Dividend Taxation and Intertemporal Tax Arbitrage*

Anton Korinek[†] and Joseph E. Stiglitz[‡]
Columbia University

November 26, 2006

Abstract

Dividend taxation makes it costly for firms to distribute funds to shareholders; on the other hand, agency problems limit the amount of funds that shareholders are willing to leave with firms. Firms have to balance agency costs and tax consequences to determine their optimal amount of working capital.

Building on these considerations, we develop a life-cycle model of the firm to analyze the effects of dividend tax policy on aggregate investment. We find that new firms raise less equity and invest less the higher the level of dividend taxes, in accordance with the traditional view of dividend taxation. However, the dividend tax rate is irrelevant for the investment decisions of internally growing and mature firms, as postulated by the new view of dividend taxation. Since aggregate investment is dominated by these latter two categories, the level of dividend taxation as well as unanticipated changes in dividend tax rates have only a minor impact on aggregate investment and output.

Anticipated dividend tax changes, on the other hand, allow firms to engage in inter-temporal tax arbitrage so as to reduce investors' tax burden. This can significantly distort aggregate investment. Anticipated tax cuts (increases) delay (accelerate) firms' dividend payments, which leads them to hold higher (lower) cash balances and, for capital constrained firms, can significantly increase (decrease) aggregate investment for periods after the tax change.

Furthermore, we show that the analysis of dividend taxation in a contestable democracy has to take into account expectations about future regime changes and the ensuing dividend tax changes. This can significantly change the evaluation of a given dividend tax policy.

JEL Codes: G35, G38, H32

Keywords: dividend taxation, capital constraints, aggregate investment, political economy of taxation

---

*The authors would like to express their thanks for helpful comments to Tim Besley, Raj Chetty, James A. Mirrlees and Bernard Salanie as well as to participants at the conference in Honor of Mirrlees' 70th birthday at Cambridge University, a workshop at the University of Manchester, and the NBER Public Economics Program Meeting in Nov. 2006.

[†]Korinek acknowledges financial support from the Austrian Academy of Sciences. Address for correspondence: 814 Uris Hall, Columbia University, 3022 Broadway, New York, NY 10027, USA. Telephone number: +1 (917) 543-5237. Email address: ak2038@columbia.edu

[‡]Stiglitz acknowledges financial support from the Ford, MacArthur, and Mott foundations.

1 Introduction

The economic effects of dividend taxation have been at the center of a fierce academic debate for decades. Proponents of the 'traditional view' of dividend taxation (as discussed e.g. in Poterba and Summers, 1985) stress that it raises the cost of equity finance. Hence, they argue, it distorts firms' investment decisions, and higher dividend taxes decrease the long-run capital intensity of an economy. The 'new' or 'tax capitalization view,' by contrast, which was developed in King (1977) and Auerbach (1979), extending an earlier argument of Stiglitz (1973),[1] assumes that most firms use retained earnings as the marginal source of investment, and for the investment decisions of these firms the level of dividend taxation is thus irrelevant.

In this paper, we explicitly take this controversy into account by modeling the life cycle of firms. When new firms are started, their marginal source of finance is equity markets, i.e. they issue new equity. As predicted by the traditional view, dividend taxation is distortionary for these firms. Next, firms grow internally by retaining their earnings until they reach their optimal size. At this point, they start paying out dividends.[2]

As postulated by the new view of dividend taxation, the level τ of dividend taxes does not affect investment decisions of internally financed growing and mature firms: it reduces the marginal cost and the marginal return of firm investment by an identical factor $1-\tau$. In firms' maximization problem, this factor cancels out and leaves optimal investment unchanged. Since growing and mature firms dominate aggregate investment, we argue that the effect of dividend taxation on investment is very low at a macroeconomic level.

We also assign an important role to capital market imperfections: In order to emphasize the importance of working capital for firms' investment decisions, we assume that firms need to carry cash on their balance sheet in order to take advantage of investment opportunities. However, since holding cash is costly and investment op-

[1] Stiglitz (1973) demonstrated that it was more tax-efficient to distribute funds through buying back shares than by paying dividends. The fact that so much money is nonetheless distributed through dividends has subsequently come to be called the dividend puzzle (see Black, 1976).

We do not address the problem of the dividend puzzle in this article – we simply assume that all payouts from the firm are subject to the dividend tax, as if there were no stock repurchases. The results of our analysis are valid so long as firms engage – at the margin – in intertemporal arbitrage. Even if they irrationally distribute funds in the form of dividends, they can be and should be sensitive to the fact that by distributing dividends in low-tax periods they increase the value of the firm, because the present discounted value of the after-tax dividend stream is higher.

[2] We are, of course, not the first to develop a life cycle theory of the firm - with differences in finance (and therefore in the impact of taxation) in different stages of the firm. In Stiglitz (1973, 1976) and Atkinson and Stiglitz (1976), the typical life cycle has three stages - firms first raise money through equity; they then finance investment through retained earnings plus borrowing; then they only invest retained earnings.

Sinn (1991) formulates a general equilibrium model in which the firm goes through these three stages: in the first it issues equity, in the second it grows by reinvesting all earnings; in the final stage, it pays out its earnings in the form of dividends.

Neither model examines corporate finance in a stochastic environment. In such a world, if profits exceed desired investment (defined by the point where the real interest rate equals the expected marginal product of capital), the excess is either retained and held in T-bills, or distributed, depending on details of relative tax rates and investors' relative valuation of dividends versus cash inside the firm.

portunities arrive randomly, firms choose an equilibrium level of cash balances that trades off the expected benefit of investing against the opportunity cost of holding cash balances, and this entails that they are capital constrained in equilibrium.

We then employ the outlined model to investigate the effects of *changes* in the dividend tax rate on macroeconomic variables such as aggregate investment and output. Our analysis shows that the impact of unanticipated tax changes is generally small, since they affect only new firms that access equity markets. By contrast, anticipated tax changes entail effects that can be an order of a magnitude higher: they allow firms to engage in inter-temporal tax arbitrage by shifting dividend payments from high-tax periods to low-tax periods. This involves significant deviations from firm's optimal steady state level of cash holdings, and since firms are capital constrained, changes in their cash holdings can significantly distort aggregate investment and output.

An anticipated dividend tax cut allows firms to reduce investors' tax bill by postponing dividend payments to the period in which the tax cut takes place. This implies that firms carry larger cash balances in the meantime, which allows them to make larger investments when an investment opportunity arrives. By the same token, anticipated dividend tax increases create an incentive for firms to pay out a large special dividend before the tax increase takes place, which leaves them with lower cash balances, that lead in turn to lower investment and output.

Temporary dividend tax changes can be viewed as an unanticipated change followed by a second anticipated change in the opposite direction. As we argued, the effects of this second anticipated change are likely to far outweigh any effects of the preceding unanticipated change. A temporary dividend tax cut, for example, is thus likely to have an overall negative effect on aggregate investment and output. By the same token, uncertainty about an impending dividend tax increase can lead firms to reduce their cash holdings, with similar negative macroeconomic effects.

Lastly, we analyze the political economy of dividend taxation in a contestable democracy with two parties, denoted as conservatives and social democrats, that have different preferences regarding the level of the dividend tax rate. In the non-cooperative game between the two, each party adjusts the dividend tax rate when it comes to power. Paradoxically, aggregate investment is higher under social democratic rule, since firms expect a tax cut when they lose power and thus hold higher cash reserves. Under a conservative regime, firms fear that dividend taxes will go up when the party loses power and thus pay out a larger amount of dividends, which depresses their cash holdings and reduces aggregate investment.

Aside from providing new insights into the consequences of the temporary dividend tax cuts that were introduced in the United States in 2003 and extended in 2006, this paper also makes a methodological contribution to the theory of political economy of more general import. In contestable democracies, the effects of policies have to be analyzed in the context of an environment in which market agents expect (stochastically) a change in decision makers. It is wrong simply to assume, as much of the political economy literature has done heretofore, that there is a set of well-defined preferences for political decision makers.[3] Because the alternation of decision makers – and decisions – affects, of course, the behavior of economic agents, it also affects the

[3]Of course, the nature of the stochastic process describing changes in the party in power is a matter for further study, and is itself endogenous. The simple model presented here abstracts from these complexities.

behavior of decision makers: they have to take into account the consequences of any policy under the assumption that economic agents will not believe that the policy is permanent (no matter what the government might say.) Economic agents, of course, know that, and rationally take that into account in making their decisions; and political decision makers know that too – and take into account the economic agents' rational responses.[4]

In the United States dividends used to be taxed at the personal income tax rate of up to 38.6% until 2003. CBO (2003) estimated the effective overall tax rate on dividends back then to be around only 19%, since many savings vehicles and shareholders were exempt from dividend taxation. The Jobs and Growth Tax Relief Reconciliation Act of 2003 (JGTRRA) reduced the personal income tax on corporate dividends to a nominal rate of only up to 15% or an estimated effective rate of 5% until 2008 (CBO, 2003). As a result, national income 2004 included a record amount of $493bn or 4.8% of dividend income, a fraction not seen since 1937.[5] $32 billion alone originated from a special dividend paid by Microsoft. Blouin et al. (2004) find that the increase in aggregate dividend payments after the 2003 tax cut was concentrated on firms paying large special dividends. This evidence supports our view that companies engaged in inter-temporal tax arbitrage by shifting dividend payments from high- towards low-tax periods.[6]

Several studies, such as Chetty and Saez (2005), indicate that the tax reduction had a significant effect not only on firms' dividend payments but also on share prices, which entails that firms' marginal cost of equity finance was reduced. This is generally viewed as proof that the dividend tax cut had an overall positive effect on aggregate investment.

Our paper argues that this view is incomplete, if not incorrect. We demonstrate that the small positive effects of the observed increase in share prices and new equity issuance are likely to be outweighed by the negative effects created by the expectations about impending dividend tax increases in case party rule in the Senate changes or

[4]The similarity between this problem and that in the standard analysis of models with asymmetric information is obvious: there the informed party knows that the uninformed will be making inferences based on his actions; and the uninformed party knows that the informed party knows that. The self-selection/signaling equilibrium resolves this infinite regress, as does the contestable political equilibrium of section 7 in this paper.

[5]Data from the webpage of the Bureau of Economic Analysis, U.S. Department of Commerce, http://www.bea.gov, table 1.12.

[6]Using a longer sample, Chetty and Saez (2006) find that there was also a significant increase in regular dividend initiations. While some of this is certainly attributable to a shift from stock repurchases to dividend payments, our model does not preclude this outcome. We assume that firms maximize the present discounted value of the net (after tax) income of their owners. But with different individuals facing different tax rates (and even different discount rates, alternative investment opportunities), the policies that different shareholders would want the firm to pursue differ. This is especially relevant if shareholders differ in their judgments of risk and risk aversions (Grossman and Stiglitz, 1977); the theory of managerial capitalism emphasizes that managers may choose the policy that maximizes their own well-being, instead of that of their shareholders. Managers with significant shareholdings are more likely to distribute profits in tax-efficient ways and have incentives to respond more strongly to tax changes. Managers who hold options (with a strike price that is typically not adjusted for dividend payments) may be less inclined to pay out dividends, even if a change in tax rates would make this attractive for shareholders. This is consistent with the evidence presented in Chetty and Saez (2006).

when the tax cut expires in 2010.

When firms anticipate a dividend tax increase, they have an incentive to pay out higher dividends as long as the low tax rate is in effect. If they are subject to capital constraints, then the ensuing reduction in their cash balances implies that aggregate investment and output are likely to be negatively affected, as compared to the benchmark case of no change in dividend taxation.

Furthermore, we demonstrate that the temporary tax cuts of 2003 and 2006 increased share prices, i.e. lowered the cost of equity, mainly for those firms that do not issue equity (mature firms). By contrast, they hardly affected share prices and the cost of equity for those firms that do issue equity (new firms). The reason is that new firms typically do not pay out dividends in the first few years of their existence, i.e. when the dividend tax rate is low, since they retain all their earnings in order to grow. By the time that they start paying dividends, the temporary tax cut is likely to be reversed and shareholders pay the high dividend tax rates. Since investors rationally foresee this, they are not willing to provide more capital to new firms if they know that a given dividend tax cut is only temporary. It is noteworthy that this policy increased the wealth of shareholders in mature "old" industries, while not having much impact on new and growing firms.[7]

In the next section we develop a basic model of capital constrained firms subject to random investment opportunities. Sections 3 and 4 discuss the steady state of such firms and add an analysis of how new firms access equity markets. Section 5 investigates the effects of unanticipated and anticipated dividend tax changes on firm behavior. In section 6 we discuss how dividend taxation and changes in the tax rate affect aggregate investment and output. Section 7 presents a simple model of the political economy of dividend taxation. Section 8 concludes.

2 Model

2.1 Motivation

Underlying the model developed here are three key economic considerations:

(a) Holding cash is costly because of agency factors. Investors follow the principle that "one bird in your hand is worth more than one in a bush." Because of the agency problems that arise when firms hold a large amount of cash[8] and because of the economic risks that firms are subject to, investors value cash on hand more highly than cash on companies' balance sheets.

(b) On the other hand, because of capital market imperfections, firms cannot quickly borrow or raise new capital for investment purposes – they have to rely on their working capital instead. In our model, these constraints become particularly important because we assume that investment opportunities arrive at random.[9]

[7] Thus our paper provides some support for one of the political critiques of the 2003 dividend tax cut, that it benefitted the interests of old and stagnant industries rather than those of young and dynamic industries.

[8] The dangers of excessive cash holdings became particularly clear in the case of oil companies, after the oil price shocks of the 1970s. For an overview of the issue, see e.g. Jensen (2001).

[9] Theories of asymmetric information, such as Stiglitz and Weiss (1981) and Myers and Majluf

(c) Firms thus have an incentive to hold on to working capital – to give them the resources to take advantage of new opportunities as they arrive – even though the rate of interest that they earn on cash is lower than their discount rate. But this has to be counterbalanced with the risk of their cash balances being squandered – the agency costs. Taxes complicate the analysis further, especially when they are expected to change. If they are expected to decrease, firms have an incentive to hold more cash balances – not to distribute – waiting until dividend taxes decrease. There is a cost to delay; but, up to a point, the benefit exceeds the cost. On the other hand, if taxes are expected to increase, firms have an incentive to distribute now, before the increase, even though that means that the firms might be unable to take advantage of some investment opportunities that arise in the future.

Our model allows us to analyze precisely these intertemporal trade-offs. Suppose that the amount of cash a firm has on hand at the beginning of period t is M_t. The firm decides how much to pay in dividends D_t and keeps its remaining cash holdings $M_t - D_t$ on the balance sheet.

The observation that investors value cash on hand more highly than cash on firms' balance sheets corresponds to using a discount factor β for firms' future distributions that is lower than the risk-free discount factor $\frac{1}{1+r}$, where r is the risk-free interest rate. The premium reflects uncertainty about managerial behavior and may also include a compensation for investors' risk aversion, if they cannot perfectly diversify their risk.

To capture the random arrival of investment opportunities, we assume that a Bernoulli variable $\tilde{\lambda}_t$ indicates every period whether a firm has an investment project in which it can invest. $\tilde{\lambda}_t$ takes on the value of 1 with probability p, which indicates that the firm can invest, and 0 with probability $(1-p)$, if it cannot invest.

If the firm has an investment opportunity, then investing I_t dollars in the project yields a certain pay-off of $F(I_t)$ at the end of period t, with $F(\cdot)$ being a neoclassical production function, e.g. of the form $F(I_t) = AI_t^\alpha$. For notational convenience we define the corresponding net production function, or net profits $G(I_t) = F(I_t) - (1+r)I_t$.

The firm's investment is limited by the amount of cash that it kept on its balance sheet

$$I_t \leq M_t - D_t \tag{1}$$

(1984), have explained why firms may face credit and equity constraints. Our analysis does not depend on the particular explanation for the constraints, only on the fact that such constraints exist.

If it does not have an investment opportunity, then it keeps the amount of $M_t - D_t$ on its bank account and earns interest at rate r on it.[10] And thus

$$\tilde{M}_{t+1} = (1+r)\left[M_t - D_t\right] + \tilde{\lambda}_t G\left(I_t\right) \tag{2}$$

This is the basic law of motion of our model. Investment is limited to cash on hand as given in equation (1). We we will show that normally, the constraint is binding. This means that the key decision variable is D_t. Given the structure of our model, D_t depends simply on M_t, i.e. it is a function $D(M_t)$. Technically, this paper analyzes the optimal dividend policy of the firm, $D^*(M_t)$, and, in particular, how it is affected by dividend taxation.

Before proceeding to the analytical part of the paper, let us note that our assumptions regarding investment opportunities, access to capital markets, and incentives for dividend payments have been made only for analytical simplicity. There are the following three dimensions along which the presented model could be generalized without affecting the qualitative predictions of the paper:

Firstly, instead of a binary Bernoulli variable determining an investment opportunity, we could allow for opportunities of different magnitudes, e.g. that firms face a production function of the form $\tilde{\eta}_t F(I_t)$, where $\tilde{\eta}_t \in [0, \infty)$ is a random variable. Alternatively, we could allow firms to carry unused opportunities into future periods at some cost. The assumption that postponing investment is costly corresponds to the real world, since firms are subject to e.g. changing competition, including competition from abroad, and evolving consumer tastes. As long as there is at least some cost associated with not fully seizing an investment opportunity in the period in which it is optimal to do so, then changes in firms' cash balances induced by changes in dividend taxation can have real effects.[11]

Secondly, the extreme case of having no access to capital markets once the Bernoulli variable, which indicates whether the firm has an investment opportunity, has materialized could easily be relaxed. We could for example assume that firms maintain a given debt equity ratio, and that this is already built into the production function $F(\cdot)$ given above. Alternatively, if firms changed their debt equity ratio in response to tax changes, the discussed results could be attenuated, but in general they will not

[10] In our model, we assume that the rate of interest is fixed, e.g. as would be the case in a small open economy. Qualitative results would be similar if the real rate of interest fell as firms in the aggregate invested more; the magnitude of the change in output would, of course, be smaller; conversely for a reduction in investment.

When firms engage in intertemporal arbitrage by paying out money to households, some of the money paid out may be invested elsewhere. Empirically, there is some question about the extent to which this may be so. For instance, dividends received by financial intermediaries may be held in cash (if they believe that it may be better to invest in equities at some later date); or it may be invested in equities, driving up the price of equities – but the evidence suggests that the response of investment to these increases in equity prices is limited, slow, and uncertain. For those dividends going directly to households, the marginal propensity to invest (directly or indirectly) is likely to be markedly lower than had the money remained inside the firm.

Furthermore, it is important to note that the intertemporal arbitrage can exacerbate capital market inefficiencies, i.e. the marginal return inside the firm may be increased relative to investment opportunities elsewhere in the economy. And this will be true regardless of whether interest rates are fixed and flexible.

[11] Obviously, the magnitude of the welfare loss is smaller the smaller the diminution in value as a result of postponement.

disappear. All that is required for our results is that at least in the short run, external finance is an imperfect substitute for internal finance, e.g. because there are costs to accessing capital markets. This is certainly true in the real world.[12]

Thirdly, regarding firms' payout policy, the assumption that shareholders discount future dividend payments at a higher rate than the risk-free interest rate can be replaced by any other theory of why firms make dividend payments, so long as intertemporal changes in tax rates affect firms' marginal incentive to pay out dividends.[13] An example for a theory of dividend payments other than agency costs is e.g. the accumulated retained earnings tax, which punishes firms for holdings excessive cash balances. As long as future changes in tax rates lead firms to alter their payout behavior, firms' cash balances and thus their investment and output will be affected.

We begin our analysis by looking at firms' payout policy under a constant dividend tax rate. We will show that the behavior of firms can be analyzed by looking at three stages: (a) new firms, for which $M_t = 0$; (b) young firms: there exists a value of $M_t = M^*$, such that for $M_t \leq M^*$, $D_t = 0$; and (c) mature firms, which distribute all excess cash, i.e. $D_t = M_t - M^*$. M^* can be derived as the optimal amount of cash balances that firms hold in steady state.

2.2 Model Setup

The representative firm in our model maximizes its stock market value, i.e. the discounted stream of dividend payments to its shareholders, who are subject to a dividend tax rate of $\tau \in [0, 1)$:

$$V(M_0) = \max_{\{D_t, I_t, M_{t+1}\}_{t=0}^{\infty}} E\left\{\sum_{t=0}^{\infty} \beta^t (1-\tau) D_t\right\} \tag{3}$$

$$\text{s.t. } M_{t+1} = (1+r)\left[M_t - D_t\right] + \tilde{\lambda}_t G(I_t) \tag{4}$$

$$M_t \geq I_t + D_t \tag{5}$$

$$D_t \geq 0 \tag{6}$$

$$\text{with } M_0 > 0 \text{ given.}$$

Constraint (4) describes the law of motion for the firm's cash holdings; constraint (5) captures the fact that dividend payments plus investment cannot exceed the firm's cash holdings. Finally, (6) is the dividend non-negativity constraint.[14]

In recursive formulation, the maximization of (3) is equivalent to the following problem, where we have substituted the law of motion for cash holdings (4) into the

[12] Even if we allowed firms to access equity markets again in case a very large investment opportunity $\tilde{\eta}_t$ arrives, firms' cost of equity would be $1+\kappa$ as opposed to $\frac{1}{1-\tau}$ for the cost of retained earnings, and so there would be a real cost involved with being capital constrained.

[13] Normally, this will be the case, so long as firms are concerned about the present discounted value of dividend taxes. Our analysis is, in particular, consistent with the signalling theory of dividends, if we allow that investors consider special dividends when tax arbitrage opportunities arise differently from regular dividend payments.

[14] This is needed since a firm that is short of cash would otherwise want to issue new equity through a negative dividend payment, but in the algebraic representation above this would cost shareholders only $(1-\tau)D_t$, i.e. they would receive a subsidy of τD_t. This clearly needs to be ruled out. We discuss the dynamics of raising equity in the next section.

problem:

$$V(M_t) = \max_{D_t, I_t} (1-\tau) D_t + \beta E V \left\{ (1+r) [M_t - D_t] + \tilde{\lambda}_t G(I_t) \right\} \quad (7)$$
$$\text{s.t. } (5), (6)$$

We can write the Lagrangian of this maximization problem as:

$$\mathcal{L}_t = (1-\tau) D_t + \beta E V' \left\{ (1+r) [M_t - D_t] + \tilde{\lambda}_t G(I_t) \right\} + + \mu_t [M_t - D_t - I_t] + \xi_t D_t \quad (8)$$

In this formulation, μ_t is the shadow value of cash M_t on the firm's balance sheet. ξ_t is the shadow value of the dividend non-negativity constraint. The envelope condition plus the resulting set of first-order conditions for this maximization problem are:

$$\text{EV: } V'(M_t) = \beta E V' \left(\tilde{M}_{t+1} \right) (1+r) + \mu_t \quad (9)$$

$$\text{FOC}(D_t)\text{: } 1 - \tau = \beta E V' \left(\tilde{M}_{t+1} \right) (1+r) + \mu_t - \xi_t \quad (10)$$

$$\text{FOC}(I_t)\text{: } \mu_t = \beta E \left\{ V' \left(\tilde{M}_{t+1} \right) \cdot \tilde{\lambda}_t G'(I_t) \right\} \quad (11)$$

$$\text{K.T.: } \mu_t [M_t - D_t - I_t] = 0 \quad (12)$$

$$\xi_t D_t = 0 \quad (13)$$

The envelope condition (9) expresses that in equilibrium, the marginal value of adding one dollar to the firm's balance sheet today equals the discounted marginal increase in firm value tomorrow plus the shadow value μ_t of holding the additional dollar on the balance sheet today. First order condition (10) states that the value to the shareholder of paying out one more dollar in dividends, $1 - \tau$, has to equal the marginal value of keeping the dollar in the firm, which consists of the expected increase in firm value next period plus the shadow value of holding the dollar on the balance sheet this period minus the shadow price of the dividend non-negativity constraint.

Equation (11) describes the optimal point of investment as the one where the marginal shadow price of holding cash on the balance sheet, μ_t, equals in expectation the increase in firm value next period, which is due to random investment opportunities. Note that combining equations (9) and (10) yields the following condition for the slope of the value function:

$$V'(M_t) = 1 - \tau + \xi_t \quad (14)$$

A direct result of this is the subsequent lemma:

Lemma 1 *The value function $V(M_t)$ is strictly increasing.*

Proof. The proof of this statement follows directly from (14) in conjunction with the observations that (a) dividend taxes are less than 1 and (b) the Lagrange multiplier ξ_t is by definition non-negative. ■

The Cash Constraint

If the firm was unconstrained in its access to capital, then it would chose to invest up to the point where $G'(I_t) = 0$ or $F'(I_t) = 1 + r$, i.e. where an additional dollar of investment yields no additional net return.

On the other hand, if $F'(I_t) > 1 + r$, then additional cash could be profitably invested if an investment opportunity arises, i.e. the given level of investment is constrained. Since the cash constraint (5) requires that $I_t \leq M_t$ and $F(\cdot)$ is convex[15], investment is always constrained when $F'(M_t) > 1 + r$.

Lemma 2 *The firm's level of investment is unconstrained, i.e.* $G'(I_t) = 0$ *if and only if the Lagrange multiplier* $\mu_t = 0$*, i.e. if holding an additional dollar of cash on the balance sheet is of no value, since it cannot be profitably invested.*

Proof. Resolving the expectations term in first-order condition (11) yields

$$\mu_t = p\beta V'\Big((1 + r)\left[M_t - D_t\right] + \ \ G(I_t)\Big) \cdot G'\left(I_t\right)$$

Lemma 1 that $V(\cdot) > 0$ then implies the stated result. ■

The Dividend-Nonnegativity Constraint

The Lagrange multiplier ξ_t on the firm's dividend nonnegativity constraint can be either positive or zero. If $\xi_t > 0$, then $D_t = 0$ and the firm would prefer to issue new equity rather than paying dividends. This situation arises when the firm's cash reserves M_t are low.

On the other hand, if $\xi_t = 0$, then the firm pays a non-negative amount of dividends $D_t \geq 0$. This corresponds to the case when the firm's cash holdings M_t are high. There is a cut-off value of cash holdings M^*, at which the dividend non-negativity constraint is marginally non-binding.

3 Steady State Analysis

Once the firm has reached M^*, its cash holdings will never fall below M^* again, given constant parameters values. To see this, note that the law of motion (4) implies that the only reason for cash reserves to fall is dividend payments. The marginal value of dividends to the investor is $1-\tau$, but equation (14) implies that the loss in the market value from paying dividends when $M_t < M^*$ is greater than $1 - \tau$, since $\xi_t > 0$. Hence no firm pays out more than $M_{t+1} - M^*$ in dividends in steady state. This allows us to conclude that $\xi_t = 0$ implies $\xi_{t+1} = 0$.

Using this result, we can substitute from equation (14) into the envelope condition to obtain an expression for the steady state shadow value of cash:

$$\mu_t = \mu^* = (1 - \tau)\left[1 - \beta(1 + r)\right] \tag{15}$$

Note that this equation confirms that dividend taxation reduces the shadow value of cash holdings proportionately by a factor of $1 - \tau$.

[15] The convexity of $F(\cdot)$ implies that $F'(I_t) \geq F'(M_t) > 1 + r$ for any level of investment that satisfies $I_t \leq M_t$.

According to lemma 2, the firm will operate at an unconstrained level of investment if $\mu^* = 0$ and at a constrained level if $\mu^* > 0$. We can see from the equation above that $\mu^* = 0$ holds if and only if the discount rate β applied to the firm's dividends equals the risk-free discount rate, or $\beta(1+r) = 1$.

Let us now turn to the determination of the amount I^{SS} out of the firm's cash holdings that is reserved for investment.

Proposition 1 *In steady state a firm allocates an amount of I^{SS} for investment purposes, which is defined by a marginal product of:*

$$G'(I^{SS}) = \frac{1-\beta(1+r)}{\beta p} \tag{16}$$

Proof. As discussed above, in steady state $\tilde{M}_{t+1} \geq M^*$, and this implies that $V'(\tilde{M}_{t+1}) = 1-\tau$. We can substitute this into the first-order condition for investment (11) and resolve the expectations term to obtain:

$$\mu_t = \beta p(1-\tau)G'(I_t) \tag{17}$$

After replacing $\mu_t = \mu^*$ from equation (15), we can solve the equation for $G'(I^{SS})$ to obtain the stated result. ■

Note that the dividend taxation term $(1-\tau)$ in (17) cancels out when we substitute for μ^*: Dividend taxation reduces the firm's shadow cost of cash μ^* and the expected marginal return to investment, i.e. the right-hand side of (17), by the same amount. This is our result on the neutrality of dividend taxation for firms in steady state:

Corollary 1 *The level of dividend taxation is irrelevant for the investment decisions of a firm in steady state.*

Using $F' = G' + 1 + r$ we can reformulate equation (16) as:

$$\beta\left[pF'(I^{SS}) + (1-p)(1+r)\right] = 1$$

In this form, the equation states that the expected marginal product of holding cash (which is F' if an investment opportunity arises or $1+r$ if it doesn't) discounted by β equals one.

Let us define I^* as the optimal level of investment under perfect capital markets, i.e. such that $F'(I^*) = 1+r$ or equivalently $G'(I^*) = 0$. Then we can see that equation (16) directly implies another important result:

Proposition 2 *A firm in steady state is constrained, i.e. $I^{SS} < I^*$ if and only if the discount factor β applied to the firm's stream of dividends is lower than the risk-free discount factor $\frac{1}{1+r}$.*

If $\beta = \frac{1}{1+r}$, the investor values cash in her pockets and cash on a firm's balance sheet equally. However, if $\beta < \frac{1}{1+r}$, then the investor discounts the firm's future cash flows – because of agency factors – at a higher rate than the return that the firm can earn on cash in its bank account. The firm thus has incentive to pay out more, and is cash-constrained when an investment opportunity arrives, so that $I^{SS} < I^*$.

Throughout the rest of the paper we limit our attention to the case where agency costs exist so that $\beta(1+r) < 1$. We argued in the introduction that agency concerns lead firms to pay out cash to reduce the amount of "free cash" that managers might abuse, and thus firms are in equilibrium cash constrained. Inducing them to reduce their cash holdings further, e.g. because of intertemporal tax arbitrage, has a very high cost.

There may be some firms, however, where agency problems are so severe that managers can avoid this "discipline." This assumes, of course, that take-over mechanisms do not work well – for reasons suggested by e.g. Grossman and Hart (1980) and Edlin and Stiglitz (1995). In that case firms may have excess cash from the perspective of economic efficiency: managers derive utility from the extra discretion that a reservoir of cash provides. It is still the case that the temporary reduction in dividend tax reduces investment, but in circumstances where investment was excessive this may be welfare increasing.[16]

The 2001 temporary dividend tax cut was introduced to stimulate the economy, in the belief that the aggregate level of investment was too low. Thus, the argument that it would have been a good thing for investment to be reduced runs at least counter to the professed argument for the tax cut.

This discussion highlights a more general point: different firms are in markedly different situations, and are therefore likely to be affected differently. Whether this increases or decreases the overall efficiency of the economy is a difficult and complex subject. For instance, among the more dynamic sectors of the economy are the "new economy" sectors, where any payouts are largely in forms that are subjected to capital gains taxation. (These "smart" sectors are not only smart in innovation; they have also figured out the dividend paradox.) Capital market imperfections (absence of good risk markets, etc.) mean that there may be underinvestment in this part of the economy, relative to the "old" and mature dividend paying firms. This is true without agency problems, but even more so with agency problems. Lowering the taxation of dividends advantages this old and mature economy, as opposed to firms in the new economy, and thus may exacerbate this inefficiency.[17] Our paper does not attempt to address these more general points; but by emphasizing that tax policy will affect firms that are in different stages of their life-cycle, it highlights the need for more detailed micro-economic research into the impacts of tax changes.

3.1 Payout Policy

The assumption $\beta(1+r) < 1$ entails that the Lagrange multiplier μ_t is positive, and hence the associated cash constraint (5) must hold with equality. Since the amount of cash with which the firm enters a period, M_t, is given, and the amount it keeps

[16]Some of the proponents of permanent dividend tax cuts were concerned with the inefficiency of capital allocation. They worried that the dividend tax induced firms to retain excess earnings, and this led to excessive investment. They thus seemed to side with the agency theory just discussed. But it is hard to know how seriously to take this rhetoric; as Stiglitz (2003) points out, there were ways of eliminating the problem of double taxation and excessive inside build-up while at the same time retaining the progressivity of the tax system. Given that argument, the dividend tax cut should have been limited to those firms that were in fact paying corporate income taxes.

[17]Indeed, this was one of the criticisms of the temporary dividend tax cut that was raised in the political debate

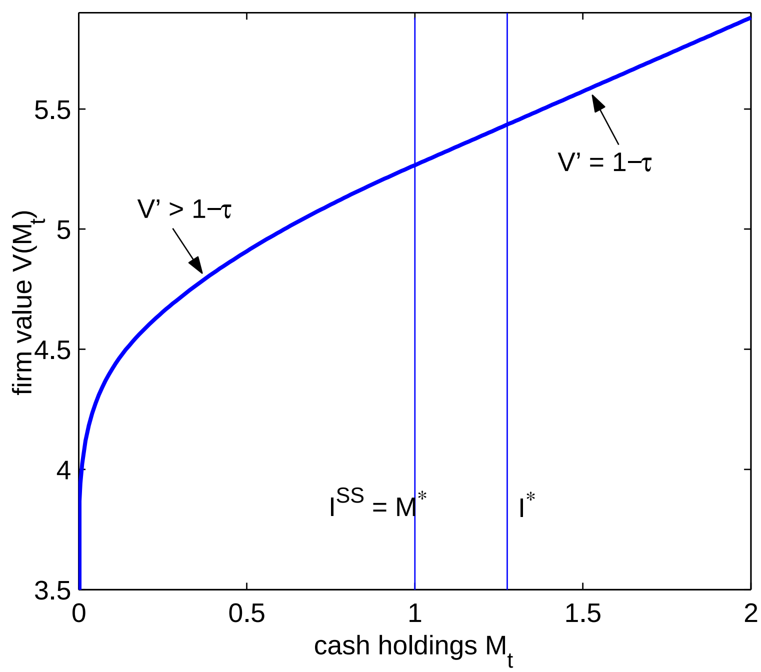

Figure 1: The value function $V(M_t)$ can be determined iteratively from firms' maximization problem (7). Parameter values: $\alpha = 1/2$, $\beta = 0.93$, $r = 0.01$, $p = 1/2$, $\tau = 38.6\%$, A calibrated so that $I^{SS} = 1$.

aside for investment purposes, I^{SS}, is defined by equation (16), the optimal level of dividends paid is:[18]

$$D_t = M_t - I^{SS} \tag{18}$$

Every period, the firm pays out the amount of cash that it holds in excess of its steady state amount of investment I^{SS}. If an investment opportunity arises, then it invests I^{SS} and earns $G(I^{SS})$. It carries this amount forward, and at the beginning of the next period $t+1$ it pays dividends equal to $G(I^{SS}) - I^{SS}$. Otherwise, if no investment opportunity arises, it earns interest rI^{SS}. It pays out its interest earnings at the beginning of period $t+1$ in order to keep its cash holdings equal to its steady state level of investment I^{SS}.[19]

Now we can conclude that $M^* = I^{SS}$ as defined above in (16). A firm pays out positive dividends whenever $M_t > M^* = I^{SS}$ and uses all of its remaining cash holdings for investment when an investment opportunity arises. Also, in steady state $\xi_t = 0$, as we argued above.

Using this insight and integrating equation (14) yields the following expression for the value function in the domain of $M_t \geq M^*$:

$$V(M_t) = V(M^*) + \int_{M^*}^{M_t} (1-\tau)dM = V(M^*) + (1-\tau)\left[M_t - M^*\right] \tag{19}$$

[18]Note that for $\beta(1+r) = 1$, the firm's cash holdings are indeterminate as long as $M_t \geq I^*$ – the firm is not penalized for holding larger cash holdings than its unconstrained optimal amount of investment I^*.

[19]This yields the result that dividends are variable. In practice, dividends often show a high degree of stability related to their signaling role, as discussed e.g. in Ross (1977) or Bhattacharya (1979) and evidenced in DeAngelo and DeAngelo (1990). Introducing this would complicate the analysis without altering the basic results of our paper.

In other words, above the threshold of M^*, the firm's value increases linearly in its cash holdings at a rate of one minus the dividend tax rate, as shown in figure 1.

We can determine $V(M^*)$ from the definition of the value function:

$$V(M^*) = \beta p V\left(G(M^*)\right) + \beta(1-p)V\left((1+r)M^*\right)$$

Using equation (19) to solve for $V(M^*)$ yields

$$V(M^*) = (1-\tau)\frac{\beta}{1-\beta}\left\{pG(M^*) + (1-p)rM^*\right\} \tag{20}$$

It can be easily verified that $V'(M^*) = 1-\tau$, as predicted by equation (14) for a firm in steady state.

4 Dynamics of Young Firms

When a firm starts out with an amount of cash below M^*, the dividend non-negativity constraint (6) is binding, and hence $\xi_t > 0$. (The optimal policy for the firm would be to raise new capital through negative dividend payments, which we have ruled out.)

$V'(\cdot)$ increases according to equation (14), and this in conjunction with the higher marginal product of investment $G'(I_t)$ raises the shadow value of cash holdings μ_t as described by first order condition (11).

Proposition 3 *If the firm's cash holdings are less than M^*, then it (i) does not pay any dividends and accumulates interest and profits until it reaches the steady state M^*. (ii) Along the way, it invests all its available cash whenever an investment opportunity arises, i.e. $I_t = M_t$.*

Proof. For claim (i) observe that $\mu_t > \mu^*$, and so first order condition (10) requires that $\xi_t > 0$. Consequently, no dividends are paid and all cash is accumulated as detailed in constraint (4).

Claim (ii) states that firms do not waste cash. Since $G'(M_t) > G'(M^*) \geq 0$ for $M_t < M^*$, it would not be optimal to invest less than the total available amount of cash. Analytically, note that $\mu_t > 0$ from first order condition (11) implies that the cash constraint (5) has to hold with equality, and since $D_t = 0$ it follows that $I_t = M_t$. ■

The described process continues until cash holdings attain or exceed the level M^*, at which point the firm reaches the steady state and starts paying dividends.

The value function $V(M_t)$ for $M_t < M^*$ does not have an explicit algebraic representation. However, figure 1 shows the relationship for typical parameter values. Figure 2 depicts the slope $V'(M_t)$ of the function.

Before we proceed, let us characterize the value function in more detail:

Lemma 3 *The value function is concave. The slope of the value function $V'(M_t)$ is (i) strictly decreasing and larger than $1-\tau$ over the interval $(0, M^*)$ and (ii) constant and equal to $1-\tau$ over the interval $[M^*, \infty)$.*

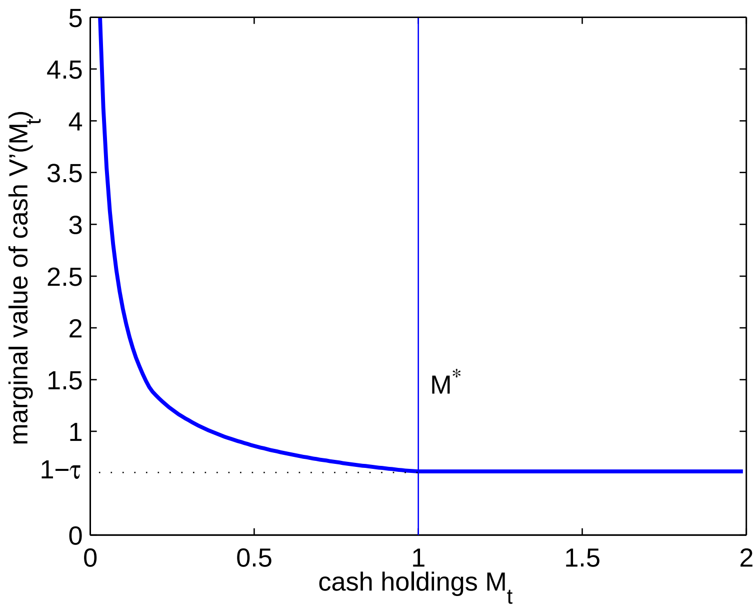

Figure 2: The derivative of the value function $V'(M_t)$ is higher than $1-\tau$ below M^* and equals $1-\tau$ thereafter. Same parameter values as in figure 1.

Proof. (i) The fact that $V'(M_t) > 1 - \tau$ for $M_t < M^*$ follows directly from (14). For the rest of the argument, we substitute the first order condition for investment (11) into the envelope condition (9) in order to eliminate μ_t. This yields the following expression for $V'(M_t)$, and after differentiating for $V''(M_t)$:

$$\begin{aligned} V'(M_t) =& \beta p V'\left(G(M_t)\right)\left[F'(M_t)\right] + \beta(1-p)V'\left((1+r)M_t\right)(1+r) \\ V''(M_t) =& \beta p V''\left(G(M_t)\right) G'(M_t) F'(M_t) + \\ & + \beta p V'\left(G(M_t)\right) F''(M_t) + \beta(1-p) V''\left(M_t(1+r)\right)(1+r)^2 \end{aligned}$$

Let us first focus on $M_t \in \mathcal{I}_1 = (\frac{M^*}{1+r}, M^*)$: We know that for this interval, all the $V''(\tilde{M}_{t+1})$-terms appearing in the equation above are zero since $\tilde{M}_{t+1} > M^*$. Consequently, the first and the last addend in the equation disappear. Since $G''(M_t) < 0$ for all neo-classical production functions, it follows that $V''(M_t)$ is negative over the observed interval.

Next, we extend the argument from interval $\mathcal{I}_k = (\frac{M^*}{(1+r)^k}, \frac{M^*}{1+r}^{k-1}]$ to $\mathcal{I}_{k+1} = (\frac{M^*}{(1+r)^{k+1}}, \frac{M^*}{(1+r)^k}]$. We know already that $V''(M_t)$ is non-positive for $M_t > \frac{M^*}{(1+r)^k}$. From $F'(M_t) > 0$, $G'(M_t) > 0$ and $G''(M_t) < 0$ we can conclude that the stated result holds for interval $\mathcal{I}_{k+1}$. Repeating this argument for $k \to \infty$ yields that the slope of the value function is strictly decreasing over the entire interval $(0, M^*)$.

(ii) Since the firm makes positive dividend payments for $M_t > M^*$, we know that the Lagrange multiplier ξ_t is zero. Hence equation (14) implies the stated result that $V'(M_t) = 1 - \tau$ for $M_t > M^*$. ■

Let us define $V(M_t; \tau)$ as the value of a firm with cash holdings M_t, given a constant dividend tax rate of τ. This allows us to state the following lemma:

Lemma 4 *The value function of a firm with cash holdings M_t under a constant dividend tax rate of τ can be expressed as $V(M_t;\tau) = (1-\tau)V(M_t;0)$, where $V(M_t;0)$ is the value function under no dividend taxation.*

Proof. For $M_t > M^*$ and $M_t = M^*$, the claim follows directly from equations (19) and (20). Since $V(M_t), M_t < M^*$ is derived iteratively from $V(M_t), M_t \geq M^*$, the same results have to apply. ■

4.1 Raising Equity

In this subsection we modify our analysis above to analyze the behavior of new firms, in particular to allow firms to issue equity.

Suppose that an entrepreneur develops a new business idea and starts a firm. The firm needs an initial investment of new equity N to start out. We assume that issuing new equity is costly. For every dollar that the firm receives in funding, the investor has to pay an additional premium of κ as a transactions cost. This can include e.g. underwriting fees (see Chen and Ritter, 2000) as well as underpricing because of agency problems (see e.g. Asquith and Mullins, 1986).

Proposition 4 *(i) For a new firm raising equity, the optimal amount of new equity $N^*(\tau)$ is such that $V'(N^*(\tau);\tau) = 1+\kappa$. (ii) $N^*(\tau)$ is decreasing in the dividend tax rate τ and in the cost of raising equity κ.*

Proof. (i) The proof follows from the convexity of $V(\cdot)$. Assume that investors have injected an amount of new equity N. Then adding one additional dollar of cash on the firm's balance sheet costs \$1 + κ, but would raise the value of the firm by $V'(N;\tau)$. If $N < N^*(\tau)$, then it would clearly be profitable to raise more cash, since $V'(N;\tau) > 1+\kappa$. On the other hand, if $N > N^*(\tau)$, then $V'(N;\tau) < 1+\kappa$ and it would be profitable to reduce the amount of equity raised.

(ii) Observe that $\frac{dV'(M_t;\tau)}{d\tau} < 0$. Since $V'(M_t;0)$ is strictly decreasing for $M_t < M^*$, the claim follows directly from lemma 4. ■

In the graph in figure 3, we have indicated the points where the slope of the value function reaches $1+\kappa$ and where it reaches 1, the latter being the case without transactions costs, which enables the entrepreneur to raise a higher amount of cash.

In short, the life-cycle of a firm in our model can be split into three distinct regions:

External financing region: For $M_t < N^*(\tau)$ the firm's marginal value of cash is higher than its external financing costs $1+\kappa$. The firm thus raises cash from equity markets up to the point $N^*(\tau)$.[20]

[20] In our analysis, firms only issue equity once. In practice, young firms often issue equity in several tranches, and firms sometimes borrow money, but return to equity markets sometime later in their development. The first problem is related to asymmetries of information: insiders may know (or believe) that information will become available that will reveal that the true value of the company is higher than currently believed; current owners thus want to minimize the dilution required by postponing raising funds from the market.
The second problem can be modeled as a simple extension of the model presented here. Randomly, firms may face a positive (publicly observable) shift in their production function (G); when that shift is large enough, it may pay the firm to raise new equity in the market (in spite of the high

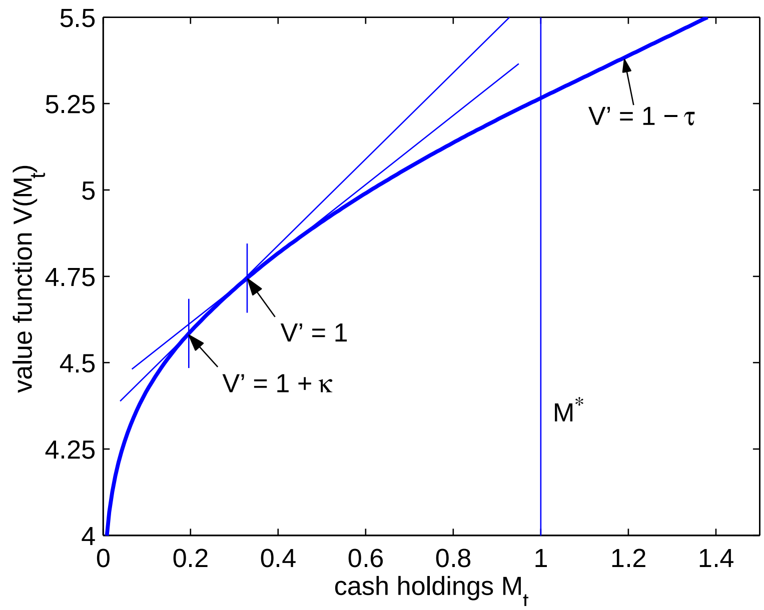

Figure 3: New equity issues: For a new firm, it is optimal to raise equity N^* up to the point where $V'(N^*(\tau);\tau) = 1 + \kappa$, or 1 if there are no costs to issuing equity. Parameter values used: $\kappa = 25\%$, $\tau = 36.8\%$.

Internal growth region: For $M_t \in [N^*(\tau), M^*)$ the firm's marginal value of funds is too low to raise new equity, but too high to pay out dividends. In this region, the firm grows internally by investing and retaining its earnings until it reaches the steady state with $M_t = M^*$. Also, the firm does not pay dividends in the internal growth region.

Steady state: For $M_t \geq M^*$ the firm is in its steady state. It regularly pays out its cash holdings in excess of the holdings M^* in the form of dividends.

Since the arrival of investment opportunities is random, the path along this life-cycle will differ from firm to firm. To give an example, we have depicted the average cash balances of a new firm over its first 15 periods in figure 4. As can be seen, almost all new firms in our example have reached the mature stage with $M_t = M^*$ after 10 periods under our choice of parameters.

Analytically, all firms have reached the steady state after

$$T = \left\lfloor \frac{\ln M^*/N^*(\tau)}{\ln(1+r)} \right\rfloor \tag{21}$$

periods, where we define the operator $\lfloor x \rfloor$ as the smallest integer above or equal to x. After T periods even the 'unlucky' firm that never had an investment opportunity reaches its steady state through the effect of compounded interest payments, i.e. $N^*(\tau)(1+r)^T \geq M^*$.

cost of doing so), because the opportunities forgone otherwise – while the firm builds up internally its working capital – is simply too great. (Recall our simplifying assumption that the firm cannot borrow. If the cost of borrowing is high – but not as large as the cost of equity – then, of course, the firm will resort to borrowing. If there is, for instance, a high cost of bankruptcy, then the firm may wish to limit its borrowing and so, even though it could borrow, will choose to raise new equity.)

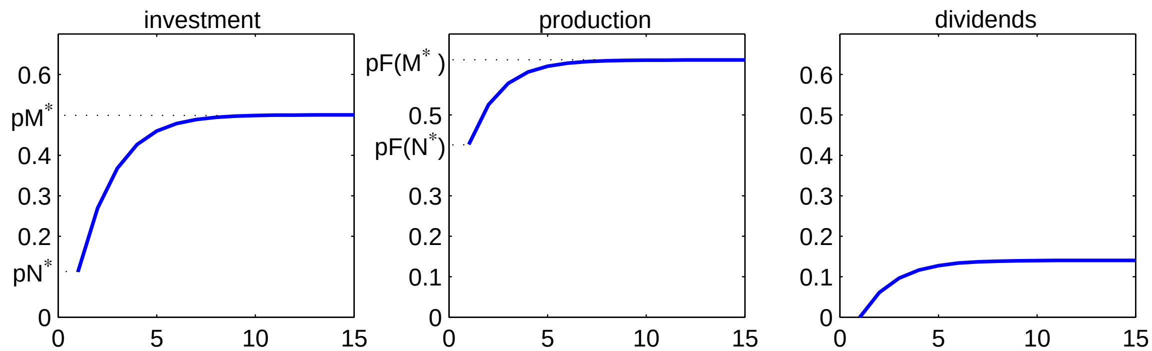

Figure 4: Average Evolution of Firm: The figure shows how investment, production, and dividend payments for a new firm evolve on average over the first 15 periods. In the first period, the firm raises $N^*(\tau)$ where we chose $\tau = 38.6\%$. It invests on average pN^* and produces $pF(N^*)$. In the 15th period, its expected investment and production are almost indistinguishable from the optimal values pM^* and $pF(M^*)$.

5 Changes in the Dividend Tax Rate

Let us now use the model that we developed in sections 2 to 4 to perform an analysis of how capital constrained firms react to changes in the rate of dividend taxation. As we discussed before, unanticipated tax changes are likely to have only a small impact on aggregate investment, since they affect only young firms. On the other hand, in the case of anticipated tax changes, firms engage in tax arbitrage, which affects their optimal level of cash balances and hence has potentially strong implications for aggregate investment.

5.1 Unanticipated Dividend Tax Changes

Unanticipated cuts and increases in the dividend tax rate affect the value function of firms in a similar fashion:

Proposition 5 *An unanticipated permanent dividend tax cut (increase) from τ_1 to τ_2 scales firms' value function $V(M_t; \tau_1)$ proportionally upwards (downwards) to $V(M_t; \tau_2)$. However, the optimal investment and payout policy of firms is not affected by an unanticipated change in dividend taxation.*

Proof. Firms' value function is defined in equation (3). It is easy to see that

$$V(M_t; \tau) = \max E\left\{\sum_{t=0}^{\infty} \beta^t (1-\tau) D_t\right\} = (1-\tau) \max E\left\{\sum_{t=0}^{\infty} \beta^t D_t\right\}$$

The factor $1-\tau$ entails just a linear transformation, and so it is easy to see that firms' optimal choice variables are independent of the dividend tax rate:

$$\underset{\{D_t, I_t, M_{t+1}\}}{\arg\max}\ E\left\{\sum_{t=0}^{\infty} \beta^t (1-\tau) D_t\right\} = \underset{\{D_t, I_t, M_{t+1}\}}{\arg\max}\ E\left\{\sum_{t=0}^{\infty} \beta^t D_t\right\}$$

Since firms' optimal investment and payout behavior is unaffected by the dividend tax, changes in dividend taxation simply scale their value function by the following factor:

$$V(M_t;\tau_2) = \frac{1-\tau_2}{1-\tau_1} V(M_t;\tau_1)$$

■

The amount of equity that new firms raise is also affected by tax cuts and increases. Furthermore, because of the asymmetries in raising and lowering equity (distributions to shareholders are diminished by dividend taxation, and there are costs to issuing new equity), there is an additional effect that arises only in the case of tax cuts:

Proposition 6 *An unanticipated permanent cut (rise) in dividend taxes from τ_1 to τ_2 increases (decreases) the amount of funds that new firms raise in equity markets from $N^*(\tau_1)$ to $N^*(\tau_2)$. Furthermore, in the case of a dividend tax cut, young firms with sufficiently low levels of cash holdings $M_t < N^*(\tau_2)$ access equity markets again and raise more cash.*

Proof. Proposition (4) entails that the amount of equity that new firm raise is decreasing in the dividend tax rate. In the case of a tax cut, there might be some young firms for which $M_t < N^*(\tau_2)$, i.e. for which one marginal dollar of new equity raises their value by more than the cost of issuing equity, i.e. $V'(M_t;\tau_2) > 1+\kappa$. These firms issue new equity up to the point where the marginal benefit and cost of new equity are equalized. ■

The reason for the asymmetry between tax cuts and tax increases is the following: After a tax cut, young firms with very low cash holdings access equity markets again, since the amount of equity N^* that firms raise increases. On the other hand, after a tax increase N^* falls, but firms do not pay out extra dividends, since M^* remains constant. This follows from the asymmetry in the tax treatment of equity issues, which carry no taxes, whereas dividend payments are taxed.[21]

For the investment decisions of mature firms a change in the dividend tax rate in either direction is neutral. This is typical for models that follow the 'new view' of dividend taxation.

5.2 Anticipated Dividend Tax Increase

Let us contrast these results to the case where an increase in dividend taxes is anticipated. In the following two propositions, we analyze the behavior of mature and young firms in the period immediately preceding an anticipated tax hike. After that, we discuss how firms' behavior in earlier periods is affected.

Throughout the discussion, we denote by $V_s(M_t;\tau_L,\tau_H)$ the value function of a firm that is subject to a dividend tax rate of τ_L for s periods and to a rate τ_H for all following periods, i.e. a firm that expects a dividend tax increase after s periods. Similarly, we denote by $M_s^*(\tau_L,\tau_H)$ the firm's optimal cash balances s periods before a tax increase from τ_L to τ_H. Note that, according to these definitions, $V_0(M_t;\tau_L,\tau_H) \equiv V(M_t;\tau_H)$ and $M_0^*(\tau_L,\tau_H) = M^*$.

[21] The marginal value of dividends to shareholders is $1-\tau$; the marginal cost to issuing equity is $1+\kappa$. Both the dividend tax and the cost involved in issuing equity thus increase the wedge between the cost of new equity and the cost of capital in the form of retained earnings.

Proposition 7 *An anticipated permanent increase in dividend taxes from τ_L to τ_H in period t induces mature firms to pay out a special dividend in the period $t-1$ preceding the tax increase. They reduce their cash holdings to $M_1^*(\tau_L, \tau_H) < M^*$, which is defined implicitly by the equation*

$$V'\left(M_1^*(\tau_L, \tau_H); \tau_H\right) = 1 - \tau_L$$

This leads to a suboptimal amount of investment and lower dividends in period t and the following periods, until firms have restored their optimal level of cash balances M^* through internal savings.

The intuition for this result is that firms can engage in tax arbitrage: shareholders can save taxes if firms pay out a higher amount of dividends in the period preceding a tax increase. However, this tax arbitrage is limited by the fact that firms lose from the sub-optimally low amount of cash balances in case an investment opportunity arises. The optimal amount of cash balances $M_1^*(\tau_L, \tau_H)$ balances these two factors off against each other.

Proof. Observe first that in the period preceding the increase, the value function for all firms paying positive amounts of dividends is:

$$\begin{aligned} V_1(M_{t-1}; \tau_L, \tau_H) = (1-\tau_L)D_{t-1} + \\ \beta E\left\{V\left((1+r)[M_{t-1} - D_{t-1}] + \tilde{\lambda}_{t-1}F(M_{t-1} - D_{t-1}); \tau_H\right)\right\} \end{aligned} \quad (22)$$

Maximizing this expression with respect to D_{t-1} yields the first order condition

$$1 - \tau_L + \beta \frac{dE\left\{V\left(\tilde{M}_t; \tau_H\right)\right\}}{dD_{t-1}} = 0$$

which can, by the definition of the value function $V(M_{t-1}; \tau_H)$, be re-written as

$$1 - \tau_L = \beta \frac{dE\left\{V\left(\tilde{M}_t; \tau_H\right)\right\}}{dM_{t-1}} = V'(M_{t-1}; \tau_H) \quad (23)$$

Since $V'(M_{t-1}; \tau_H)$ is decreasing, $M_1^*(\tau_L, \tau_H) < M^*$, i.e. firms' special dividend leaves them with a level of working capital that is lower than the steady-state optimum. ■

Graphically, firms' optimal cash balances $V_1(M_{t-1}; \tau_L, \tau_H)$ in the period prior to an anticipated dividend tax increase can be determined as the point where the slope of the value function after the tax increase, $V'(M_t; \tau_H) = 1 - \tau_L$, i.e. the investor's valuation of dividend payments under the old dividend tax rate. In figure 5 we have indicated this point by the vertical line labeled $M_1^*(\tau_L, \tau_H)$.

The firm's value function $V_1(M_{t-1}; \tau_L, \tau_H)$ in the period before the tax increase can thus be described as follows: For $M_{t-1} \leq M_1^*(\tau_L, \tau_H)$ it coincides with the value function $V(M_t; \tau_H)$ under the new dividend tax rate. For all $M_{t-1} > M_1^*(\tau_L, \tau_H)$ it increases at the constant rate $1 - \tau_L$, since dividends in this period are still taxed at the low rate τ_L.

Proposition 8 *In the period prior to an anticipated dividend tax increase to τ_H, new firms only raise the same amount of new equity as if the higher tax rate was already in effect, i.e. $N_1^*(\tau_L, \tau_H) = N^*(\tau_H)$.*

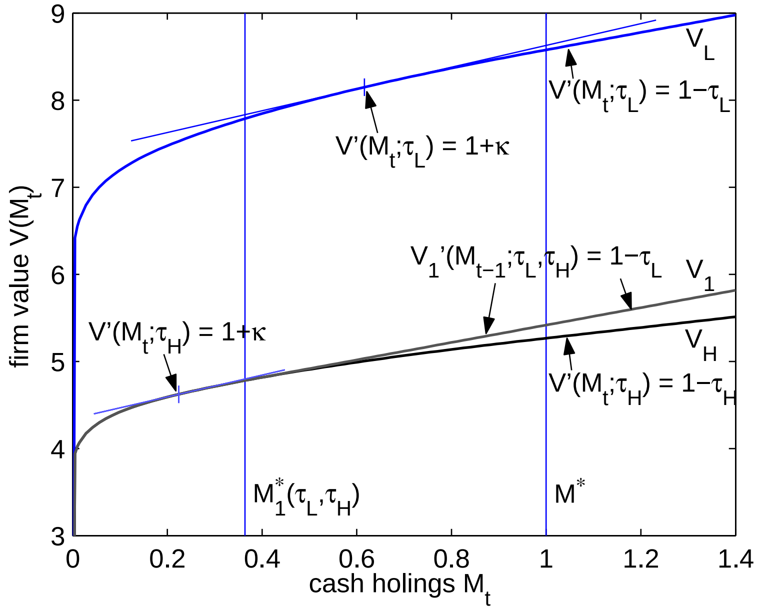

Figure 5: Anticipated Dividend Tax Increase, 1 Period Ahead: In the period before a dividend tax increase from $\tau_L = 0\%$ to $\tau_H = 38.6\%$, firms pay out all their cash holdings above $M_1^*(\tau_L, \tau_H)$ in the form of a special dividend. Their value function $V_1(M_t; \tau_L, \tau_H)$ coincides with the new value function V_H up to that point and increases at slope $1 - \tau_L$ (steeper than the new value function) thereafter.

Proof. The proof follows directly from our discussion above on firms' value function in the period prior to a tax increase. Since all dividends to investors that pay in new equity will be subject to the higher tax rate τ_H, the amount of new equity that firms can issue is $N^*(\tau_H)$. ■

One implication of these findings is that the value of young firms is more strongly impacted by an expected increase in dividend taxes than that of old firms – for firms with $M_t \leq M_1^*(\tau_L, \tau_H)$, the effect of a one-period-ahead expected increase in taxes is as strong as if the increase was already enacted. For firms with more cash, the value function lies above the value function under the new tax rate. In other words, expectations about tax increases hurt young firms more strongly than old firms. However, recall from the previous subsection that once the tax cut is enacted, all firms' value functions shift down proportionally as compared to the situation under the old tax rate.

Let us next investigate what happens in earlier periods. By definition, firms' value function in a period $t - s, s > 0$, given a dividend tax hike at time t, can be denoted iteratively as

$$V_s(M_{t-s}; \tau_L, \tau_H) = \max(1 - \tau_L) D_{t-s} + \beta E \left\{ V_{s-1} \left(\tilde{M}_{t-s+1}; \tau_L, \tau_H \right) \right\} \tag{24}$$

where $V_1(M_{t-1}; \tau_L, \tau_H)$ is defined above in (22).

We have depicted the value functions V_1 to V_{15} for the 15 periods preceding an anticipated dividend tax hike in figure 6. The upmost value function in the figure represents firms under the old tax rate τ_L, the lowest function represents firms under the new tax rate τ_H.

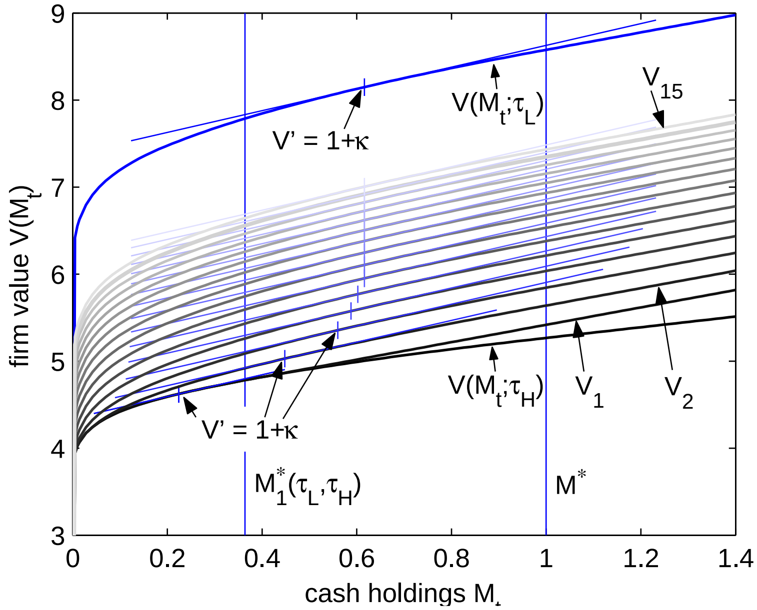

Figure 6: Anticipated Dividend Tax Increase, Value Functions: This figure shows the value functions $\{V_1, ..., V_{15}\}$ of firms in the 15 periods prior to an anticipated dividend tax increase from $\tau_L = 0\%$ to $\tau_H = 38.6\%$. When the future tax increase is announced, the value function falls immediately from $V(M_t; \tau_L)$ to V_{15}. Firms' value function then moves a small step closer to $V(M_t; \tau_H)$ every period.

If an increase in the dividend tax rate is announced s periods ahead, this implies that firms' value function immediately jumps down to $V_s(\tau_L, \tau_H)$. As can be seen in the figure, the series of value functions $V_k(\tau_L, \tau_H)$, $k = s, s-1, \ldots, 1$ moves every period a little closer to the final value function $V(M_t; \tau_H)$, which will prevail under the new dividend tax rate.

Firms' optimal amount of cash balances $M_s^*(\tau_L, \tau_H) = M^*$ for $s \geq 2$, since the tax rate between periods $t-s$ and $t-s+1$ remains unchanged and there is thus no tax arbitrage opportunity: Firms accumulate any cash balances below M^* and distribute amounts higher than this, paying a dividend tax rate of $1-\tau_L$. To see this point more formally, we maximize (24) with respect to D_{t-s}, which yields a first-order condition of

$$1 - \tau_L = \beta p V_s' \left(F(M_{s-1}^*) \right) F'(M_{s-1}^*) + \beta (1-p) V_s' \left((1+r) M_{s-1}^* \right) (1+r)$$

By the definition of M_s^* it follows that $V_s'(M_s^*) = 1 - \tau_L$. Furthermore, observe that $M_{s-1}^* \geq M_s^*$ for $s \geq 1$ when a dividend tax increase is anticipated, since the incentive to engage in tax arbitrage is strongest in the period directly preceding the tax increase. Since $F(M_{s-1}^*) > (1+r)M_{s-1}^* > M_{s-1}^*$ it follows that in the equation above, both $V_s'(\cdot)$-terms equal to $1 - \tau_L$. After canceling this term out of the equation, we obtain equation (16), which defined M^* in the absence of changes in dividend tax rates.

This implies that for $M_{t-s} \geq M^*$, the slope of the value functions is $V_s(M_t - s; \tau_L, \tau_H) = 1 - \tau_L$ for all $s \geq 2$. By lemma 4, $V'(M_t; \tau_H) = \frac{1-\tau_H}{1-\tau_L} V'(M_t; \tau_L)$, where $\frac{1-\tau_H}{1-\tau_L} < 1$. According to (24) and the ensuing discussion, the slopes of the value functions $V_s(\cdot)$ must lie in between $V(\cdot; \tau_H)$ and $V(\cdot; \tau_L)$. As a result, the amount of new equity $N_s^*(\tau_L, \tau_H)$ that new firms issue falls the smaller s, i.e. the less time is left under the low dividend tax. Furthermore, the proportional change in value for young firms is stronger than that for mature firms.

In figure 6, the amount of equity that new firms started in a given year s would issue is depicted as the tangents with slope $1+\kappa$ to the various value functions $V_s(\cdot)$. Clearly, $N_s^*(\tau_L, \tau_H)$ decreases the closer we come to the dividend tax increase.

5.3 Probabilistic Dividend Tax Increase

Firms react in a similar way if there is uncertainty about whether a dividend tax increase will occur in a future period. Such a situation can arise for example if a dividend tax cut has been implemented in the past and firms expect (e.g. for political economy reasons or because of large fiscal deficits) that the tax cut will be undone in a future period.

Proposition 9 *Suppose that in every time period, there is a probability π that the prevailing dividend tax rate τ_L will rise to τ_H next period, until the increase is finally realized. Then mature firms reduce their cash balances to $M_\pi^*(\tau_L, \tau_H) < M^*$ (or in short M_π^*) for as long as the low tax rate is in effect.*

Proof. We can denote the value function of firms in such a situation as

$$V_\pi(M_t; \tau_L, \tau_H) = (1-\tau_L)D_t + \beta\pi E\left\{V\left(\tilde{M}_{t+1}; \tau_H\right)\right\} + \beta(1-\pi)E\left\{V_\pi\left(\tilde{M}_{t+1}; \tau_L, \tau_H\right)\right\} \tag{25}$$

For $M_t \le M^*$ we can simplify $\beta E\left\{V\left(\tilde{M}_{t+1}; \tau_H\right)\right\} = V(M_t; \tau_H)$ in the equation above, since firms would not be willing to pay out dividends as long as they are below their steady state amount of cash. We use this observation in (25) above and maximize the expression with respect to D_t for firms that pay out dividends, i.e. for which $M_\pi^*(\tau_L, \tau_H) \le M_t \le M^*$, to obtain

$$1-\tau_L = \pi V'(M_\pi^*; \tau_H) + \beta(1-\pi)(1-\tau_L)\left[pF'\left(M_\pi^*\right) + (1-p)(1+r)\right]$$

This equation implicitly defines a unique value of $M_\pi^*(\tau_L, \tau_H)$, the optimal amount of cash when a dividend tax increase is expected with a constant probability of arrival of π. It is straightforward to show that for $\pi = 0$, $M_\pi^* = M^*$ and, using the implicit function theorem, that

$$\frac{dM_\pi^*(\tau_L, \tau_H)}{d\pi} < 0 \qquad \text{and} \qquad \frac{dM_\pi^*(\tau_L, \tau_H)}{d\tau_H} < 0$$

■ A direct implication of the observation that $M_\pi^*(\tau_L, \tau_H) < M^*$ is the following:

Corollary 2 *If there is uncertainty about a future dividend tax increase, aggregate investment by mature firms can be increased by raising dividend taxes immediately.*

Perhaps even more striking: whether the government raises or leaves the tax rate (permanently) unchanged, tax revenues will increase. Once the uncertainty about a possible dividend tax increase is resolved (in either direction), firms increase their cash holdings back to M^*.

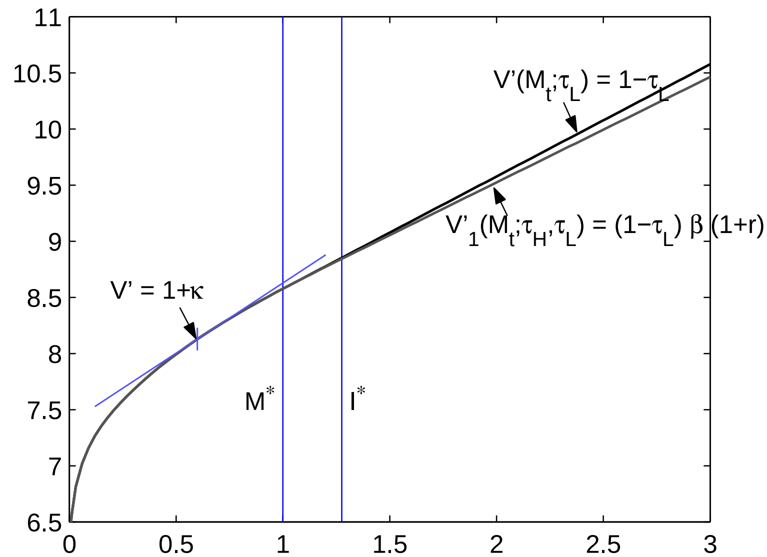

Figure 7: Dividend Tax Cut, Anticipated 1 Period Ahead: The value function $V_1(M_{t-1}; \tau_H, \tau_L)$ in the period preceding a large dividend tax cut coincides with the value function after the tax cut up to M^*, then it is concave between M^* and I^*. Above this threshold, it increases at the constant slope $(1-\tau_L)\beta(1+r)$.

5.4 Anticipated Dividend Tax Cut

In this subsection we investigate the case that a dividend tax cut from τ_H to τ_L is anticipated s periods ahead. As in the case of the anticipated dividend tax increase, we focus first on the period immediately preceding the cut, and then discuss the behavior of firms in earlier periods. The general idea is that an anticipated dividend tax cut allows firms to increase investors' net return by delaying dividend payments into periods where the tax rate is lower. This induces firms to accumulate higher cash balances than M^*, which in turn raises investment when an investment opportunity arises.

Proposition 10 *The value function $V_1(M_{t-1}; \tau_H, \tau_L)$ of firms in the period $t-1$ preceding a dividend tax cut coincides with the value function $V(M_t; \tau_L)$ after the tax cut for $M_{t-1} \leq M^*$. This implies that new firms issue the same amount of equity in the period preceding the tax cut as new firms after the tax cut.*

Proof. For $M_{t-1} \leq M^*$ firms do not pay dividends under constant dividend taxes, and their incentive to distribute cash under falling dividend taxes is even lower. For $M_{t-1} \leq M^*$, we can thus rewrite

$$V_1(M_{t-1}; \tau_H, \tau_L) = \beta E\left\{V\left(\tilde{M}_t; \tau_L\right)\right\} = V(M_{t-1}; \tau_L)$$

As discussed in proposition 4, this implies that new firms can raise an amount of equity of $N_1^*(\tau_H, \tau_L) = N^*(\tau_L)$ where $V'(N^*(\tau_L); \tau_L) = 1 - \tau^L$. We have illustrated this situation in figure 7. ■

Proposition 11 *For $M_{t-1} > M^*$ we have to distinguish two cases: (i) If the tax cut is large enough that $1 - \tau_H < (1-\tau_L)\beta(1+r)$, then $M_1^*(\tau_H, \tau_L) = \infty$, i.e. firms carry* *any cash holdings over into the next period and do not pay dividends. (ii) In the other case, firms keep a maximum amount $M_1^*(\tau_H, \tau_L) > M^*$ of cash on their balance sheets, which is defined by*

$$1 - \tau_H = (1 - \tau_L)\beta \left[pF'(M_1^*) + (1-p)(1+r)\right]$$

and return any excess cash to their shareholders immediately.

Proof. (i) If $1 - \tau_H < (1 - \tau_L)\beta(1 + r)$, a firm would never pay out dividends in period $t - 1$. It could raise its value even in the absence of investment opportunities by accumulating interest on its cash holdings. For $M_{t-1} \in [M^*, I^*]$, where I^* is the maximum amount of investment that does not yields a negative net marginal product, i.e. $G'(I^*) = 0$, firms can also earn additional cash by raising their investment up to $I_{t-1} = M_{t-1}$. In that interval, firms' value function is thus concave. Above I^* the value function increases at the constant slope $(1 - \tau_L)\beta(1 + r)$. Firms still retain all their cash holdings, but invest only I^* in case an investment opportunity arises. They earn regular interest on the remainder.

(ii) With $1 - \tau^H > (1 - \tau^L)\beta(1 + r)$, it would not pay firms to keep an arbitrarily large amount of cash holdings on their balance sheets, since at the discount rate β their future dividends are worth less than current dividends, despite of the somewhat lower future dividend tax rate. However, an increase in firms' cash balances allows them to raise their investment I_{t-1} in case an investment opportunity arises, and since $F'(M^*) > 1 + r$, they can profitably do so up to some point $M_1^*(\tau_H, \tau_L) > M^*$. This latter inequality must hold, since we derived M^* in the previous section under the assumption that holding cash for one period would return only $\beta(1+r) < 1$ in case no investment opportunity arises, whereas this return is effectively $\frac{1-\tau_L}{1-\tau_H}\beta(1+r) > \beta(1+r)$ now.

We can find the cut-off value $M_1^*(\tau_H, \tau_L)$ by evaluating the derivative in condition (23), which describes the optimal payout policy in the period preceding a dividend tax change:

$$1 - \tau_H = \beta \frac{dE\left\{V\left(\tilde{M}_t; \tau_L\right)\right\}}{dM_{t-1}} = \beta(1 - \tau_L)\left[pF'(M_1^*) + (1-p)(1+r)\right]$$

Solving this equation for the marginal product of investment yields

$$\beta\left[pF'(M_{t-1}^*) + (1-p)(1+r)\right] = \frac{1 - \tau_H}{1 - \tau_L}$$

All cash above this amount is paid out immediately at the prevailing dividend tax rate of τ_H, which also defines the slope of the value function above M_1^*. Note that M_1^* is always in the interval $[M^*, I^*]$, since its marginal product is between $F'(M^*)$ and $1 + r$. ■

Naturally, whether an economy is in case (i) or (ii) depends on the size of the dividend tax cut relative to the equity premium. Figure 7 illustrates the first case. Finally, when the tax cut comes into effect in period t, all firms with $M_t > M^*$ pay out their accumulated excess holdings of cash in the form of a special dividend $D_t = M_t - M^*$.

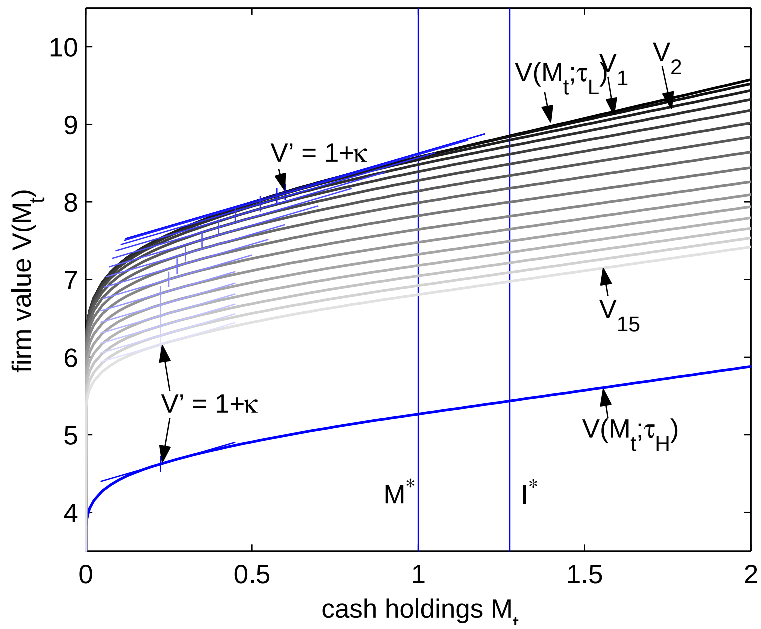

Figure 8: Anticipated Dividend Tax Cut, Value Functions: The value functions $\{V_1, ..., V_{15}\}$ of firms in the 15 periods prior to an anticipated dividend tax cut move closer and closer to the value function $V(M_t; \tau_L)$ prevailing under the new dividend tax rate. The amount of equity that new firms can issue, indicated by the tangent with slope $1 + \kappa$ to each value function, rises.

The value functions for periods $t - s$ more than one period from the anticipated dividend tax cut can be determined iteratively. For the 15 periods preceding a dividend tax cut, this is depicted in figure 8.

Note that, as for anticipated tax increases, the value of young firms (i.e. $V_1(M_t; \cdot)$ for low M_t) is affected relatively more strongly by anticipations of an impending dividend tax cut than that of mature firms. However, once the cut is enacted, all firms' value has increased proportionally compared to the situation before the tax cut was anticipated.

In analogy to the two cases in proposition (11), it can be shown that for k periods before a dividend tax cut firms do not pay out any dividends. This k can be determined as the largest integer such that keeping cash on the balance sheet for k periods and paying it out at the lower dividend tax rate is more favorable to the investor than paying it out immediately under the high dividend tax rate, i.e. $1-\tau_H < (1-\tau_L)[\beta(1+r)]^k$. More formally,

$$k = \left\lceil \frac{\log(1-\tau_H) - \log(1-\tau_L)}{\log[\beta(1+r)]} \right\rceil \tag{26}$$

where we define the operator $\lceil x \rceil$ as the largest integer below or equal to x. Consequently, $M_s^*(\tau_H, \tau_L) = \infty$ for $1 \le s \le k$ and firms invest $I_{t-s} = \min\{M_{t-s}, I^*\}$ along the way.

In the $(k+1)^{th}$ period before the tax cut, firms keep cash balances up to the optimal $M_{k+1}^*(\tau_H, \tau_L) > M^*$, which is defined by the following equation:

$$\beta \left[pF'(M_{k+1}^*) + (1-p)(1+r) \right] = \frac{1-\tau_H}{1-\tau_L} \cdot \frac{1}{[\beta(1+r)]^k}$$

If firms' cash holdings are higher than this M^*_{k+1} in period $t-k-1$, they pay out the excess cash $M_{t-k-1} - M^*_{k+1}$. No further dividend payment is made until after the tax cut, when all the accumulated excess cash holdings $M_t - M^*$ are distributed to shareholders.

If a dividend tax cut is anticipated more than $k+1$ periods into the future, the level of the value function of firms goes up immediately, but the slope of the value function for $M_{t-s} \geq M^*$ remains unchanged. By implication, the payout policy of existing firms is unaltered. The amount of equity that new firms issue rises from period to period before a dividend tax cut.

5.5 Probabilistic Dividend Tax Cut

If there is uncertainty about whether a dividend tax cut will occur in a future period, similar conclusions hold. An example for this situation occurred in 2002, when political parties were discussing the possibility of a reduction in the dividend tax rate.

Proposition 12 *If firms expect a dividend tax cut from τ_H to τ_L with a probability of π in every time period until the tax cut is realized, they accumulate cash balances up to a level of $M^*_\pi(\tau_H, \tau_L) > M^*$ while the higher tax rate is in effect, which is defined by*

$$\begin{aligned} \beta(1 - E\tau_{t+1})[pF'(M^*_\pi) + (1-p)(1+r)] = 1 - \tau_H \quad &\textit{if} \quad 1 - \tau_H > \beta(1+r)(1 - E\tau_{t+1}) \\ M^*_\pi = \infty \quad &\textit{if} \quad 1 - \tau_H \leq \beta(1+r)(1 - E\tau_{t+1}) \end{aligned}$$

where $E\tau_{t+1} = (1-\pi)\tau_H + \pi\tau_L$ is the expected dividend tax rate next period.

In other words, firms reduce their dividend payments and increase their cash holdings in order to take advantage of the expected lower dividend tax in the future. If the tax cut is likely or large enough, as described by the second inequality, firms will not pay any dividends until the tax cut has materialized. Instead, they would be willing to accumulate arbitrarily large cash balances.

Proof. The value function of firms under these circumstances is

$$V_\pi(M_t; \tau_H, \tau_L) = (1 - \tau_H)D_t + \beta\pi E\left\{V(\tilde{M}_{t+1}; \tau_L)\right\} + \beta(1-\pi)E\left\{V_\pi(\tilde{M}_{t+1}; \tau_H, \tau_L)\right\}$$

i.e. the sum of current dividend payments plus the value of future payments if the tax cut has materialized (with probability π) plus the value of future dividends if the tax cut has not materialized (with probability $1-\pi$). The first order condition to this problem defines the optimal amount of cash balances $M^*_\pi(\tau_H, \tau_L)$ for firms in such a situation:

$$\begin{aligned} 1 - \tau_H =& \beta\pi\left[pV'(F(M^*_\pi))F'(M^*_\pi) + (1-p)V'((1+r)M^*_\pi)(1+r)\right] + \\ &+ \beta(1-\pi)\left[pV'_\pi(F(M^*_\pi))F'(M^*_\pi) + (1-p)V'_\pi((1+r)M^*_\pi)(1+r)\right] \end{aligned}$$

Now observe that for the given parameter values, $V' = 1 - \tau_L$ and $V'_\pi = 1 - \tau_H$. This allows us to simplify the expression to

$$1 - \tau_H = \beta\left[pF'(M^*_\pi) + (1-p)(1+r)\right]\left[(1-\pi)(1-\tau_H) + \pi(1-\tau_L)\right] \qquad (27)$$

which can be rearranged to yield the first equation in the proposition. It states that the optimal level of cash balances is at the point where the marginal value of dividend payments now, i.e. the left-hand side term $1 - \tau_H$, equals the expected marginal value of carrying over cash to the following period, which is an average of the case that an investment opportunity comes up and that it doesn't times one minus the expected dividend tax rate.

However, if $1-\tau_H \leq \beta(1+r)(1-E\tau_{t+1})$, then it actually pays firms to accumulate any amount of cash balances, since in expectation shareholders are better off if all dividends are paid out once the tax cut has materialized, i.e. $M_\pi^* = \infty$. Again, when an investment opportunity arises, firms invest an amount $I_t = \min\{M_t, I^*\}$ and earn the regular rate of interest $1 + r$ on any cash holdings in excess of I^*. This is the situation captured in the second equation of the proposition. ■

Corollary 3 *If there is uncertainty about whether there will be a tax decrease, it pays the government to postpone resolving that uncertainty (either way). In the interim, tax revenues are higher and output is higher.*

5.6 Temporary Dividend Tax Changes

We can now use the analysis of the previous subsections to discuss the case of a temporary dividend tax change. Essentially, a temporary change in the level of dividend taxation that takes effect immediately is equivalent to an unanticipated change in one direction plus an anticipated change in the reverse direction at a later point in time.

Let us first focus on temporary dividend tax cuts. Assume that before time t, a dividend tax rate of τ_H prevails. At time t, the dividend tax rate is unexpectedly reduced to τ_L for k periods, after which it returns to τ_H. As discussed before, the unanticipated reduction in the tax rate increases the value of all firms. For new firms, this increases the amount of equity that they raise, but it does not have any effects on the behavior of mature firms. However, note that both the increase in stock prices and in the amount of equity that new firms issue is attenuated by the fact that a dividend tax increase in period $t + k$ is anticipated.

Figure 6 illustrates that the value functions for the periods preceding a tax hike from τ_L to τ_H are actually much closer to $V(M_t; \tau_H)$, i.e. the lowest line in the figure, than to $V(M_t; \tau_L)$, the highest line. In other words, the positive effects of lower dividend taxes in terms of higher availability of equity for new firms is strongly reduced by the fact that a dividend tax hike in a future period is impending. The reason is that young firms usually do not pay out dividends in the first few years of their existence, but only once they have accumulated their steady state holdings of cash M^*. We can thus conclude that temporary dividend tax cuts reduce the marginal cost of equity $V'(M_t)$ mostly for those firms that do not need it (mature firms), but hardly affect the marginal cost of equity for new firms that need to access capital markets.

In the extreme case of a temporary one-period tax cut, firms' value function becomes $V_1(M_t; \tau_L, \tau_H)$ in the figure. This implies that new firms cannot raise more equity than if the tax rate had been kept at τ_H. At the same time, firms pay out a special dividend to reduce their cash holdings to $M_1^*(\tau_L, \tau_H)$, and cash balances will be lower for a number of periods until firms have recovered their optimum, M^*. The effect of such a one-period temporary dividend tax cut on aggregate investment and output would thus be unambiguously negative.

The opposite conclusions hold for a temporary dividend tax increase. While the amount of equity that new firms issue during the high tax period would fall, firms hold higher cash balances than M^* in anticipation of the impending dividend tax cut, and this allows them to invest higher amounts. Again, the extreme case of an unanticipated one-period increase would have unambiguously positive effects on aggregate investment, though only for one period.

5.7 US Dividend Tax Policy in the Current Decade

In analyzing the effects of dividend tax policy on firms' payout policy and the resulting macroeconomic implications, it is important to take account of both implemented tax changes and expectations thereof.

In the US, it can be argued that rational, forward-looking firms already assigned a small but positive probability to a dividend tax cut in 2000, when the Bush campaign promoted tax cuts as way of reducing the high predicted fiscal surplus over the following decade. After Bush's election victory, that probability was revised upwards. Over the course of 2001 and 2002 the specifics of which taxes were to be cut became clearer, and the probability that firms assigned to a cut increased.[22] Finally, in May 2003, the Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) was enacted and all details of the temporary tax cut, including its retroactive start at the beginning of 2003 and its expiry in 2008, became, at that point, fully known. Of course, there remained uncertainty about whether the law would be repealed and taxes increased – uncertainty which plays an important role in the sequel.

Over the period leading up to the enactment of the JGTRRA, mature firms thus had an increasing incentive to postpone dividend payments to later periods with a lower expected tax load and thus to raise their cash balances above M^*. The lower tax rates also enabled new firms to issue more equity. This had the potential to raise investment among all categories of firms.

In 2003 the temporary tax cut came into effect, with an official expiration date of 2008. While many small-government conservatives argued that the tax cut should be made permanent, an increasing number of deficit-hawks called for an early repeal of the tax cuts, and so did Senator Kerry who ran against Bush in the presidential elections. During the election campaigns of 2004, firms must thus have assigned a positive probability to all three events – the scheduled expiration of the tax cut in 2008 as well as an extension or a repeal as early as 2005. This latter possibility made it optimal to pay out large dividends in 2004, as exemplified by Microsoft's $32bn special dividend.

When Bush won re-election in 2004, expectations adjusted again. An early repeal of the tax cut was no longer considered likely, but the scheduled expiry or an extension of the tax cut beyond 2008 were both possible alternatives. Accordingly, in a special budget reconciliation bill in May 2006, the dividend tax cuts were extended until the end of 2010. President Bush continued to call for making the dividend tax reductions permanent, where "permanent" has to be understood in the context of a political

[22] For a detailed overview of the relevant events see table 1 in Auerbach and Hassett (2007).

economy in which future governments are likely to change course again, as we will discuss below in section 7.

6 Aggregate Investment and Output

Let us now investigate the effects of dividend tax policy on aggregate investment.[23] Define $\bar{I}_a := pE[I_a]$ as the expected amount of investment by a firm of age a, i.e. a firm that was started a periods ago.[24] Evidently, this definition entails that $\bar{I}_0 = pN^*$ and $\bar{I}_a = pM^*$ for all $a \geq T$, where we defined T above in (21) as the number of years after which a firm has reached its steady state cash holdings even if it never had an investment opportunity.

Next, define $\bar{Y}_a := pE\{G(I_a)\}$ as the expected net output by a firm of age a. Again, $\bar{Y}_0 = pG(N^*)$ and $\bar{Y}_a = pG(M^*)$ for all $a \geq T$. But note that $\bar{Y}_a < G(\bar{I}_a)$ for $0 < a < T$, since the firm's cash balances are a random variable and Jensen's inequality applies. Similarly, we denote $\bar{D}_a := E[D_a]$ for a firm of age a, and we note that $\bar{D}_0 = 0$ and $\bar{D}_a = p\,[G(M^*) - M^*] + (1-p)rM^*$. These $\bar{I}_a$, $\bar{Y}_a$ and $\bar{D}_a$ are what we had depicted in figure 4.

Assume that at any given time t there exists a mass z_{t-1} of previously existing firms, and a mass $\Delta z_t = \gamma z_{t-1}$ of new firms is started, yielding a total mass of $z_t = (1+\gamma)z_{t-1}$. Firms are indexed by i and follow the maximization problem and the resulting rules described in the previous sections. The arrival of investment opportunities is independent among firms.[25]

In steady state the fraction of firms at a given stage of development, say age a, is then constant at $\frac{\gamma}{(1+\gamma)^{a+1}}$. However, the total mass of firms is increasing. At time t the total mass of firms of age a is $\frac{\gamma}{(1+\gamma)^{a+1}} z_t$.

Aggregate investment AI_t in the economy at time t can thus be expressed as

$$AI_t = \int_0^{z_t} pI_{t,z}dz = z_{t-T}\bar{I}_T + \Delta z_{t-T+1}\bar{I}_{T-1} + \cdots + \Delta z_t \bar{I}_0 =$$

$$= z_{t-T}\bar{I}_T + \sum_{a=0}^{T-1} \Delta z_{t-a}\bar{I}_a = z_t \left[\frac{\bar{I}_T}{(1+\gamma)^T} + \sum_{a=0}^{T-1} \frac{\gamma \bar{I}_a}{(1+\gamma)^{a+1}} \right]$$

By the same token, aggregate output AY_t and aggregate dividend payments AD_t are

$$AY_t = z_t \left[\frac{\bar{Y}_T}{(1+\gamma)^T} + \sum_{a=0}^{T-1} \frac{\gamma \bar{Y}_a}{(1+\gamma)^{a+1}} \right] \quad \text{and} \quad AD_t = z_t \left[\frac{\bar{D}_T}{(1+\gamma)^T} + \sum_{a=0}^{T-1} \frac{\gamma \bar{D}_a}{(1+\gamma)^{a+1}} \right]$$

In the remainder of this section we use this very simple model of the aggregate economy to simulate the impact of dividend tax policy on aggregate investment. Since dividend taxes are only a comparatively small part of government revenue, we disregard any effects of dividend taxation on the government's budget position. Following Mehra

[23] The following macroeconomic analysis ignores possible offsetting monetary policy actions.

[24] Note that in this expression, the amount of cash that firms have reserved for investment purposes is multiplied by the probability p that an investment opportunity arises and cash can actually be invested.

[25] By the law of large numbers, this assumption yields sure values for all macroeconomic variables, even though the underlying firm-level variables such as output and investment are random.

and Prescott (1985, 2003) we chose $\beta = 0.93$ and $r = 1\%$. In order to replicate a typical growth rate of the economy we calibrated $\gamma = 3\%$.[26] For the other structural parameters we used values of $\alpha = 1/2$ and $p = 1/2$, and A was chosen so that $M^* = 1$. However, our results are robust to alternative calibrations.

6.1 Aggregate Effects of Unanticipated Dividend Tax Changes

Proposition 13 *In response to unanticipated changes in dividend taxes, aggregate investment AI_t, aggregate output AY_t and aggregate dividend payments AD_t in the economy are lower the higher the tax rate.*

Proof. As discussed in proposition 4, $dN^*/d\tau < 0$. Since $\bar{I}_0 = N^*$, the claim is evident for firms of age zero. Now note that we can express $\bar{I}_1$ in the following way:

$$\bar{I}_1 = E[I_1] = p \min\{F(N^*), M^*\} + (1-p) \min\{(1+r)N^*, M^*\}$$

Again, since $dN^*/d\tau < 0$ it also holds that $d\bar{I}_1/d\tau \leq 0$. An analogous argument can be made for all $\bar{I}_t$, $1 < t \leq T$. Since AI_t is a weighted sum of these terms, it follows that $dAI_t/d\tau < 0$. The same argument applies to AY_t and AD_t. ■

However, we derived in section 3 that the steady state level of cash M^* in mature firms is unaffected by dividend taxes. Hence these firms' contribution to aggregate investment is also unaffected. Since the majority of firms in an economy are mature firms, we can thus expect the impact of dividend taxation on aggregate investment, output, and dividends to be rather small.

We depict an example of the unanticipated reduction of a $\tau_H = 38.6\%$ dividend tax to $\tau_L = 15\%$ in figure 9. The tax cut occurs unexpectedly in period 4 and immediately raises the amount of new equity N^* that new firms issue. Since the optimal amount of cash balances M^* for mature firms remains unchanged, the average firm in the economy has only slightly higher cash balances, as a result of the higher N^*. As a result, aggregate investment AI, production AY, and gross dividend payments AD increase very modestly. However, the distribution of firms' dividends changes significantly: shareholders net dividend receipts $(1-\tau)AD$ increase steeply; the government's revenue from dividend taxation τAD falls sharply. Thus the redistributory effect of changes in dividend taxation is an order of magnitudes stronger than any efficiency effects.[27]

In our simulation model, the described cut of a 38.6% dividend tax by more than half increased aggregate investment, output, and dividends only by 1.3%, 0.9% and 2% respectively. Recall in this context that due to various tax deductions, the estimated effective dividend tax rate in the US economy is much lower than the nominal rate.

6.2 Aggregate Effects of an Anticipated Tax Increase

As we discussed in the previous section, the results for an anticipated tax change differ substantially.

[26]Empirically, the value of new equity issued fluctuates wildly from year to year, but $\gamma = 3\%$ is a reasonable approximation to the average long-run amount of equity issuance.

[27]If the higher government revenue allows for higher government expenditure and if there is a multiplier associated with government expenditure, then the dividend tax cut would potentially have a strongly negative effect on output. However, we do not explore this issue further in the given paper.

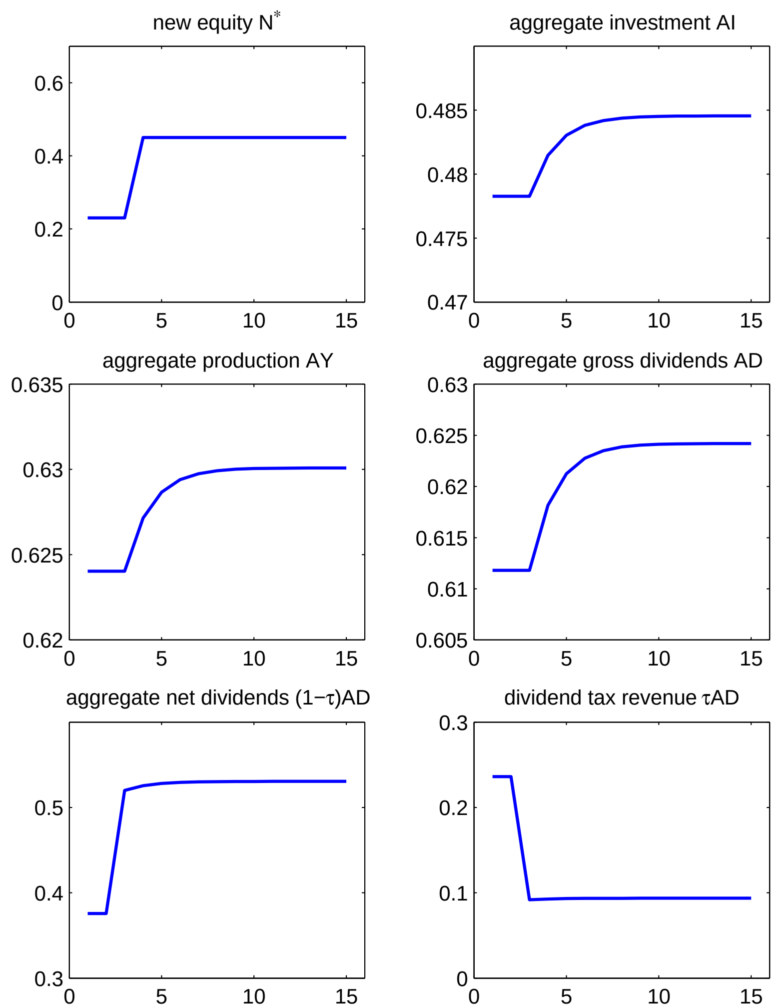

Figure 9: Unanticipated Dividend Tax Cut: After an unanticipated dividend tax cut from 38.6% to 15% in period 4, the amount of new equity N^* that new firms issue rises significantly. However, as can be seen from the scale of the Y-axes, aggregate investment AI, production AY, and gross dividend payments AD increase only very modestly. The redistributory effect of the tax change, on the other hand, is very strong. Note that we have normalized all aggregate variables in the figure by z_t in order to distinguish the effect of tax policy from the general growth of the economy.

We have depicted an example of a tax increase from $\tau_L = 15\%$ to $\tau_H = 38.6\%$ in figure 10. Starting from period 1, a tax increase is anticipated to occur in period 8. As shown in the figure, the effects of such an anticipated dividend tax increase on aggregate macroeconomic variables are twofold: Firstly, new firms issue less and less equity both in the periods leading up to the increase and under the new tax rate, which reduces their investment and prolongs the time that they need to reach the steady state. Secondly, all mature firms pay out a special dividend in the period immediately preceding the tax increase, which reduces their investment for periods to come. Since most firms in the economy are mature, this second effect is much more pronounced than the first one. However, as can be seen from the two graphs at the bottom, the redistributory effects of the tax increase are even stronger than the direct effects on investment and output.

Let us now turn again to the case where firms expect a dividend tax increase from τ_L to τ_H to happen any period with probability π. New firms issue less equity in such a situation than if the low dividend tax rate τ_L was expected to prevail indefinitely, but more than when the high tax rate τ_H comes into effect. However, the large part of the economy are mature firms. For these, investment is reduced when there is a tax increase impending, and thus aggregate investment is most likely depressed and will increase once the dividend tax increase has occurred. After that, mature firms retain all their earnings until they reach the new steady state cash holdings $M^* > M^*_\pi$.

An example of this is given in figure 11. In the first three periods of the figure, firms anticipate that there is a 25% risk of a dividend tax increase in each of the next periods. This reduces their optimal cash balances to M^*_π. In period 4, the tax increase materializes, and firms' optimal level of cash holdings reverts to $M^* > M^*_\pi$, which increases investment activity among mature firms. Note, however, that the amount of equity that new firms issue decreases in period 4 because of the higher dividend tax rate.

6.3 Aggregate Effects of an Anticipated Tax Cut

An example of firms' reaction to an anticipated dividend tax cut is given in figure 12: Agents foresee starting from period 1 that dividend taxes will fall from $\tau_H = 38.6\%$ to $\tau_L = 0\%$ in period $t = 11$. The amount of equity that new firms issue slowly increases between periods 1 and 11, as was also shown by the tangents in figure 8.

To determine the behavior of mature firms, we first calculate that $k = 5$ according to condition (26). Recall that mature firms' payout policy is unchanged up to period $t - k - 1 = 5$, in which they reduce their dividend payments. From period $t - k = 6$ to period $t - 1 = 10$, firms make no dividend payments and accumulate an increasing amount of cash on their balance sheets. This allows them to invest more in case an investment opportunity arises, but only up to a maximum of I^*, since the marginal product of capital turns negative after this point.

During these periods, aggregate investment AI and output AY is significantly above its long-run equilibrium value. In period $t = 11$ mature firms pay a special dividend, which returns them to their steady state for the periods to come. New firms can now issue a slightly larger amount of equity, but as the figure illustrates, aggregate investment and output are lower once the anticipated dividend tax cut has materialized than during the periods when it was expected.

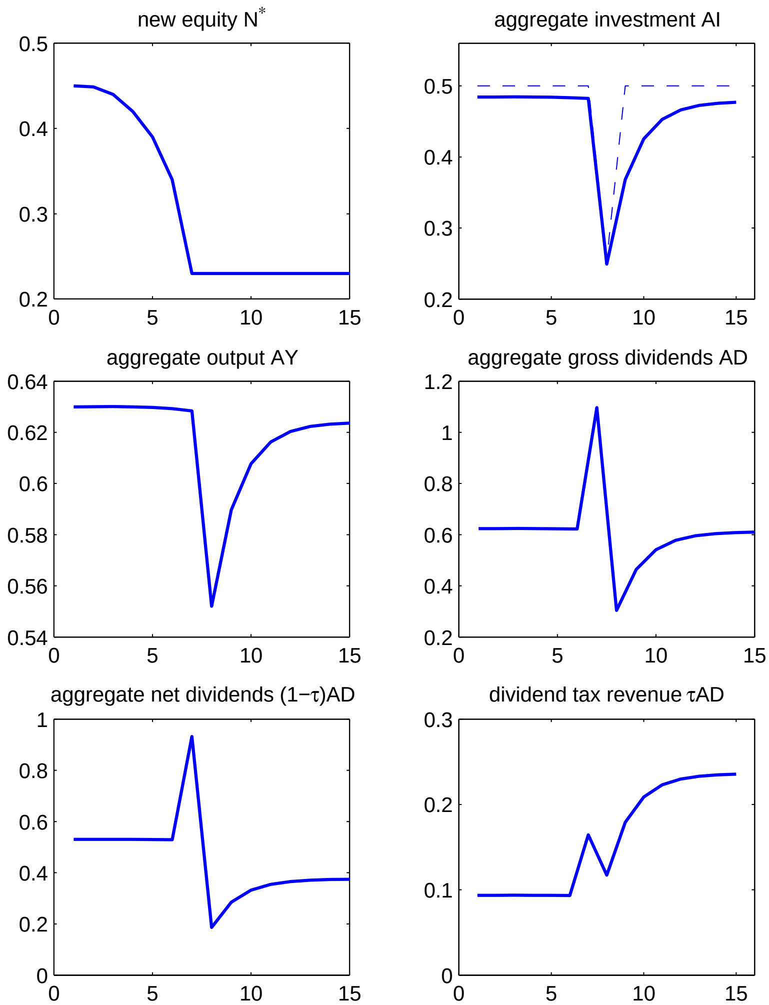

Figure 10: Anticipated Dividend Tax Increase: The amount of equity N^* that new firms issue starts falling when the tax increase is announced. This causes a modest decline in aggregate investment, output, and dividend payments, which is hardly recognizable in the figure. In the period preceding the dividend tax increase, here period 7, firms pay out a special dividend. This reduces aggregate investment AI_t, output AY_t and dividends AD_t for periods to come. (The dotted line in the graph for investment represents the path for the optimal cash holdings M^*. It lies above AI since not all firms are in steady state.)

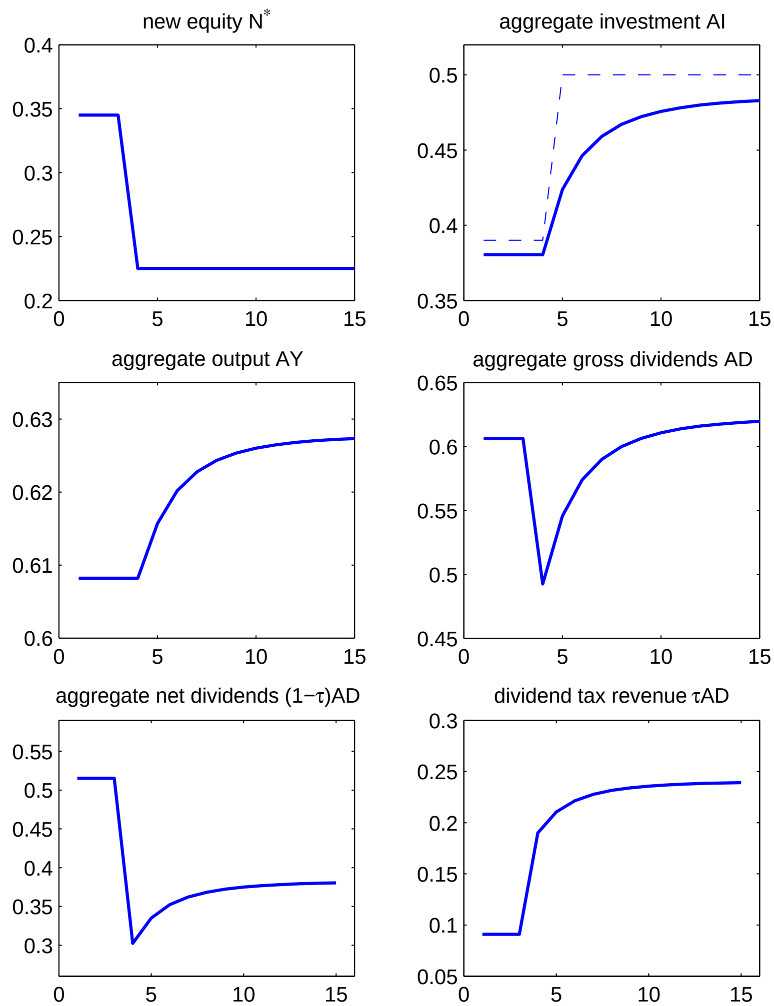

Figure 11: Risk of Dividend Tax Increase: In periods 1 to 3 firms anticipate that there is a $\pi = 25\%$ risk of a dividend tax increase from 15% to 38.6%. This reduces their cash reserves for investment purposes to M_π^* and depresses aggregate investment and output. In the given example, the tax increase materializes in period 4, and firms' optimal cash reserves (the dotted line in the upper right pane) increase to M^*. Aggregate investment increases even though new firms issue less equity N^* under the new dividend tax rate.

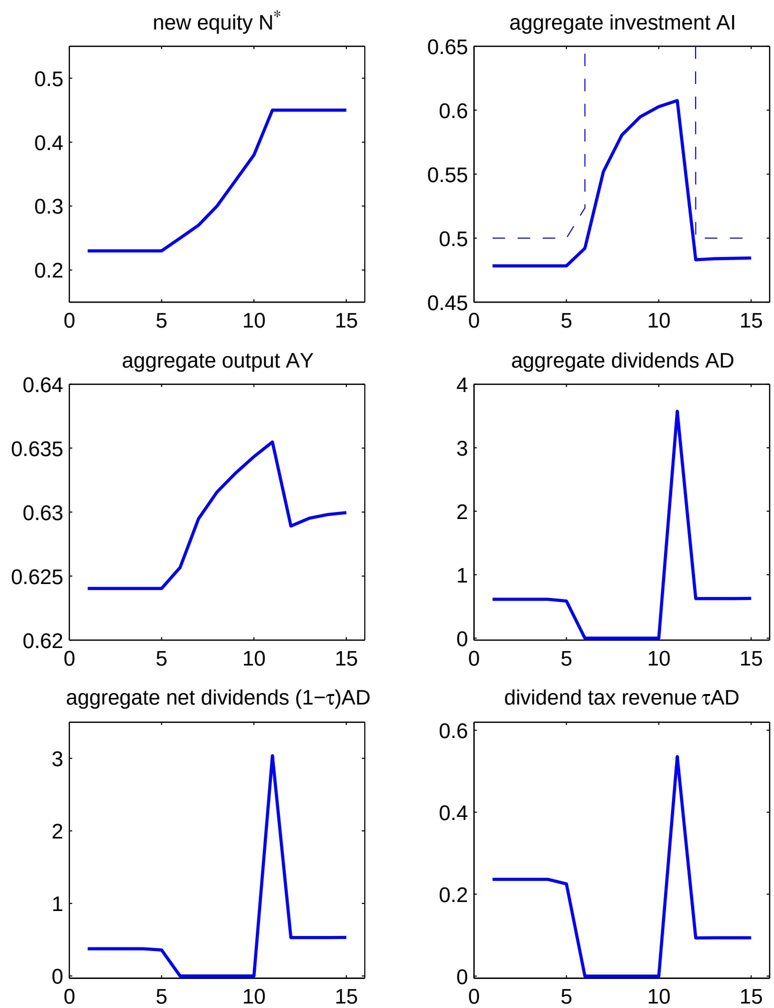

Figure 12: Anticipated Dividend Tax Cut: At $k = 5$ periods before an anticipated dividend tax cut from 38.6% to 15%, firms stop paying dividends AD and accumulate cash instead. This increases the aggregate amounts of investment AI and production AY that firms can engage in. After the tax cut is enacted, here in period 11, firms pay out all their excess cash and return to the steady state. The amount of equity that new firms issue, N^*, increases as a result of the tax cut, but the overall effect on aggregate investment once the cut has materialized is lower than in the periods of anticipation.

In comparing this example with the case of the unanticipated tax increase of figure 9, we can see that the effects on all aggregate variables are by an order of a magnitude higher in the case of anticipated tax changes. This section thus has five very clear implications for policymakers:

1. Unanticipated changes in dividend taxes have only insignificant macroeconomic effects, but strong redistributory consequences.

2. If policymakers want to increase dividend taxes *or* if firms expect an increase in the future, it is better to enact it immediately.

3. If policymakers want to reduce dividend taxes *or* if firms expect a reduction in the future, back-load the tax cut and keep firms waiting.

4. Unanticipated temporary dividend tax cuts are equivalent to an immediate unanticipated tax cut followed by an anticipated tax increase at a later time; they have an overall negative effect on investment and output.

5. Conversely, unanticipated temporary dividend tax increases have an overall positive effect on investment and output.

7 Political Economy of Dividend Tax Changes

In the previous section, we analyzed the effects of a tax increase which is expected with a certain probability π every period. In democratic societies, there are frequent changes in parties in power, with different views of taxes, and these parties find it tempting to adjust the level of dividend taxes according to their preferences when they come to power. Evaluating the impact of tax policies thus requires an analysis in a dynamic context that takes the possibility of regime changes and future changes in tax policy into account. In this section, we explore the consequences.

Our analysis here is of even greater importance, since we have shown in the previous sections that the level of dividend taxation itself does not have strong effects on aggregate investment and output, but that anticipated changes in the tax rate, or – as demonstrated in propositions 9 and 12 – even uncertainty about future changes in the tax rate can introduce significant distortions into the economy.

7.1 Markov Switching Between Two Tax Regimes

Here, we simply assume there are two representative parties in the political spectrum, labeled conservative C and social democrat S, and that each has committed itself to a tax rate of τ_L and τ_H respectively, where $\tau_L < \tau_H$. Party rule follows an exogenous symmetric Markov process, i.e. there is a probability π of a switch in party rule every period.[28] We want to investigate the dynamic consequences for investment, output, and dividend payments in the periods in which the two parties are in office.

[28] This is an abstraction from reality, since we would expect a party that better implements the preferences of its constituents to stay in power for longer. However, it can be argued that the level of dividend taxation is such a specific topic that elections are not usually lost or won because of this single issue.

Suppose first that the social democratic party is in power. Then firms recognize that every time period, there is a probability of π of tax rates being lowered. Accordingly, as proposition 12 established, mature firms increase their holdings of cash, which allows them to invest more when an investment opportunity comes up and raises output. We denote their optimal cash balances in this setting of Markov switching as $M^*(\tau_H, \tau_L; T_\pi)$ or M_H^* in short, where $M_H^* > M^*$ and $T_\pi = \begin{pmatrix} 1-\pi & \pi \\ \pi & 1-\pi \end{pmatrix}$ is the transition matrix containing the probabilities of keeping/switching party rule. As long as the social democrats stay in power, firms will only pay out dividends if their cash holdings exceed M_H^*.

If the conservatives are elected, taxes are immediately reduced to τ_L, and mature firms instantaneously recognize that they have too much cash. In anticipation of the risk of an increase in taxes next period, they pay out a large dividend now, bringing their cash holdings down to $M^*(\tau_L, \tau_H; T_\pi)$ or M_L^* in short, where $M_L^* < M^*$. As a result, investment and output fall. The next time that the social democrats win, there is an immediate tax increase, dividend payments get reduced, and cash balances increase, as do investment and output.

This short description shows the anomalous consequences of intertemporal tax arbitrage in the given setting: output and investment are higher under the high tax regime of the social democrats, lower under that of the conservatives. Through their commitment to lower taxes, the conservatives help the social democrats; but through their commitment to raise taxes, the social democrats hurt the conservatives. The conservatives exert a positive externality on the social democrats; the social democrats a negative externality on the conservatives.

Analytically, let us denote firms' value function under a conservative government as $V(M_t; \tau_L, \tau_H; T_\pi)$, or $V_L(M_t)$ in short, and that under a social democratic government as $V(M_t; \tau_H, \tau_L; T_\pi)$ or $V_H(M_t)$. We can then find for $i, j \in \{L, H\}, i \neq j$:

$$V_i(M_t) = \max_{D_t} (1-\tau_i)D_t + \beta\pi E V_j\left(\tilde{M}_{t+1}\right) + \beta(1-\pi)EV_i\left(\tilde{M}_{t+1}\right)$$

Taking the first order conditions of this maximization problem, we obtain

$$1-\tau_i = \beta\pi\left[pV_j'\left(F(M_i^*)\right)F'(M_i^*) + (1-p)V_j'\left((1+r)M_i^*\right)\right] + \\ + \beta(1-\pi)\left[pV_i'\left(F(M_i^*)\right)F'(M_i^*) + (1-p)V_i'\left((1+r)M_i^*\right)\right] \quad (28)$$

Now note that our observation that firms hold less cash under conservative rule than under social democratic rule, or $M_H^* < M^* < M_L^*$, implies that $V_L'(M_H^*) = 1-\tau_L$, but $V_H'(M_L^*) > 1-\tau_H$. We can use the result on $V_L'(M_H^*)$ in equation (28) to derive the marginal value of the optimum cash balances M_H^* under conservative rule as the following expression:

$$pF'(M_H^*) + (1-p)(1+r) = \frac{1-\tau_H}{\beta\left[\pi(1-\tau_L) + (1-\pi)(1-\tau_H)\right]} \quad (29)$$

As in our discussion on probabilistic tax cuts in subsection 5.5, if the switch from the high to the low tax regime is sufficiently large and/or likely, firms accumulate *all* their earnings as long as the high dividend tax rate is in effect, i.e. $M_H^* = \infty$. We can determine the threshold for this by setting $F' = 1+r$ in the equation above. Firms

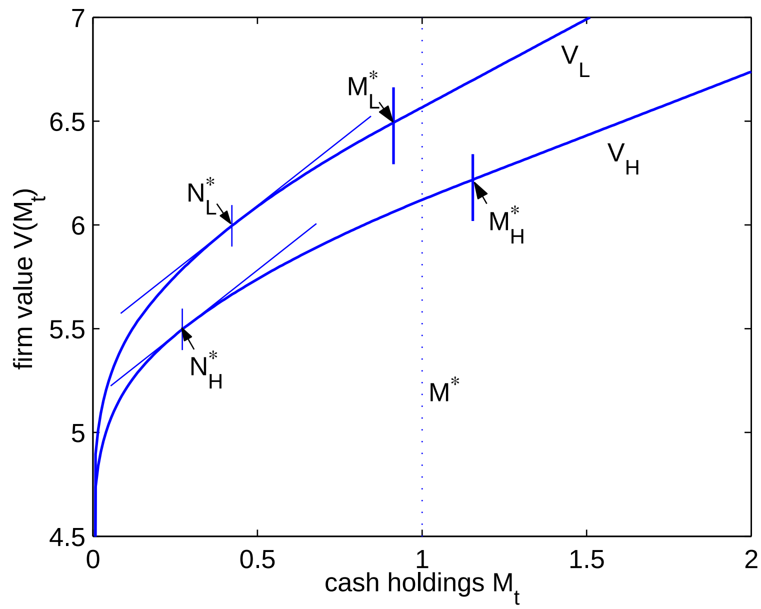

Figure 13: Value Function of Firms Under Conservative and Social Democratic Governments: Party rule changes with an exogenous probability of π. The conservative government cuts taxes upon entering office, which shifts the value function up to V_L and increases the amount of equity that firms are willing to issue. However, because of the risk of a dividend tax increase when the next social democratic government comes to power, firms' optimal cash holdings are depressed to M_L^*. The opposite results apply for the value function V_H under social democratic rule.

will accumulate all their cash holdings ($M_H^* = \infty$) if the marginal value of immediate dividend payments at the prevailing high tax rate is less than the marginal value of holding cash at interest rate r and postponing payment in expectation of a lower dividend tax rate in the future:

$$1 - \tau_H < \beta(1+r)(1 - E\tau_{t+1})$$

where $E\tau_{t+1} = (1-\pi)\tau_H + \pi\tau_L$ is the expected one-period-ahead tax rate.

The value for M_L^* cannot be expressed explicitly, but it can be obtained by using numerical methods. We have plotted an example of the value functions as well as M_L^* and M_H^* under Markov switching between two tax regimes in figure 13.

The figure also illustrates some implications for aggregate investment dynamics in an economy where the dividend tax rate fluctuates according to the described Markov process. Since $V(M)$ is concave, the decrease in dividend taxes that is expected during social democratic rule raises cash reserves more than the expected increase during conservative rule reduces them. Consequently, aggregate investment in the economy is on average higher if the economy constantly switches between low and high dividend taxation during conservative and social democratic rule than if the dividend tax rate remained constant at the average of the two rates.[29]

[29] Since production is concave in investment, the implications for aggregate output depend on the exact parameters of the model. In simulation results we found some pairs of tax rates which implied that expected output was higher if the economy switched between two tax rates than if either of the

7.2 Parties' Optimal Tax Rates

The previous subsection assumed that the tax rates of each party were given; but in fact, they are a matter of choice. In this subsection, we model the game between a conservative party C and a social democratic party S in a perfect information setup. As parties choose a tax rate, they recognize that they cannot bind successors, that firms know that, and that this will affect the behavioral response to their policies. Heuristically, this means that the conservatives recognize that the lower they set their taxes, given the policy of the social democrats, the more investors will fear a victory of the social democrats; the higher the dividend payments and the lower investment during conservatives' rule. Hence, their optimal tax rate is higher than it would be without strategic considerations.

By the same token, the social democrats recognize that the higher they set their taxes, given the policy of the conservatives, the lower dividends firms pay and the higher cash balances they accumulate. Social democrats thus gain from higher dividend tax rates in terms of both investment and output. Hence, the presence of conservatives who will subsequently lower taxes means that social democrats impose taxes that are even higher than they would be if social democrats were permanently in power.

The bias towards high taxes is further exacerbated if each party takes into account the consequences of its actions on the *relative* performance of the economy under its regime. That is, social democrats will raise taxes, recognizing not only that this improves the economy while they are in office, but also that it worsens the economy during their rival's regime. The conservatives raise taxes further, recognizing that in doing so, they lower the gain that the social democrats have while in office. Such a focus on relative performance might be particularly important if relative performance plays an important role in determining voting behavior.[30]

In this subsection, we analyze two kinds of political equilibria: first the ex ante non-cooperative Nash equilibrium, in which both parties choose a dividend tax rate that is the best response to their rival's tax rate. Second we investigate cooperative equilibria, which address the discussed bias towards excessive taxation. We limit our attention to equilibria in which each party implements a single fixed tax rate when it comes to power and leaves this rate unchanged until it looses power, i.e. we abstract from strategies that involve randomization or variations in tax rates without a change in party rule.[31]

Let us first define parties' preferences. At the most general level, we can postulate that each party's utility is a function of a vector X_t of macroeconomic variables in every period and an indicator function 1_t^i that takes the value of one when party i is

two tax rates was kept indefinitely. This seems to be an example where random taxation yields a welfare-superior outcome, as discussed in Stiglitz (1982).

[30] In the analysis of this section, however, we assume that this issue does not affect the probabilities of getting elected or reelected – this would complicate the analysis further without altering the qualitative results. The assumption may be justified on the grounds that the outcome of elections is determined by such a large array of issues that dividend taxation and its consequences can be ignored.

[31] This restriction can be justified by the observation that dividend tax rates have changed only infrequently in the past, perhaps because bringing the necessary measure through the legislative branch of government requires a certain fixed cost in terms of political capital.

in power and zero otherwise.

$$U^i = U^i \left(\{X_t, 1_t^i\}_{t=1}^{\infty} \right) \tag{30}$$

X_t can contain variables such as output, profits, tax revenues (in a more general model, wages), and parties have rational expectations about firms' behavior and the macroeconomic outcomes of this behavior. A party can be more sensitive to the value of those variables in the periods in which it is in office; alternatively, it may attach value to differences between the value of the relevant variables when it is in office and when it is not. One party (here we assume the conservatives) may prefer lower tax rates, say because it reflects more corporate interests, the other may prefer higher tax rates, because it values more the public goods which can be purchased with the tax revenue.

Each party controls, of course, only the tax rate τ_t when it is in power, i.e. when $1_t^i = 1$. The party that is in power sets the dividend tax rate that maximizes its expected utility:

$$\max E \left[U^i \left(\{X_t, 1_t^i\}_{t=1}^{\infty} \right) \right]$$

The vector X_t is generated by firms' behavior, which in turn depends on the vector of all firms' cash holdings $\{M_{t,z}\}$ with $z \in [0, z_t]$, on the current dividend tax rate, and on firms' beliefs about future dividend taxes, which we represent by the distribution function $\Gamma_t(\tau_{t+1}, \tau_{t+2}, ...)$:

$$X_t = X \left(\{M_{i,t}\}_i, \tau_t, \Gamma_t(\tau_{t+1}, \tau_{t+2}, ...) \right)$$

Firms form rational expectations about parties' dividend tax policies and about potential changes in party rule, i.e. $\Gamma_t(\cdot)$ is the result of parties' optimizing behavior. Given its current cash holdings $M_{t,z}$, the current dividend tax rate and beliefs about future tax rates, each firm $z \in \{0, z_t\}$ chooses its optimal investment $I_{t,z}$ and dividend payments $D_{t,z}$ by maximizing the following optimization problem subject to the standard cash-flow and dividend non-negativity constraints (5), (6):

$$V(M_{t,z}) = \max_{D_{t,z}, I_{t,z}} (1 - \tau_t) D_{t,z} + \beta EV \left\{ (1 + r) [M_{t,z} - D_{t,z}] + \tilde{\lambda}_{t,z} G (I_{t,z}) \right\} \tag{31}$$

The result is an infinite horizon rational expectations game between the two parties and all firms in the economy. An equilibrium in this game can be defined as

- a series of tax rates $\{\tau_t\}_{t=1}^{\infty}$ which satisfy in every period t the optimization problem (30) of the party i that is in power that period, given party i's beliefs on the future behavior of firms and of its rival
- a series of vectors of firms' dividends $\{D_{t,z}\}$, investment decisions $\{I_{t,z}\}$ and money balances $\{M_{t,z}\}$ which satisfy firms' optimization problem (31), given firms' beliefs on the future dividend policy
- in which both parties' beliefs about their rival's behavior and firm behavior are consistent with their rival's optimization problem, firms' maximization problem, and the Markov chain that determines the probabilities of regime change

- and firms' beliefs are consistent with both parties' optimization problem (30) and the Markov chain that determines the probabilities of regime change.

In order to characterize this game in more detail, we make a number of simplifying assumptions. Firstly, suppose that whenever conservatives or social democrats are in power, they implement the constant tax rates τ_C and τ_S respectively. We can then denote the utility function for the conservatives as $U^C(\tau_C, \tau_S)$. The optimal tax rate of any party (in power) depends on its beliefs of what its rival will do when it gets in power. In a more extended version of this model, optimal policies would depend on how long the previous party was in power (and more generally, on the whole history, perhaps summarized through a state variable, like the vector of firms' cash balances), as well as on beliefs, not just about what policies a party's successor will undertake when it is in power, but what the party itself will do when it regains power, and what its rival will do when it retakes power, and so forth. The whole past history and details of future beliefs matter. But given our simplifying assumption, the entire structure is Markovian, so that each party knows that what is optimal for it to do now (given expectations of its rivals actions) is the same as what is optimal when it retakes power, and similarly for its rival. It is important to recognize that we are not assuming that the party is making a decision today that commits it to a tax rate whenever it comes into power; it is only taking an action today, but thinking through the full consequences, knowing the Markovian structure of the model.

We assume $U^C_{11} < 0$, so there is a (unique) optimal tax rate τ_C for a given level of τ_S. Moreover, $U^C_{12} > 0$, for $\tau_S > \tau_C$. The marginal benefit to the conservatives of increasing the tax rate is increased when the social democrats increase their tax rate. This is because the larger the difference in tax rates, the lower cash balances held by firms will be while conservatives are in power, and hence the lower incomes will be, the lower their tax revenues will be, and the higher the marginal value of increasing taxes.[32]

We define $U^S(\tau_S, \tau_C)$ analogously and assume that it has similar properties. In particular, $U^S_{12} < 0$ for $\tau_S > \tau_C$. When conservatives increase their tax rate, firms reduce cash balances under conservatives' rule by less and raise them by less under social democratic rule. This reduces the marginal benefit that social democrats receive from raising their tax rate.

7.2.1 Non-Cooperative Nash Equilibrium

Let us now analyze the non-cooperative game in which the two parties play the constant tax rates τ_C and τ_S. Given the utility function $U^C(\cdot)$ from above, we define τ_C^* as

$$\tau_C^* = \arg\max_{\tau_C} U^C(\tau_C, \tau_C)$$

[32]Note that this implicitly assumes that strategic considerations regarding the effect of tax rates on the level of output are more important than considerations of the effect of tax rates on government revenue or capital income, which could potentially change the assumed sign of the derivative.

i.e. as the tax rate which, if the conservatives could choose a single tax rate forever, would maximize their welfare. We define τ_S^* similarly.[33] We assume furthermore that

$$\tau_C^* < \tau_S^*$$

i.e. the conservatives prefer lower tax rates. This might be because they weigh capital income (net of taxes) more highly or – under certain conditions – because they value government revenue less in their utility function. In figure 14, the tax rates τ_C^* and τ_S^* are represented by crosses on the dotted diagonal defined by $\tau_S = \tau_C$.

Next we define $\hat{\tau}_C(\tau_S)$ as the conservatives' best response function to the social democrats' constant tax rate τ_S, i.e.

$$\hat{\tau}_C(\tau_S) = \arg\max_{\tau_C} U^C(\tau_C, \tau_S)$$

and similarly for $\hat{\tau}_S(\tau_C)$. Both functions are depicted in figure 14. We have not yet made an assumption on timing here, i.e. on which party is in power first. The described setup would work both for the case that one party is known to be the first mover and for the case that parties have to commit to a constant tax rate *ex ante*, i.e. before they know which one will be in power first, such as on the eve of an election. Naturally, the chosen tax rates would be different depending on this factor, though the differences would be limited by the fact that the long-run probability for any party to be in power is one half. This is in turn guaranteed by the fact that the symmetric matrix of transition probabilities in the Markov chain that determines party rule.

The results of the analysis below depend critically on the following two properties:

Property 1 *:* $\hat{\tau}_C(\tau_S) > \tau_C^*$ *for* $\tau_S > \tau_C^*$

Because the social democrats will set a tax rate τ_S higher than the conservatives' optimal rate τ_C^* when they are in power, the conservatives will want to increase the tax rate which they levy when they are in power. The reason is because of the intertemporal arbitrage: knowing that the tax rate will be higher in some later period (when the social democrats are in power), investors distribute more dividends now, lowering investment and income today. By raising the tax rate (closer to the rate levied by the social democrats), the incentive to distribute dividends is reduced, and national income during the period in which the party is in power is increased.

While the social democrats exert a negative externality on the conservatives, the conservatives exert a positive externality on the social democrats. If the conservative party cares about the relative performance of the economy under its rule, then this provides a further argument for them for increasing the tax rate: for the lower the tax rate, the larger the "arbitrage," i.e. the higher the retained earnings of firms under social democratic rule, and hence the higher investment and output.

Property 2 *:* $\hat{\tau}_S(\tau_C) > \tau_S^*$ *for* $\tau_C < \tau_S^*$

By a similar argument, because the conservatives have lowered the tax rate below the optimal level τ_S^* of the social democrats, the latter benefit from increasing the tax rate, which raises investment and output under their rule.

[33] Under some parameter values it might be optimal for a party to randomly switch tax rates even if it knew that it would stay in power forever. However, in the analysis here we rule out the possibility that random taxation can increase a party's welfare.

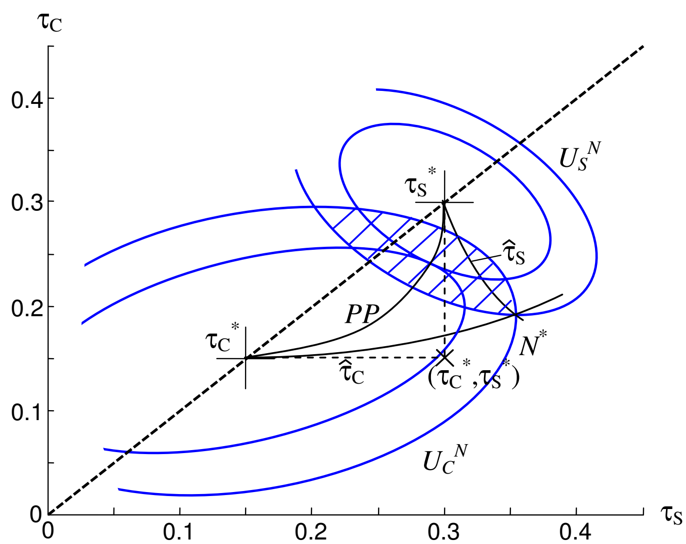

Figure 14: Parties' Iso-Utility Functions in the Nash Equilibrium: The graph depicts conservatives' and social democrats' dividend tax rates on the horizontal and vertical axis respectively. The preferred tax rates, in the given example $(\tau_C^*, \tau_S^*) = (15\%, 30\%)$ are indicated by crosses and are surrounded by concentric indifference curves, with parties' utility declining the further away they move from their optimum. At the Nash equilibrium, here at $(\tau_C, \tau_S) = (20\%, 35\%)$, parties' iso-utility curves U_C^N and U_S^N intersect in a right angle: no party can alter its tax rate individually without loosing utility. This point marks also the intersection of parties' reaction functions. On the other hand, if both parties cooperate to lower their tax rates, then they can both raise their utility: all pairs of tax rates inside of the shaded lens formed by the two iso-utility curves can be reached in a cooperative agreement.

The Nash equilibrium $\mathcal{N}^*$ of the game is the pair of tax rates (τ_C^N, τ_S^N) that solves the equations

$$\hat{\tau}_C(\tau_S^N) = \tau_C^N \quad \text{and} \quad \hat{\tau}_S(\tau_C^N) = \tau_S^N$$

or, combining the two equations,

$$\hat{\tau}_C(\hat{\tau}_S(\tau_C^N)) = \tau_C^N$$

It is characterized by $\tau_C^N > \tau_C^*$ and $\tau_S^N > \tau_S^*$, i.e. the non-cooperative equilibrium entails higher tax rates than the parties would have supported on their own and is clearly not Pareto efficient. Figure 14 depicts our analysis graphically.

7.2.2 Cooperative Equilibria

In the above analysis, we had each party choose a constant tax rate, taking the (constant) tax rate chosen by the other as given. Making policy depend on the previous history of tax policies opens up the possibility for a richer set of equilibria, in particular the possibility that parties can cooperate on how they set their tax rates.

We thus investigate next how parties can improve upon their welfare through cooperation. We noted that both parties have a bias towards excessive dividend taxation, since each party's choice of a dividend tax rate imposes an externality on the rival that induces the other party in turn to raise their tax rate. A cooperative agreement between the two can mitigate this bias.

While the described game allows for a very rich set of possible equilibrium outcomes, we focus again on pure strategy equilibria with a fixed dividend tax rate for each party. Among these equilibria, we will further describe the set of Pareto optimal equilibria.

Let us first determine the participation constraints for any such cooperative agreement. We define $\mathcal{C}^*$ as the set of all pairs of dividend tax rates $(\tau_C^{CO}, \tau_S^{CO})$ for which both parties' ex-ante utility is greater or equal than in the non-cooperative Nash equilibrium $\mathcal{N}^*$:

$$U_C(\tau_C^{CO}, \tau_S^{CO}) \geq U_C(\tau_C^N, \tau_S^N) \quad \text{and} \quad U_S(\tau_S^{CO}, \tau_C^{CO}) \geq U_S(\tau_S^N, \tau_C^N)$$

Figure 14 shows the Nash equilibrium $\mathcal{N}^*$ as well as the corresponding iso-utility curves U_i^N of the two parties, which indicate all tax pairs for which party i reaches the same utility as in the Nash equilibrium. The shaded lens between the two iso-utility curves depicts the set $\mathcal{C}^*$ of possible cooperative pairs of tax rates. Clearly, any cooperative tax rate must be lower than the corresponding tax in the non-cooperative Nash equilibrium $\mathcal{N}^*$, i.e. $\tau_i^{CO} < \tau_i^N$. This is because lowering taxes from $\mathcal{N}^*$ reduces the externalities that parties impose on each other.

The set $\mathcal{C}^*$ does not include the point where each party plays its preferred tax rate (τ_C^*, τ_S^*) and the point where both parties play conservatives' preferred tax rate (τ_C^*, τ_C^*). However, depending on the parameters, it may include the point (τ_S^*, τ_S^*) – in figure 14 it does not. If some part of the diagonal $\tau_C = \tau_S$ is included in $\mathcal{C}^*$, then a cooperative agreement in which both parties keep a constant, identical dividend tax rate is feasible – in the given example in the figure this is the case. Based on the folk theorem we can derive the following result for all pairs of tax rates in $\mathcal{C}^*$:

Proposition 14 *For each pair of taxes* $(\tau_C^{CO}, \tau_S^{CO}) \in \mathcal{C}^*$ *and for sufficiently low discount rates, the following strategy constitutes a cooperative equilibrium for* $i, j \in \{C, S\}, i \neq j$*: (1) Play* τ_i^{CO} *in the first period and as long as the rival does not deviate from* τ_j^{CO}*. (2) Play* τ_i^N *forever if the rival has ever deviated from* τ_j^{CO} *in the past.*

The utility of both parties in such an equilibrium is weakly higher than in the non-cooperative Nash equilibrium $\mathcal{N}^*$, and thus – given a high enough discount factor – they both have an incentive to make a cooperative agreement sustaining one of the tax pairs in set $\mathcal{C}^*$. A deviation from the agreement would return them to $\mathcal{N}^*$ forever after.

The upward-sloping $\mathcal{PP}$ line in figure 14 depicts the set of Pareto-optimal pairs of tax rates, i.e. of all pairs (τ_C, τ_S) such that no change in tax rates can improve on one party's utility without reducing its rival's utility. Along this locus, the iso-utility curves of both parties are tangents to each other. We denote this set as $\mathcal{PP}$. Analytically, it is defined as the locus of all pairs (τ_C, τ_S) such that

$$\frac{\partial U_C(\tau_C, \tau_S)}{\partial \tau_C} \cdot \frac{\partial U_S(\tau_S, \tau_C)}{\partial \tau_S} = 1$$

which is simply the condition for the indifference curves U_C and U_S to be tangents. Note that only those Pareto-optimal pairs of tax rates that are also in set $\mathcal{C}^*$, i.e. all pairs $(\tau_C, \tau_S) \in \mathcal{C}^* \cap \mathcal{PP}$, can be sustained through cooperative agreements.

We can describe various mechanisms for choosing which of the cooperative equilibria is chosen, for example the one according to Rubinstein (1982). However, note that cooperative equilibria in the sequential game between two political parties are not renegotiation-proof: Whenever party rule changes and party j comes to power, it is optimal for party j to re-negotiate its agreement with party i, since it now has greater bargaining power. Similarly, party i will find it optimal to accept the request for a new agreement, but in turn re-negotiate when party i comes to power again. As a result, both parties know that their rivals will re-negotiate any agreement when they come to power, and all the discussed cooperative equilibria break down. This implies that the Nash equilibrium $\mathcal{N}^*$ is the only equilibrium in the described game that is renegotiation-proof.

In order to maintain a cooperative equilibrium, it is necessary to find some external mechanism to enforce it, such as a constitutional law that cannot be changed unilaterally at a later time, or a sufficiently high penalty for a party that deviates from an agreement and attempts to renegotiate when it comes to power. In a world in which explicit penalties for breaking the "contract" cannot be imposed, similar results can be obtained by policies which have commitment-like effects, i.e. which impose a cost on the party deviating from the cooperative equilibrium. Constitutional amendments that prevent changes in tax rates are very costly, because they do not allow changes in response to changing circumstances. But a provision that e.g. delayed the implementation of any tax change for a number of years would reduce the benefits of, say, the Social Democrats from raising taxes above the cooperative level, because there is a high probability that they would no longer be in office when the change comes into effect.

7.3 Generalizations

We have modeled the political game between two parties that have different preferences over dividend tax rates, and for which economic performance under one party's rule is affected by the other party's choice of dividend tax rates. Similar considerations hold in any situation, in which

1. Governments are contestable, i.e. the political parties in power and thus government policies change over time.
2. These changes in policies make it optimal for private economic agents to shift certain policy-relevant actions across time.
3. The change in private sector behavior affects macroeconomic variables, which in turn has an impact on parties' utility.

Other examples for such situations include:

- income taxes and reallocations in labor supply/compensation
- income taxes and payments into/withdrawals from tax-deferred IRAs, after-tax IRAs
- capital gains taxes and stock sales
- sales/VAT-taxes and purchases of durable consumer goods
- corporate profit taxes and a variety of corporate decisions, such as corporate investment, repatriations of foreign profits, executive compensation
- public infrastructure and complementary private investment
- environmental taxes/regulations and investment in green technologies

For each of these examples, there is empirical evidence that changes in policy have intertemporal effects that could lead to strategic "arbitrage." In general, the importance of these strategic effects of policies will be larger the higher private agents' ability (or the lower their cost) of shifting their actions across time. Furthermore, the strategic importance increases the larger the effect of these changes in private sector behavior on parties' utility.

It can also be shown that the introduction of certain devices to shift tax load across periods, such as e.g. IRAs, benefits the utility of one party at the expense of the other party.

We leave a more detailed analysis of the political game between parties in such situations for future research.

8 Conclusions

This paper has presented a model of the life cycle of capital constrained firms that allowed us to investigate the dynamic effects of dividend taxation on investment and output. Firms in our model start out by issuing equity, then they accumulate more funds through retaining their earnings, i.e. internal saving, and when they reach the mature stage, they pay out dividends.

The arguments of the traditional view of dividend taxation apply to firms that are in the first stage: since the level of the dividend tax rate affects the valuation of firms, it has an impact on how much equity new firms issue. On the other hand, the new view of dividend taxation applies to firms in the second and third stage. Their corporate saving and investment decisions are not distorted by dividend taxes, so long as the rate is constant, even though their (after tax) value to shareholders decreases proportional to $1-\tau$. Since only a small fraction of firms in a typical economy are in the first stage, we have argued that the level of dividend taxation has only a minor impact on aggregate investment and output.

We then used the model to investigate the effects of changes in the dividend tax rate. Unanticipated tax changes have a small impact on aggregate investment, since they affect how much equity new firms issue. However, anticipated dividend tax changes create opportunities for inter-temporal tax arbitrage, and this can distort firms' cash balances and investment decisions significantly. An anticipated dividend tax cut makes mature firms delay their dividend distributions to the period when the tax rate is reduced, which implies that they have more cash on hand for investment purposes when an investment opportunity arises. At the same time, anticipated tax cuts increase the value of firms already in the periods ahead, which raises the amount of equity that new firms issue. Overall, anticipated tax cuts strongly raise aggregate investment in the periods prior to the cut. Similarly, an anticipated dividend tax increase causes mature firms to pay out excessive dividends, which reduce aggregate investment for a number of periods after the tax hike until firms have recovered their optimal level of cash holdings. Anticipated tax hikes also depress the stock price of all firms for several periods before they come into effect, which reduces the amount of equity that new firms can raise. In total, anticipated increases in the dividend tax rate reduce aggregate investment before and after the increase comes into effect.

An unanticipated temporary dividend tax cut, such as the one enacted in 2003 in the United States, can be seen as an unanticipated tax cut, which has small positive effects on aggregate investment, followed by an anticipated tax increase at the expiration of the law, which has strong negative effects. Overall, the net effect of such a tax change on aggregate investment is very likely to be negative.

By the same token, uncertainty about whether or how long a government will keep a low dividend tax rate in place can lead to a phenomenon similar to the peso problem: firms assign a certain probability to a dividend tax increase, which leads them to pay out more cash than optimal in every period until the expectations of the tax increase have realized. In such a situation, it would be optimal to bring the tax increase forward – this would increase aggregate investment starting in the following period.

In a contestable democracy, the analysis of dividend taxation over time has to take into account that different parties with different preferences regarding dividend taxes come to power. This can lead firms to expect changes in the dividend tax rate whenever party rule changes. Since conservatives have incentives to reduce the expected hike in dividend taxes when they lose power and, conversely, social democrats have an incentive to increase the expected reduction in the tax rate when they lose power, both parties have an incentive to raise the dividend tax rate under their rule to an inefficiently high level.

While the structure of the problem analyzed here highlights the importance of intertemporal interactions in a very specific setting, they arise more generally. Our analysis suggests strongly that the analysis of behavioral responses to policies which typically assume such changes to be permanent are seriously flawed when there are dynamic interactions. In democratic societies with contestable elections, policies will vary with the party in office; no matter what assurances are given about changes being permanent, market agents rationally do not believe politicians' assurances. The analysis of the consequences for both behavioral responses of economic agents and the political economy responses of political actors is a rich field to be explored in future research.

References

Asquith, Paul and David Mullins (1986), "Equity Issues and Offering Dilution," *Journal of Financial Economics*, 15(1/2), pp. 61 - 89.

Atkinson, Anthony B. and Joseph E. Stiglitz (1976), "The Design of Tax Structure: Direct Versus Indirect Taxation," *Journal of Public Economics*, 6(1-2), pp. 55 – 75.

Auerbach, Alan J. (1979), "Wealth maximization and the cost of capital," *Quarterly Journal of Economics*, 93(3), pp. 433 – 446.

Auerbach, Alan J. (2002), "Taxation and Corporate Financial Policy," in *Handbook of Public Economics III*, edited by Alan J. Auerbach and Martin Feldstein, Amsterdam: North-Holland, pp. 1251 – 1292.

Auerbach, Alan J. and Kevin A. Hassett (2002), "On the marginal source of investment funds," *Journal of Public Economics*, 87, pp. 205 – 232.

Auerbach, Alan J. and Kevin A. 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, MA: Cambridge University Press.

Bhattacharya, Sudipto (1979), "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," *Bell Journal of Economics* 10(1), pp. 259 – 270.

Black, Fischer (1976), "The Dividend Puzzle," *Journal of Portfolio Management* 2(2), pp. 5 – 8.

Blouin, Jennifer L., Jana S. Raedy and Douglas A. Shackelford (2004), "Did Dividends Increase Immediately After the 2003 Reduction in Tax Rates?" NBER Working Paper w10301.

Congressional Budget Office (2003), "How CBO Analyzed the Macroeconomic Effects of the President's Budget," Washington DC.

Chen, Hsuan-Chi and Jay R. Ritter (2000), "The Seven Percent Solution," *Journal of Finance* 55(2), pp. 1105 – 1131.

Chetty, Raj and Emmanuel Saez (2005), "Dividend Taxes and Corporate Behavior: Evidence from the 2003 Dividend Tax Cut," *Quarterly Journal of Economics* 120(3), pp. 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, pp. 124 – 129.

DeAngelo, Harry and Linda DeAngelo (1990), "Dividend Policy and Financial Distress: An Empirical Investigation of Troubled NYSE Firms," *Journal of Finance* 45(5), pp. 1415 – 1431.

DeAngelo, Harry, Linda DeAngelo and Rene Stulz (2004), "Dividend Policy, Agency Costs, and Earned Equity," NBER Working Paper w10599.

Edlin, Aaron and Joseph E. Stiglitz (1995), "Discouraging Rivals: Managerial Rent-Seeking and Economic Inefficiencies," *American Economic Review* 85(5), pp. 1301 – 1312.

Fama, Eugene F. and Kenneth R. French (2001), "Disappearing dividends: Changing firm characteristics or lower propensity to pay?" *Journal of Financial Economics* 60, pp. 3 – 43.

Feldstein, Martin S. and Jerry R. Green (1983), "Why do companies pay dividends?" *American Economic Review* 73(1), pp. 17 – 30.

Grossman, Sanford J. and Oliver D. Hart (1980), "Takeover Bids, The Free-Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics* 11(1), pp. 42 – 64.

Grossman, Sanford J. and Stiglitz, Joseph E. (1977), "On Value Maximization and Alternative Objectives of the Firm," *Journal of Finance* 32(2), pp. 389 – 402.

Hettich, Walter and Stanley L. Winer (1999), *Democratic Choice and Taxation: A Theoretical and Empirical Analysis*, Cambridge University Press.

Hoon, Hian Teck and Edmund S. Phelps, "Tax Cuts, Employment and Asset Prices: A Real Intertemporal Model," Columbia University, mimeo.

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

Jensen, Michael C. (2001), *A Theory of the Firm: Governance, Residual Claims, and Organizational Forms*, Harvard University Press.

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

Mehra, Rajnish and Edward C. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15, pp. 145 - 161.

Mehra, Rajnish and Edward C. Prescott (2003), "The Equity Premium in Retrospect," in George M. Constantinides, Milton Harris, and Rene M. Stulz (eds.), *Handbook of the Economics of Finance*, Amsterdam, Netherlands: North-Holland.

Miller, Merton and Myron Scholes (1978), "Dividends and Taxes," *Journal of Financial Economics* 6(4), pp. 333 – 364.

Miller, Merton and Myron Scholes (1982), "Dividnes and Taxes: Some Empirical Evidence," *Journal of Political Economy* 90(6), pp. 1118 – 1141.

Myers, Stewart C. and Nicholas S. Majluf (1984), "Corporate financing and investment decisions when firms have information that investors do not have," *Journal of Financial Economics* 12, pp. 187 – 221.

Poterba, James and Lawrence Summers (1985), "The Economic Effects of Dividend Taxation," in Edward Altman and Marti Subrahmanyam (eds.), *Recent Advances in Corporate Finance*, Homewood, IL: Richard D. Irwin Publishers, pp. 227 – 284.

Ross, Stephen (1977), "The Determination of Financial Structure: An Incentive-Signalling Approach," *Bell Journal of Economics* 8(1), pp. 23 – 40.

Rubinstein, Ariel (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50(1), pp. 97 - 109.

Sinn, Hans-Werner (1991), "The vanishing Harberger triangle," *Journal of Public Economics* 45, pp. 271 – 300.

Stiglitz, Joseph E. (1973), "Taxation, Corporate Financial Policy, and the Cost of Capital," *Journal of Public Economics* 2, pp. 1 – 34.

Stiglitz, Joseph E. (1976), "The Corporation Tax," *Journal of Public Economics* 5(3-4), pp. 303 – 311.

Stiglitz, Joseph E. (1982), "Utilitarianism and horizontal equity: The case for random taxation," *Journal of Public Economics* 18(1), pp. 1 – 33.

Stiglitz, Joseph E. (2003), "Bush's Tax Plan – The Dangers," *New York Review of Books* 50(4).

Stiglitz, Joseph E., and Andrew Weiss (1981), "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71, pp. 393 – 410.