as to why this specific money-burning strategy has been chosen seems to be that cash dividends are the “cleanest” way of burning money, i.e. they allow the cleanest inference of the signal.

All in all, we find that modeling based on a violation of the “single-crossing” property turns out to be productive in explaining the role of taxable cash dividends.
Appendix

This example illustrates the model in subsection 4a. It is assumed that:

\[ x = 100 \quad k = \frac{1}{2} \quad (1 - \alpha) = 0.25 \quad F(0, \theta) = F(0, \bar{\theta}) = 0 \]

\[
F_I(I, \theta) = \begin{cases} 
1.5 & \text{for } 0 \leq I < 110 \\
0.75 & \text{for } 110 \leq I < 120 \\
0.5 & \text{for } 120 \leq I \\
2 & \text{for } 0 \leq I < 90 \\
1.5 - 0.05(I - 90) & \text{for } 90 \leq I < 110 \\
0.5 & \text{for } 110 \leq I 
\end{cases}
\]

We note that the production function fulfills the conditions stated on page 21 and that \( I^* = 90 \).

The first best solutions are given by:

\[
I_{FB}^*(\theta) = 110 \quad B_{FB}^*(\theta) = -10 \quad D_{FB}^*(\theta) = 0 \quad \pi_{FB}^*(\theta) = -10 + 165 = 155
\]

\[
I_{FB}^*(\theta) = 100 \quad B_{FB}^*(\theta) = 0 \quad D_{FB}^*(\theta) = 0 \quad \pi_{FB}^*(\theta) = 0 + (180 + 12.5) = 192.5.
\]

If the low type imitates the payout pattern of the high type \( B_{FB}^*(\theta) = D_{FB}^*(\theta) = 0 \), his gross profit is:

\[
0 + \frac{1}{2} F(100, \theta) + \frac{1}{2} F(100, \bar{\theta}) = \frac{1}{2} \cdot 192.5 + \frac{1}{2} \cdot 150 = 171.25.
\]

Hence, the low type earns a positive profit of 16.25 (= 171.25 – 155) by imitating the first best level of share repurchases of the high type. This demonstrates that a production function exists such that the low type can earn a positive profit by imitating the first best level of share repurchases of the high type.

Since \( F_I(I, \theta) < F_I(I, \bar{\theta}) \) for \( 90 < I \leq 100 \), the most favorable means of deterrence is the share repurchase/under-investments strategy until \( I^* = 90 \) has been reached. If the high type
increases share repurchases by 10 (decreases investments by 10), this costs the high type 2.5 (= 192.5 – 10 – 180). When the low type imitates this increase in share repurchases, it costs the low type $\frac{1}{2} \cdot 2.5 + \frac{1}{2} \cdot 5 = 3.75$, i.e. a favorable trade-off. According to this, the imitation profit of the low type has been reduced to $16.25 – 3.75 = 12.5$. Since $I^* = 90$ now has been reached, additional share repurchases are no longer a favorable strategy.

The low type’s remaining imitation profit of 12.5 is now eliminated by means of a payout of taxable cash dividends of 50 financed by a share issue, which costs both types 12.5.

The profit of the high type is then given by:

$$\pi(\tilde{\theta}, B = 10 - 50, D = 50) = 0.75 \cdot 50 - 40 + F(90, \tilde{\theta})$$

$$= 37.5 - 40 + 180 = 177.5 (= 192.5 - 2.5 - 12.5).$$

The profit of the low type (when he imitates) is given by:

$$\pi(\theta, B = 10 - 50, D = 50) = 0.75 \cdot 50 - 40 + \frac{1}{2} F(90, \theta) + \frac{1}{2} F(90, \theta)$$

$$= 0.75 \cdot 50 - 40 + \frac{1}{2} \cdot 180 + \frac{1}{2} \cdot 135 = 155 (= 171.25 - 3.75 - 12.5),$$

confirming that the low type will experience that imitating the high type is not profitable.

If the high type deters imitation solely by means of the share repurchase strategy, a share repurchase of 26.67 is needed. The profit of the high type is then given by:

$$\pi(\tilde{\theta}, B = 26.67, D = 0) = 26.67 + F(73.33, \tilde{\theta}) = 173.33 < 177.5.$$

The profit of the low type (when he imitates) is given by:

$$\pi(\theta, B = 26.67, D = 0) = 26.67 + \frac{1}{2} F(73.33, \theta) + \frac{1}{2} F(73.33, \theta) = 155,$$

confirming that the low type will experience that imitating the high type is not profitable.

Hence deterrence can be achieved most efficiently by a combination of share repurchases and money burning.
References


455-476.


Bhattacharya, S. (1979), “Imperfect Information, Dividend Policy, and ‘the Bird in the Hand’ 


New York.

of Finance*, 45: 993-1019.


Chapter 23 in *Finance* by Jarrow, Maksimovic, and Ziemba, editors (1995).


Hypothesis,” *Journal of Finance*, 57: 1649-1684


**Figure 1a:** The S-shaped production function, $F(I)$, as a function of the level of real investments, $I$.

**Figure 1b:** The first best level of investments, $I_{FB}$, given by $F'(I_{FB}) = 1$, i.e. by point A. This defines the corresponding first best levels of share repurchases, $B_{FB}$, for the two types of firms.
Figure 2: The gain to type $x$ when imitating types $x'$ and $x''$, $x < x' < x''$. When imitating type $x'$, the gain corresponds to area A, while the gain by imitating type $x''$ is area $A + B - C + D$. Note that $F'(\bar{x} - B^{FB}(\bar{x}) - \tau) = 1$ for $\tau = 0$.

Figure 3a: The high type, $x$, prevents imitation from the low type, $\bar{x}$. Point A corresponds to the case where the high type holds on to the first best solution. Point B corresponds to the case where the low type imitates the payout of the high type. When the high type increases the level of share repurchases, $B(x)$, points A and B move to the left. At point D, which corresponds to the low type when he imitates the share repurchase level of $B'(\bar{x})$, it is assumed that the low type no longer has any gain by imitating.
Figure 3b: At point B the marginal productivity of the low type when he imitates is higher than the productivity of the high type (point A). With increased share repurchases/under-investments, points A and B move to the left. At points C and D, which correspond to a repurchase level of $B'(x)$, it is assumed that the low type still has a gain when imitating. At points C and D increased share repurchases will imply that the marginal productivity of the low type when he imitates is lower than the productivity of the high type.

Figure 4: Point A corresponds to type $x+dx$ when he chooses the same level of investments as type $x$, and point B corresponds to type $x$ when he profitably imitates this level of share repurchases, $B(x+dx)$. Point C shows the situation for type $x–dx$ when he imitates type $x+dx$. 
Figure 5: With the share repurchase strategy, movement is from point A and to the left. Point B corresponds to the level of investments for the critical type $x_c$. With the share repurchase strategy, the level of investments will reach point B before it reaches $I_{\text{max}}$.

Figure 6a: Point A corresponds to the optimal strategy of type $x$ and point B corresponds to the strategy of type $y$ when he is on the verge of imitating type $x$. At point B the marginal productivity is lower than at point A. Point C corresponds to the situation for type $x+dx$ if he follows the optimal strategy of type $x$. 

Figure 6b: Point A corresponds to the optimal strategy of type \( x \) and point B corresponds to the strategy of type \( y \) when he is on the verge of imitating type \( x \). At point B the marginal productivity is equal to the marginal productivity at point A. Point C corresponds to the situation for type \( x+dx \) if he follows the optimal strategy of type \( x \).

\[
F'(I) \\
B \quad A \quad C \\
\theta \quad \text{IF} \quad \text{IF} \quad \text{IF} \\
I \\
y - D^*(x) - B^*(x) \quad x - D^*(x) - B^*(x) \quad x + dx - D^*(x) - B^*(x)
\]

Figure 7: This figure shows a simple example of a production function fulfilling the assumptions in subsection 4.a for two different types of firms; a high quality firm \( \bar{\theta} \) and a low quality firm \( \underline{\theta} \).
Figure 8: This figure shows the marginal productivities of the two types based on a production function fulfilling the assumptions in subsection 4.a. $I^*$ is the unique level of investments where $F_I(I^*, \theta) = F_I(I^*, \bar{\theta})$. Point A corresponds to the case where the high (low) type has chosen the first best level of investments, whereas B corresponds to the case where the low type has chosen the first best level of investments. Point C corresponds to the case where the low type imitates the first best level of investments for the high type.
| WP 2000-1 | Bjarne Astrup Jensen and Carsten Sørensen: Paying for minimum interest rate guarantees: Who should compensate who? |
| WP 2000-2 | Jan Jakobsen and Ole Sørensen: Decomposing and testing Long-run Returns with an application to initial public offerings in Denmark. |
| WP 2000-9 | Claus Munk and Carsten Sørensen: Optimal Consumption and Investment Strategies with Stochastic Interest Rates. |
| WP 2000-11 | Henrik Lando: Ny lov om jordforurening i økonomisk belysning. |

| WP 2001-1 | Michael Møller, Claus Parum og Thomas Sørensen: Den ny pensionsafkastbeskatningslov. |
WP 2001-3 Caspar Rose: Impact of Investor Meetings/Presentations on Share Prices, Insider Trading and Securities Regulation.

WP 2001-4 Caspar Rose: Corporate Financial Performance and the Use of Takeover Defenses.

WP 2001-5 Shubhashis Gangopadhyay and Clas Wihlborg: The Impact of Bankruptcy Rules on Risky Project Choice and Skill Formation under Credit Rationing.

WP 2001-6 Claus Munk, Carsten Sørensen & Tina Nygaard Vinther: Portfolio Choice under Inflation: Are Popular Recommendations Consistent with Rational Behaviour?


WP 2001-8 Michael Møller & Caspar Rose: Legal pre-emption rights as call-options, redistribution and efficiency loss.


WP 2002-1 Bjarne Astrup Jensen: On valuation before and after tax in no arbitrage models: Tax neutrality in the discrete time model.


WP 2002-3 Steen Thomsen and Caspar Rose: Foundation ownership and financial performance. Do companies need owners?

WP 2002-4 Martin Richter and Carsten Sørensen: Stochastic Volatility and Seasonality in Commodity Futures and Options: The Case of Soybeans.


WP 2002-6 Caspar Rose: Impact of Takeover Defenses on Managerial Incentives.

WP 2002-7 Ken L. Bechmann og Peter Løchte Jørgensen: Optionsaflønning i danske børsnoterede selskaber.

WP 2002-8 Jesper Rangvid: Output and Expected Returns – a multicountry study.


WP 2004-2 Caspar Rose: Bestyrelsesarrangement og finansiel performance i danske børsnoterede virksomheder - Er Nørbyrapportens anbefalinger til gavn for aktionærerne?


WP 2005-3: Jensn Lunde: The owner-occupiers’ capital structure during a house price boom -Does negative equity exist as a permanent feature in the Danish housing market?