VATT Working Papers 41

Non-linear dividend tax and the dynamics of the firm

Seppo Kari
Jussi Laitila

GOVERNMENT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

41

Non-linear dividend tax and the dynamics of the firm

Seppo Kari
Jussi Laitila

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2014

Second edition, updated 26.11.2014.

ISBN 978-952-274-129-5 (PDF)
ISSN 1798-0291 (PDF)

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, November 2014

Non-linear dividend tax and the dynamics of the firm

Government Institute for Economic Research
VATT Working Papers 41/2014

Seppo Kari – Jussi Laitila

Abstract

We analyze the implications of a non-linear tax scheme for dividends using a life-cycle model of a firm. In this model new firms first enter markets, then grow internally financing from retained earnings and finally distribute their profits in the steady state. We find that under a non-linear tax the owners prefer a smooth flow of dividends, which encourages firms to begin distributions right from the start. This early distribution incentive (EDI) slows down investments and leads to delayed growth. Our calculations indeed confirm that a revenue-neutral switch from linear to progressive tax exacerbates production losses. We further demonstrate that this distortion can be reduced by carrying forward unused tax allowances with interest, as proposed e.g. by Mirrlees et al. (2011).

Key words: dividend tax, progressive tax, investment, firm behavior, early distribution incentive

JEL classification numbers: D92, G35, H24, H32, G35

Contents

1. Introduction

The last decade has witnessed a wave of reforms of dividend taxation in developed economies.[1] These reforms have attracted new research, in particular empirical, to better understand the consequences.[2] The theoretical debate over the impacts of dividend taxation has been long and rich in content. One of its main elements is the so-called new view, which emphasizes that personal taxes on dividends do not distort the investment or dividend decisions of a mature firm (King 1974, Auerbach 1979, Bradford 1981). Its main contender, the so-called old view, holds that dividend taxes affect both investment and pay-out decisions (Harberger 1966, Poterba and Summers 1985). Sinn (1991a) extends the analysis to the life cycle of a corporate firm and shows that the new-view neutrality result only applies in the long-run equilibrium. The firm's life cycle consists of three stages: start-up, internal growth and steady state. In the two early stages dividend taxes are distortionary and slow down the growth of the firm. The signaling model proposed by Gordon and Dietz (2008) and the agency cost model of Chetty and Saez (2010) further expand the range of approaches.

One common aspect of this literature is that it focuses almost entirely on proportional tax rates and pays virtually no attention to progressive dividend tax schedules (as noted by Gourio and Miao, 2010). This is in stark contrast to the observation that most countries impose tax schemes which include progressivity due to exemptions and gradually rising marginal tax rates. In a number of countries dividends are included in broadly defined taxable income and a progressive tax is levied on them. In some others, such as Denmark, Spain and the United States, dividends are taxed using a separate progressive rate schedule (OECD, 2014). While a smaller body of literature considers the investment and efficiency implications of a specific progressive tax, the so-called "split model" of the Nordic dual income tax (Lindhe et al., 2004; Kari and Karikallio, 2007; Sørensen, 2005), we are not aware of any study focusing on the effects of a generic non-linear tax rate schedule on the decisions of firms.

In this paper, we aim to fill this apparent gap in the literature by analyzing the implications of progressive dividend taxation on the investments and financing of a corporate firm. We introduce a linear-progressive tax scheme, which combines a proportional tax rate with an exemption limit, and analyze it in a life-cycle

*Kari: Government Institute for Economic Research (VATT), PO Box 1279, 00101 Helsinki, Finland (seppo.kari@vatt.fi); Laitila: University of Helsinki, Department of Biosciences, PO Box 65, 5205 University of Helsinki, Finland (jlaitila@helsinki.fi). The authors would like to thank Giampaolo Arachi, Luis Alvarez, Essi Eerola, Dominika Langenmayr, Teemu Lyytikäinen, Jukka Pirttilä, Olli Ropponen, Jan Södersten, Alfons Weichenrieder and two anonymous referees for their valuable comments.

[1] In 2003 the United States adopted a separate progressive tax scheme for dividends. Many EU member countries have switched from the imputation system to low separate tax rates for dividends. See e.g. OECD (2014).

[2] Examples of recent empirical studies are Chetty and Saez (2005), Brown et al. (2007), Bond et al. (2007), Auerbach and Hassett (2007), Kari et al. (2008), Becker et al. (2012), Jacob and Jacob (2012) and Yagan (2014).

model of an all-equity firm. We show that this tax provides incentives to even out dividend distributions over the firm's life cycle (early distribution incentive, EDI). The tax exemption encourages a financially constrained firm to start distributing dividends right from the beginning and therefore to grow at a slower pace than under a linear tax. On the other hand, the threshold also increases the size of the optimal start-up capital stock reducing any production delay. To assess the net effect of progressivity on growth, we introduce indicators which capture these opposite effects and make it possible to analyze their relative importance. We illustrate numerically that a revenue-neutral move from linear to progressive tax unambiguously slows down the growth of production.

In the second stage we augment the basic tax scheme with a new element that allows the owner to carry forward unused tax exemptions. Hence, if the firm distributes an amount lower than the tax threshold in a particular year, the difference can be saved and used in later years. Sweden and Norway apply carry-forward in their non-linear tax schemes for capital income and Mirrlees et al. (2011) included it in their proposal for the taxation of savings.[3,4]

The question addressed here is how carry-forward interacts with the EDI generated by progressive elements in the tax rate. We observe that carry-forward transforms the original static constraint on tax-exempt dividends to a life-cycle constraint and therefore broadens the choice opportunities open to the firm. We find that carry-forward may even eliminate the effects of progressivity on the firm's pay-out and growth altogether. However, we also find that at given levels of tax rates and exemption limits, carry-forward leads to reduced tax revenue.

Our paper contributes to the extensive theoretical literature studying the effects of capital income taxes on investment and financing. To our knowledge our paper is the first account of how progressive taxes on distributions affect the investment and dividend decisions of a growing firm. Its main novelty is in the two theoretical findings concerning EDI and how this can be avoided by allowing carry-forward of unused allowances.

Our paper ties in with studies on dividend taxes in a life-cycle framework (e.g. Sinn, 1991a; Korinek and Stiglitz, 2009; Auerbach and Hasset, 2007; and Lindhe and Södersten, 2009). It extends the model of Sinn (1991a) by introducing a linear-progressive tax schedule. In this context, one of the novel contributions of this paper is the observation that theory does not necessarily rule out dividend distributions during the growth phase. In fact, the outcome depends on the shape

[3] For carry-forward of unutilized allowances, see Sørensen (2005) and Lindhe and Södersten (2012)

[4] While the closest real-life counterpart of the extended model is the tax treatment of closely held firms in Sweden, our basic model corresponds to a regime for small-business owners in Finland (see Kari et al., 2008). In this tax system dividends are exempt up to 60,000 euros and at this point the personal tax rate on dividend income jumps from 0 to 21 per cent (2012). No carry-forward of any unused part of the exemption is allowed.

of the tax schedule and in our basic model dividends are distributed throughout the firm's life-cycle. Our paper also ties in with studies discussing the non-linear tax schemes of the Nordic countries. While some previous studies have considered similar tax structures, they have not analyzed their effects on the policies of a growing firm.

Our results suggest that to avoid distortions to investment decisions one should either have a linear tax schedule for distributions or, if progression is a priority, augment a non-linear scheme with carry-forward. While the intuition of the results is clear, these policy implications should nonetheless be applied with care. The model and the tax system are simple. There is, for example, only one method of rewarding owners: dividends. Extending the model to include other ways of extracting profits would complicate the owner's choices, with possible implications for the timing of dividends. Furthermore, in the model, owners have no outside income sources and there is no heterogeneity among the owners. Relaxing any of these constraints could narrow the applicability of the results considerably.[5] However, the results may carry substantial relevance at least in the case of closely held companies. With concentrated ownership or with several owners who all receive most or all of their income from their firm, the incentive effects detected may well shape behavior in real life. Finally, the model is deterministic. Introducing uncertainty would certainly raise interesting new issues, but not necessarily undermine the current results.

The remainder of this paper is organized as follows. Section 2 sets up the dynamic model of the firm under non-linear dividend tax. Section 3 provides and discusses the results. Section 4 introduces measures of tax-induced distortion and simulates the total effect of a tax threshold on production. Section 5 introduces the carry-forward element into the non-linear tax scheme. Section 6 discusses and concludes. Two Appendices provide proofs of the main propositions.

[5] Weichenrieder (1996) demonstrates that allowing alternative means to transfer profits to owners may affect incentives in a striking way.

2. The life-cycle model of the firm

Consider the dynamics of a value-maximizing, equity-financed corporation in a continuous-time model with progressive tax on dividends as the only form of taxation. The firm is owned by a household, which is able to borrow and lend at the common market rate of interest $\rho > 0$. The firm produces output with non-depreciating capital K as the only production factor and finances its operations through profits $F(K)$ and issues of new equity Q. All prices are normalized to one. The firm spends its resources on dividends D and investments $I = \dot{K}$. The motion of capital is thus expressed by

$$F(K) + Q = D + \dot{K}.$$

The profit function F is defined for the non-negative real numbers $\mathbb{R}_0^+$, it is twice differentiable and strictly concave on the positive real line $\mathbb{R}^+$ and satisfies the usual Inada conditions. The start-up stock of capital injected at time $t = 0$ is given by $K(0) = K_0 \geq 0$. The flow of new equity is required to be non-negative $Q \geq 0$ to ensure that it is impossible for the firm to pay cash to its owner that is not taxed as dividends. Hence the model is a "trapped equity" model, where dividends are the only means of withdrawing funds from the firm. [6,7]

For simplification, we ignore taxes other than dividend tax and assume a linear-progressive tax scheme where dividends are exempt up to the threshold $E \geq 0$ and taxed at a proportional rate $\tau \in (0, 1)$ on any excess amount. We assume that the tax rate stays constant over time and denote $\theta = 1 - \tau$. We model this tax scheme by writing

$$D = D_1 + D_2,$$

where D_1 is the tax-exempt and D_2 the taxable part of dividends, $0 \leq D_1 \leq E$ and $D_2 \geq 0$. The lower bounds are included to rule out financing through negative dividends. We further constrain dividends by current profits $D \leq F(K)$. This constraint rules out distributions financed from new share issues.[8]

---

[6] See Sinn (1991b) and Lindhe and Södersten (2009) for a discussion of how well the constraint on new equity corresponds to current institutions and how relaxing the constraint affects results.

[7] We also make the following technical assumptions which are necessary for the mathematical analysis: we assume that K is continuous and that $\dot{K}$, D and Q are "almost everywhere continuous", that is, they are continuous on $\mathbb{R}_0^+$ except for at most a finite number of points in any finite subinterval and they are everywhere right-continuous. Differential equations involving these variables should be understood to hold outside the discontinuities.

[8] This constraint is introduced to avoid indifference outcomes in certain cases and hence to simplify exposition. Combined with $Q \geq 0$ it has the potentially undesirable implication that the firm's choices do not include any means of reducing *K,* even if *K* were initially higher than the long-run optimal stock of capital. However, the primitives of the model rule out such situations.

The total net cash flow Y from the firm to the shareholder becomes

$$Y = D_1 + \theta D_2 - Q.$$

The market value of the firm's shares V is determined by the no-arbitrage condition

$$Y + \dot{V} = \rho V,$$

which requires that in equilibrium the shareholder is indifferent between holding shares or exchanging them for bonds yielding ρ. We assume that V is continuous and non-negative on $\mathbb{R}_0^+$ and that $\lim_{t\to\infty} V(t)\mathrm{e}^{-\rho t} = 0$, which ensures that the firm has zero value if it never issues new shares or distributes any dividends. The initial market value of the shares is thus

$$V_0 = V(0) = \int_0^\infty Y(t)\mathrm{e}^{-\rho t}\,\mathrm{d}t.$$

The shareholder's problem is hence to find the policy $P = (K_0, D, Q) \in \mathcal{A}$ which maximizes the net value of the shares

$$W(P) = V_0 - K_0, \qquad (1)$$

where $\mathcal{A}$ is the set of all admissible policies.

Problem (1) generalizes Sinn's (1991a) life-cycle model by introducing a simple progressive dividend tax element. Note that the model contains as special cases the linear tax scheme ($E = 0$) and the trivial model of no taxation ($E = \infty$).

In spite of some aspects which distinguish the model from many neoclassical models (e.g. no adjustment costs and the optimally chosen initial capital stock), the model turns out to be useful for our analysis. As shown by Sinn (1991a), under a single dividend tax, the firm grows slowly towards the steady state, financing investment from retained earnings. This aspect is important since the special incentive effects of progressive taxation will materialize in those very circumstances. In the present framework the incentive to finance internally is produced by a tax penalty on dividends compared to retained earnings. Due to this extra cost the firm prefers retained earnings to external funds, and, after the start-up capital stock is invested, adapts its further investment program to the slowly accruing profit flow. An alternative way to model internal growth would be to assume non-tax costs for external financing as the models by Fazzari et al. (1988) and Bond and Van Reenen (2006) do. However, their approach would increase complexity without any real benefits.

3. Analysis of the basic model

This section provides an outline and discussion of the firm's optimal dynamic policy and illustrates numerically how the size of E affects the firm's growth. We start with the following result, which presents a unique solution to problem (1). We refer to Appendix A for a precise formulation and proof. In the following we denote by K^* the number that satisfies $F'(K^*) = \rho$, that is, K^* is the size of the long-run capital stock in an environment with no taxation and, with reference to Sinn (1991a), also the long-run capital stock in the presence of a single linear dividend tax.

Proposition 1. Problem (1) has an optimal solution $P = (K_0, D, Q)$ with the following properties:

i. If $E \geq F(K^*)$, then P is constant: $K = K_0 = K^*$, $D_1 = F(K^*)$, $D_2 = 0$ and $Q = 0$.

ii. If $E < F(K^*)$, then (a) at the start-up we have $K_0 = \zeta$, (b) there is a growth phase $0 < t < t^*$, where $D_1 = E$, $D_2 = 0$, $Q = 0$ and $\dot{K} = F(K) - E > 0$, (c) in the steady-state phase $t \geq t^*$ we have $D_1 = E$, $D_2 = F(K^*) - E$, $Q = 0$ and $K = K^*$. [9]

The number $\zeta \in (F(E)^{-1}, K^*)$ which appears above in case (ii) is defined by the equation

$$\int_{\zeta}^{K^*} \frac{F'(s) - \rho}{F(s) - E} \mathrm{d}s = \log\frac{1}{\theta}, \qquad (2)$$

where its existence is guaranteed by the strict concavity of F. Moreover,

$$t^* = \int_{\zeta}^{K^*} \frac{\mathrm{d}s}{F(s) - E} = \frac{1}{\rho}\log\left(\theta\frac{F(K^*) - E}{F(\zeta) - E}\right). \qquad (3)$$

We will start the discussion of Proposition 1 by first defining two special cases in the model, which serve as benchmarks for the analysis. The first case is no taxation ($E = \infty$) and the second case is linear tax ($E = 0$). When no taxes are levied on distributions, the firm's optimal investment policy follows the simple rule $K_0 = K^*$. The firm invests its capital stock at the start-up stage to the level that is optimal in the long run. When instead a linear dividend tax is levied, the

[9] Moreover, it may be worthwhile to note that in both cases the optimal value is $W(P) = \frac{F(K_0)}{\rho} - K_0$.

optimal initial stock of capital is lower than the long-run stock of capital, $K_0 < K^*$. The start-up phase is followed by a phase of internal growth where the firm invests using retained profits as the source of finance. No new equity is issued, $Q = 0$, and no dividends are distributed, $D = 0$. In the long-run equilibrium the firm's capital stock satisfies $K = K^*$ and the firm spends all profits on dividends.(Sinn 1991a)

Why slow growth, and why financing from retained earnings, when external financing is also available? The answer is that dividend tax produces a tax penalty on external equity. It reduces the opportunity cost of retained profits but leaves the cost of new equity issues unchanged. Therefore the firm uses new equity only to finance a small start-up capital stock to get its operations going. The rest of the long-run stock of capital $K*$ is financed from retained earnings. Dividend taxation has no effects in the long run since it reduces the opportunity cost of profits retained for investment in the same proportion as it reduces future returns, with the result that the effect of taxation is canceled out. (Sinn, 1991a, 1991b)

Let us now return to behavior under non-linear tax. The optimal policy is divided into two distinct cases depending on the size of E. If $E \geq F(K^*)$, the firm behaves exactly as in a world with no dividend taxation. The capital stock is set at the steady-state level from the start-up $K_0 = K^*$, using external equity as the source of finance. Tax parameters have no effects because the forward-looking firm understands that taxes on dividends are never paid.

In the more interesting case where $E < F(K^*)$, the start-up capital stock $K_0 = \zeta$ is strictly lower than the steady-state stock K^* (but $F(K_0) > E$). Therefore, the start-up stage is followed by an internal growth phase, where the firm finances investments from retained profits. This resembles optimal policy in linear taxation. Now, however, profits are spent on dividends up to the threshold E and only the remaining part is retained and spent on investments. Hence non-linear tax generates an incentive to start distributions early (EDI). Once K reaches its steady-state size K^*, the firm starts to fully pay out profits as dividends. The steady-state capital stock is undistorted, which implies that non-linear elements in the tax rate do not undermine the new-view neutrality result.

The solution implies that during the growth phase a share of the firm's profit is distributed even though the marginal return on investment is higher than the market interest rate, $F'(K) > \rho$, which the owner could earn in the financial markets. Hence the firm gives up investments with a high return in exchange for distributions with a lower return when reinvested. This inefficiency aspect is non-existent in linear taxation.

To give intuition to this outcome, let us consider how non-linear tax affects the sources and uses of funds. The firm's cost of finance depends on whether the growing firm finances investment by cutting taxable or tax-exempt dividends. If the source is the former, the cost of financing 1 unit of investment is $\theta < 1$, and if the source is tax-free dividends, the cost equals 1. The optimal use of profits can now be understood by comparing these costs to the after-tax value of the investment. According to the optimality conditions, the marginal value of investment λ satisfies $\theta < \lambda < 1$ in the growth phase (see Appendix A). Hence the value of the investment is higher than the cost of taxable dividends but lower than the cost of tax-exempt dividends. The firm's choice is therefore to distribute profits up to *E* and spend the rest on investments.

In sum, the forward-looking firm takes full advantage of the tax exemption available at every point of time. This reflects the incentive to even out distributions over time under a progressive tax. The consequence is that distributions begin right after the firm has started to operate. Dividends are preferred to investments even if the before-tax marginal return on investments is higher than their return outside the firm. The results provide some testable predictions, among them the following: in a cross-section of heterogeneous firms with different ages and different productivities, a share of all firms set their dividends exactly at $D = E$.

The optimal size of the start-up capital stock ζ and the duration of the growth path t^* depend on the tax parameters as follows:

$$\frac{\partial \zeta}{\partial E} = \frac{F(\zeta) - E}{F'(\zeta) - \rho} \int_{\zeta}^{K^*} \frac{F'(s) - \rho}{(F(s) - E)^2} \mathrm{d}s > 0, \tag{4}$$

$$\frac{\partial t^*}{\partial E} = \frac{1}{F'(\zeta) - \rho} \int_{\zeta}^{K^*} \frac{F'(\zeta) - F'(s)}{(F(s) - E)^2} \mathrm{d}s > 0, \tag{5}$$

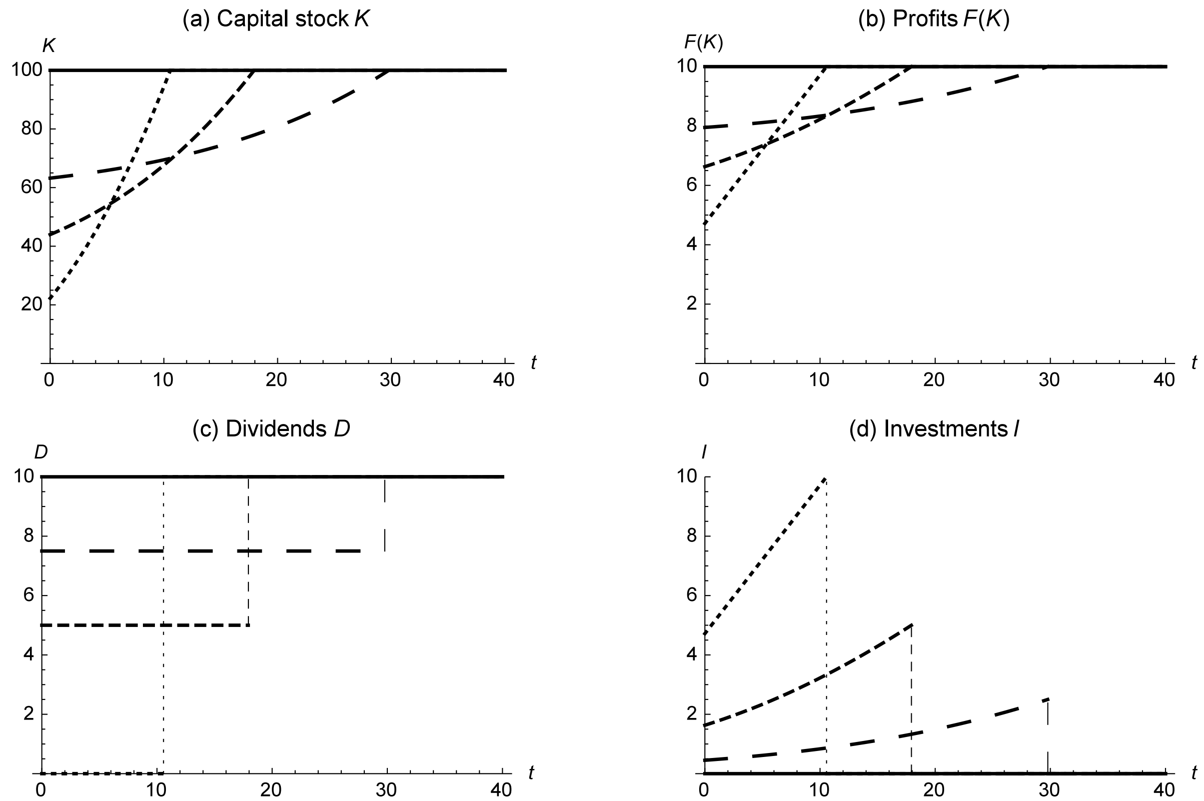
$$\frac{\partial \zeta}{\partial \tau} = \frac{F(\zeta) - E}{\theta(F'(\zeta) - \rho)} < 0,$$

$$\frac{\partial t^*}{\partial \tau} = \frac{1}{\theta(F'(\zeta) - \rho)} > 0.$$

Hence the size of the start-up capital stock ζ depends positively on E but negatively on τ, and the length of the growth path t^* depends postively on both E and τ.

Consider the intuition of (4) and (5). During the growth phase the capital stock develops according to $\dot{K} = F(K) - E$, implying that at a given level of K, the EDI induced by the tax threshold *E* reduces investment and slows down growth. This effect is reflected in equation (5), which says that with a higher threshold it takes longer to attain the steady state. The total effect on the firm's growth, however, is more complex since the threshold turns out to affect not only the pace of investment, but also the size of the initial capital stock ζ. As equation (4) shows, the latter effect is positive. In other words, a rise in E increases start-up investment, ζ, and hence reduces the production delay, but increases the delay via a longer duration of the growth phase. To illustrate the predicted impacts, Figure 1 depicts the patterns of K, $F(K)$, D and I for four values of the tax threshold E.

Figure 1. *(a) Optimal capital stock K, (b) profits $F(K)$, (c) dividends D, and (d) investments I as functions of time for the threshold values $E = 0$ (dotted), $E = 5$ (short dashing), $E = 7.5$ (long dashing) and $E = 10$ (solid).*



Here we have chosen $F(K) = K^{0.5}$, $\rho = 0.05$ and $\tau = 0.2$. Hence from $F'(K^*) = \rho$, we calculate that $K^* = 100$ and $F(K^*) = 10$. The case $E = 0$ corresponds to linear taxation. If $E \geq F(K^*) = 10$, then the marginal tax rate on dividends is zero and the effect of dividend taxation vanishes. Panel (a) shows the positive relationship between E and both K_0 and t^*. We observe that both variables are quite sensitive to the size of E. For example, at $E = 7.5$ the growth path is nearly three times as long as in the case of $E = 0$. Comparing the size of the initial capital stock K_0 in these two cases, the factor is of similar size. To interpret the patterns in panels (b)-(d), recall that the profits $F(K)$ of a growing firm are spent on dividends $D = E$ and investments I, which implies $I = F(K) -$

E. In the steady-state phase $D = F(K)$ and $I = 0$. According to panels (b)-(d), the levels of $F(K)$, D and I are also quite sensitive to the size of E.

Tax schedules differ across countries as regards the number of tax brackets and the level of tax rates.[10] Therefore it may be worthwhile to discuss the implications of some modifications to our tax schedule. Assume first that dividends are taxed at rate τ_1 up to E and at rate τ_2 on the excess amount, with the following properties $\tau_2 \geq \tau_1 > 0$. All other assumptions are as in the basic model.

The solution to this model (in the case $E < F(K^*)$) has many similarities to our basic model with $\tau_1 = 0$. The most important is that the incentive to start distributions early (EDI) still emerges. The main difference is that now there are two distinct growth phases. The firm starts its internal growth in a regime where no dividends are distributed and all accruing funds are spent on investment. Nonetheless, the second growth phase is qualitatively similar to the growth phase of the basic model: the firm distributes up to the threshold, $D = E$, and invests the remainder. This implies that the speed of growth is reduced by the incentive to distribute early. Analogously it is possible to extend the model to include $N > 2$ tax brackets with positive tax rates which would induce up to N growth regimes with different growth properties. We conclude that, while the form and magnitude of the effects may vary, the existence of EDI is not contingent on the shape of the non-linear tax schedule, in particular on our linear-progressive scheme with a zero starting rate.[11]

[10] For example, under the Swedish tax rules, dividends from a closely held company are taxed at a reduced rate of 20% up to a threshold, and at a much higher rate on any exceeding amount.

[11] The solution to the more general model is outlined in a memo available on request.

4. Total effect on production

The "nucleus model" of Sinn (1991a) implies that dividend taxation slows the growth of new firms towards long-run equilibrium, leading to delayed production activity compared to its potential. In our model the tax threshold affects the production delay in two ways: first, through the size of the initial capital stock and, second, through the rate of investment during the internal growth phase.

To further analyze the relative importance of these effects, we introduce two indicators for measuring the value of the production delay. Consider the optimal solution P to problem (1) as a function of E and τ and write $P = (K[E,\tau], D[E,\tau], Q[E,\tau])$. The present value $\pi(E,\tau)$ of the production flow is then

$$\pi(E,\tau) = \int_0^\infty F\big(K[E,\tau](t)\big)\mathrm{e}^{-\rho t}\,\mathrm{d}t.$$

Tax revenue $T_b(E,\tau)$ collected from the representative firm's dividends is

$$T_b(E,\tau) = \int_0^\infty \tau D_2[E,\tau](t)\mathrm{e}^{-\rho t}\,\mathrm{d}t = \frac{\tau}{\rho\theta}\big(F(K_0[E,\tau]) - E\big),$$

where the right-hand identity follows from equation (3) and Proposition 1.

The first indicator μ_N compares the production flow under a non-linear dividend tax to production in a world with no dividend tax. It is calculated as the ratio of the present values of distorted and non-distorted production flows. The second indicator μ_R aims to illustrate the effect on production of a revenue-neutral tax reform where a non-linear tax is substituted for a linear one while keeping the tax revenues unaffected. It is calculated as a ratio of the present values of production flows under non-linear and linear tax with the same tax proceeds. Tax revenue is balanced by adjusting the tax rate of the linear dividend tax. For fixed τ, the indicators μ_N and μ_R are defined as follows:

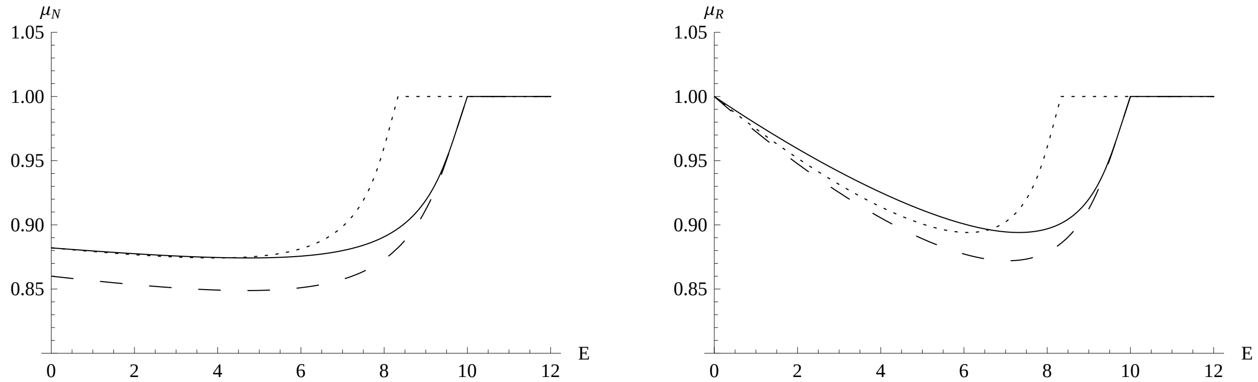
$$\mu_N(E) = \frac{\pi(E,\tau)}{\pi(F(K^*),\tau)}, \qquad \mu_R(E) = \frac{\pi(E,\tau)}{\pi(0,\tau_R[E,\tau])},$$

where $\tau_R = \tau_R[\tau, E]$ is determined by $T_b(E,\tau) = T_b(0,\tau_R)$.

The indicator μ_N takes a value between 0 and 1 with $\mu_N = 1$ if and only if K is identically K^* on Ω. Hence the lower the value of μ_N, the larger the distortion caused by E. For the revenue-neutral measure μ_R, we clearly have $\mu_R = 1$ for $E = 0$ and $E \geq F(K^*)$. In fact, in the former case no tax reform takes place and in the latter case the firm is "too small" to be affected by such a reform (see

Proposition 1). For $0 < E < F(K^*)$ the value of μ_R represents the effects on production delay of a revenue-neutral reform.

Figure 2. *Measures μ_N and μ_R as functions of E for parameter values $\tau = 0.2$ and $\rho = 0.05$ (solid lines), $\tau = 0.25$ and $\rho = 0.05$ (dashed lines) as well as $\tau = 0.2$ and $\rho = 0.06$ (dotted lines).*



In Figure 2, the measures μ_N and μ_R are illustrated as functions of E for the production function $F(K) = K^{0.5}$. The figure presents graphs for different choices of τ and ρ for comparison. We observe that the normalized measure μ_N first slowly declines with E and then, near the steady state, quickly grows and approaches one. Hence at low levels of E the distortion grows with E but is not very sensitive, implying that the effect of the start-up capital stock nearly balances the effect of the pace of investments. At higher levels of E the positive effect on the initial capital dominates, which reduces the overall distortion. Figure 1 helps to understand the relative importance of the two opposing effects.

The revenue-neutral measure μ_R has a U-shaped slope, showing that a tax reform which introduces an exemption limit to the dividend tax schedule always leads to an increased production lag. The size of the production loss is considerable. The intuition of the result is that a higher tax rate on dividends, required by the balanced budget rule, adversely affects the start-up capital stock and increases the production loss. This counteracts the balancing effect resulting from an increase in the initial capital stock, and the effect on production resulting from an increased length of the growth phase remains dominant. Hence, while the production loss caused by dividend taxation is not very sensitive to the size of E in absolute terms, a revenue-neutral tax reform where a non-linear tax is substituted for a linear one increases the production loss for all relevant values of E.

5. Non-linear dividend tax with carry-forward element

5.1 Extended model of the firm

The analysis in the previous sections shows that non-linearity in the dividend tax scheme affects the firm's behavior in two ways: it both increases the start-up capital stock and reduces investments during the internal growth phase. The latter effect is a consequence of an incentive to even out dividend payments over the firm's life cycle. It seems to be related to the static feature of the linear progressive tax scheme. The scheme is applied to periodic income with no attempts to take into account the effects of progressivity when income fluctuates over time. As pointed out in the Introduction, some recent tax policy experiments include elements which aim at cushioning such effects of progressivity. One example is Norway's taxation of income from shares, where any unused exemption for the so-called normal return on savings is carried forward with interest for future years. A similar element is in use in Sweden's taxation of dividend income received from closely held corporations as well as in the RRA model proposed by Mirrlees et al. (2011) for the taxation of risky income.

In 2006, Sweden introduced a reduced tax rate of 20% on dividends from closely held companies compared to the standard rate of 30% on capital income. Dividends subject to this rate were constrained by a normal return on new equity (general rule) or, as an optional new element, by a fixed amount of SEK 143,275 (2012; simplification rule). As already before the reform, any dividends exceeding the ceiling were taxed as labor income subject to the progressive rate. Unused allowances were carried forward with interest irrespective of the type of the allowance chosen by the owner. (Alstadsæter and Jacob, 2012)

The aim of this section is to analyze how introducing carry-forward of unused allowances affects the distortions produced by progressive elements in the tax scheme. We will proceed by augmenting the fixed allowance of the basic model with carry-forward. As a result, the tax system that we will focus on has many similarities to the Swedish simplification rule.

Earlier attempts to analyze the effects of carry-forward include Alstadsæter and Fjærli (2009), who demonstrate the neutrality of the Norwegian application with respect to the timing of dividends in a two-period model with constant returns, and Lindhe et al. (2002), who analyze Swedish dividend tax rules for closely held corporations in a dynamic investment model. The latter study focuses on the long-run cost of capital, and does not detail the potential effects of carry-forward of unused allowances. In particular, neither study considers the effects of carry-forward on growth.

As before, let E denote a tax threshold and write $D = D_1 + D_2$, where D_1 is the tax-exempt and D_2 the taxable part of dividends which are both assumed to be non-negative and satisfy $D \leq F(K)$. The tax threshold is strictly positive, $E > 0$. In the modified tax system, unused exemptions may be carried forward with interest. To illustrate how this element can be modeled in our framework, let us use a simple example in discrete time: let R_t denote the amount of accumulated unused tax exemptions at the end of period t and i the rate of interest applied to compound the stock. Tax-free dividends $D_{1,t}$ are limited to the sum of the periodic exemption E and the amount of unused allowances from previous years compounded by interest. Hence the upper constraint on $D_{1,t}$ is

$$D_{1,t} \leq E + (1 + i)R_{t-1},$$

and the reserve R_t develops over time as

$$R_t = (1 + i)R_{t-1} + E - D_{1,t}.$$

The value of the reserve at the end of period t is then the value at the end of the preceding period compounded by interest $i > 0$ plus the periodic exemption E minus dividends distributed during the period. By combining these two constraints, the limit for $D_{1,t}$ reduces to

$$R_t \geq 0.$$ [12]

To include the carry-forward in our continuous-time framework, we simply modify the basic model (1) by replacing the limit $D_1 \leq E$ by the following constraints:

$$R \geq 0, \tag{6}$$

$$R(0) = 0, \tag{7}$$

$$\dot{R} = E - D_1 + iR, \tag{8}$$

where R is a continuous function on $\mathbb{R}_0^+$.

As before, the shareholder's problem is to find the policy $P = (K_0, D, Q) \in \mathcal{A}'$ which maximizes

$$W(P) = V_0 - K_0, \tag{9}$$

[12] Lindhe et al. (2002) end up with the same formulation in their discrete-time dynamic model.

where $\mathcal{A}'$ is the set of all admissible policies $P = (K_0, D, Q)$ for the carry-forward model. A precise formulation of this problem is contained in Appendix B. Although the objective function of problem (9) is the same as for the basic model (1), the set of admissible policies is larger in the carry-forward problem.

Observe that using (7) and (8) the constraint (6) can be rewritten as

$$\int_0^t (E - D_1(s))\mathrm{e}^{-\rho s}\,\mathrm{d}s \geq 0, \tag{10}$$

for all t. The present value of the flow of dividends D_1 over interval $[0, t]$ is constrained by the present value of the exemption flow. This means that D_1 may well exceed E temporarily or over a finite period but in the long run D_1 must stay in the limits of E. Hence the carry-forward element transforms the original static constraint to a life-cycle constraint. This obviously gives the firm much new flexibility when it plans its investment and dividend policies. Note the obvious limitation, however, that the exemption can be saved for later years but not used in advance.

5.2 Optimal behavior in the extended model

In this section we describe and discuss the firm's optimal policy in the extended model. Proposition 2 below presents a solution to problem (9) which we have restricted to the case $i = \rho$ for simplicity.

To state this result we require some new notation: Let the constants ζ_0 and t_0^* refer to the quantities defined in (2) and (3) in the case of linear taxation, that is, the initial capital stock and the length of the growth phase for $E = 0$. We conclude from (3) that $F(\zeta_0)/\theta < F(K^*)$, which implies a lower bound $\tau F(K^*)$ for the distance that $F(K)$ grows under linear tax. In the case $F(\zeta_0)/\theta < E < F(K^*)$, we define $\tilde{\zeta}_E \in (\zeta_0, F^{-1}(E))$ and $\tilde{t}_E^* \in (0, t_0^*)$ to specify the size of the initial capital stock and the length of the growth phase by

$$\int_{\tilde{\zeta}_E}^{K^*} \frac{\mathrm{d}s}{F(s)} = \frac{1}{\rho}\log\frac{F(K^*)}{E},$$

$$\tilde{t}_E^* = \frac{1}{\rho}\log\frac{F(K^*)}{E},$$

where the existence again follows from the strict concavity of F.

Proposition 2. Assume that $i = \rho$. Problem (9) has the unique solution $P = (K, D, Q)$ with the following properties:

i. If $E \geq F(K^*)$, then P is constant as in Proposition 1 (i).

ii. If $F(\zeta_0)/\theta < E < F(K^*)$, then (a) $K_0 = \tilde{\zeta}_E$, (b) during the growth phase $t \leq \tilde{t}_E^*$, we have $D_1 = D_2 = Q = 0$ and $\dot{K} = F(K) > 0$ and (c) during the steady-state phase $t > \tilde{t}_E^*$ we have $K = K^*$, $Q = 0$, $D_1 = F(K^*)$ and $D_2 = 0$.

iii. If $E \leq F(\zeta_0)/\theta$, then (a) $K_0 = \zeta_0$, (b) during the growth phase $t \leq t_0^*$ we have $D_1 = D_2 = Q = 0$ and $\dot{K} = F(K) > 0$ and (c) during the steady-state phase $t > t_0^*$ we have $K = K^*$, $Q = 0$ and D_1 and D_2 satisfy the non-unique conditions (10), $D = F(K^*)$ and

$$\int_0^\infty (E - D_1(s))\mathrm{e}^{-\rho s}\,\mathrm{d}s = 0. \qquad (11)$$

The solution to Problem (9) divides into three distinct cases with clearly different behavioral implications. In the first case, which holds if the allowance is high compared to profit, $E \geq F(K^*)$, the firm behaves as if there were no taxation at all. It is equivalent to the corresponding case of the basic model (Proposition 1 (i)). If instead $E < F(K^*)$, the firm's behavior divides into two distinct patterns with the common aspect that no early distribution incentive (EDI) exists. .

If $E \leq F(\zeta_0)/\theta$ (primary case), the firm's capital stock evolves exactly as in linear taxation. The firm starts with the same optimal initial stock of capital $K_0 = \zeta_0 < K^*$ and grows, investing all profits until its capital stock attains its steady-state size K^*. No EDI exists. Hence introducing carry-forward removes all the distorting elements produced by progressivity in the tax rate. Unlike in the basic model but as in linear taxation, both the size of the start-up capital stock and the length of the growth path are independent of the size of the exemption E:

$$\frac{\partial \zeta_0}{\partial E} = 0, \qquad \frac{\partial t_0^*}{\partial E} = 0.$$

This follows from the similarity to the model with linear taxation, where $E = 0$.

If $F(\zeta_0)/\theta < E < F(K^*)$ (intermediate case), the carry-forward model produces a solution which has some similarities with the case of linear tax but is not identical. As under linear tax, there is a phase of internal growth with no dividend distributions. There are three differences however: first, the initial capital stock is higher than under linear tax $K_0 = \tilde{\zeta}_E > \zeta_0$. Second, the firm pays out its entire profit in the steady state but now it only distributes tax-free dividends with the result that tax is never paid on dividends. This policy is facilitated by the reserve of unused allowances that the firm accumulates during its growth phase: $D = F(K^*) = E + iR - \dot{R} > E$. To maintain this dividend policy the firm must collect an appropriate target level of unused allowances, $R^T = R(\tilde{t}_E^*) = (F(K^*) - E)/\rho$. This requires a correct balance between K_0 and the amount invested from retained profits during the growth path.

The third distinction from linear tax is that both the initial capital stock and the length of the growth path depend on the size of E.

$$\frac{\partial \tilde{\zeta}_E}{\partial E} = \frac{F(\tilde{\zeta}_E)}{\rho E} > \frac{\theta}{\rho} > 0, \tag{12}$$

$$\frac{\partial \tilde{t}_E^*}{\partial E} = -\frac{1}{\rho E} < 0. \tag{13}$$

As in the basic model, $\tilde{\zeta}_E$ increases in E, but, in contrast to the basic model, now $\tilde{t}_E^*$ decreases with E. Why does this happen? The outcome stems from two aspects. First, during the growth phase no dividends are distributed. Therefore, investment at a given level of K is independent of the size of E. The only channel through which E may affect growth is the size of the initial capital stock $\tilde{\zeta}_E$. The stock increases with E and therefore the length of the growth path $\tilde{t}_E^*$ decreases.

To sum up, we observe that, if $i = \rho$, carry-forward abolishes EDI and the resulting production delay. There are, however, two slightly different solutions. While the primary case is fully equal to the firm's behavior under linear tax, the intermediary case has some slight differences. In both there is a phase of internal growth where the firm finances investments from retained earnings and distributes no dividends. The regimes differ in terms of how the steady-state capital stock is financed and whether taxes are paid in the steady-state phase or not: firms in the intermediate regime have a larger start-up capital stock and their owners pay no taxes while owners of primary-regime firms do pay. Overall, the results for the present model imply that carry-forward abolishes any bunching of firms around the tax threshold predicted by the basic model. This prediction could be tested using data in future research.

To provide some intuition for how firms might divide into the two regimes of case $E \leq F(\zeta_0)/\theta$, we calculate the widths of the feasibility bands using the numerical example of Section 3. Recall that the parameter values used there ($F(K) = K^{0.5}$, $\rho = 0.05$, $\tau = 0.2$) provide a steady-state capital stock of $K^* = 100$ and a steady-state profit of $F(K^*) = 10$. Under this specification, $F(\zeta_0)/\theta = 8.29$, which implies that the primary case is valid when $0 < E \leq 8.29$ and the intermediary case when $8.29 < E \leq 10$. If, on the other hand $E > 10$, case (i) of Proposition 2 applies and taxes have no effect on behavior. We conclude that the intermediary case holds for a fairly narrow range of values of E, but it is not just a razor's edge case, and hence has some practical importance.

Finally, how does carry-forward affect tax revenue? It can be shown that for a given E with $E < F(K^*)$,

$$T_b(E) > T_{cf}(E),$$

where b refers to non-linear tax without and cf with carry-forward. Hence carry-forward always reduces tax revenue. This is most obvious for firms in the intermediate regime, where no taxes are paid under carry-forward, but also applies to the primary regime.

6. Discussion

The large theoretical literature on the effects of dividend taxes on the behavior of firms has usually assumed linear tax rates, and is therefore in clear contradiction with real-life tax systems, which commonly impose progressive tax rates on capital income. This paper addresses the question of how some of the central results of the standard literature change if we introduce a progressive tax scheme. It does this by analyzing a linear-progressive tax schedule in Sinn's (1991a) growth model of an all-equity firm. The paper also analyzes how tax rules where unused exemptions can be carried forward affect the distortions produced by non-linear tax. Allowing carry-forward of unused allowances was proposed by Mirrlees et al. (2011) and it has been implemented in some Nordic countries.

Our analysis predicts that a pure non-linear tax on dividends produces an incentive for early distributions (EDI). Under this tax the owners prefer a smooth income flow, which induces firms to start distributions right from the beginning. This leads to delayed production, particularly when compared to a linear tax with the same tax proceeds. If we equip the non-linear tax scheme with carry-forward of unused exemptions, EDI may disappear even entirely and the firm's policy may then follow the same pattern as in proportional taxation. The steady-state capital stock is not affected by dividend tax rules in either model. Hence the new-view neutrality result, well known from previous literature, is not overridden by progressive elements in the tax schedule.

Allowing carry-forward of unused allowances makes the firm's decisions much more flexible. According to the tax rules, the stock of unused allowances is compounded using a rate of interest corresponding to the firm's discount rate. This implies that the use of tax allowances can be deferred to coming years without any loss in present value. Hence allowing carry-forward with interest translates the original annual ceiling for tax-free dividends to a life-cycle constraint under which the timing of dividend distributions does not matter anymore. The firm can focus on growth by using all its internal revenue on investment and postpone dividend distribution to the steady state where no profitable investment opportunities are left.

Thus the analysis of our basic model unveils a new form of inefficiency caused by taxation of corporate- source income. Progressivity distorts the timing of dividends and investments by a financially constrained firm. This distortion can be avoided, however, by allowing carry-forward of unused allowances.

As discussed in Section 3, the main results of this paper can be easily generalized to cover dividend tax schemes consisting of several tax bands and hence in practice all possible progressive tax schemes. However, since a tax scheme with a single tax threshold already reveals essential differences compared to linear tax systems, in particular EDI, the formulation of such generalizations remains outside the scope of this paper.

The main policy contribution of our model is that, to avoid timing effects, sharp jumps in personal tax rates on dividends and other forms of corporate source income should be avoided. Similarly, a high ceiling for lump-sum exemptions is worse than a low ceiling, since it leads to a relatively higher amount of dividends brought forward and to a greater reduction in investment. Furthermore, if political priorities require progressivity, the tax schedule should be equipped with carry-forward of unused allowances.

The model is simple and lacks many important real-life aspects. If we introduce several income sources and if we let the tax system mimic comprehensive income tax, it will be less likely (but not impossible) that the exemption limit will affect the owner's choices. And if there are several heterogeneous owners with different amounts of income from varying sources, the probability may decrease further. Furthermore, when we add in several ways of rewarding the owner and when these income types are subject to different tax rates, new tax-planning opportunities may emerge. These alternative means may crowd out the timing effect found in our model. Hence, there are good grounds to expect that in a richer set up, the EDI incentive and its effects on optimal policy would not be as emphatic as in our analysis.

However, our analysis reveals some incentive aspects of progressive taxation that are also important in practice. A potential case where the results could well be of practical importance is a progressive tax on income from closely held companies. We have in mind a case where a small number of fairly similar owners derive most of their income from their company. The case is strengthened if a separate progressive tax schedule is applied to distributed profits. In this case the diluting effect of other income sources may remain small. Such rules exist in the Nordic countries but separate rate schedules appear elsewhere as well. The US and Spanish dividend taxes are two examples.

The extent to which the predictions of our model translate into behavioral changes should be explored using empirical data. Future research could use the reforms in Finland and Sweden in 2005 and 2006 as natural experiments. Both reforms introduced a non-linear tax schedule for dividend income from closely held companies; in Sweden the scheme was equipped with carry-forward while in Finland it was not. Our analysis predicts that without carry-forward, non-linear tax induces firms to bunch around the dividend tax threshold. There is indeed some evidence of bunching below the exemption limit in Finland (Kari and Karikallio, 2007; Harju and Matikka, 2013). Similarly, in line with our predictions, Alstadsæter and Jacob (2013) report no bunching around the Swedish exemption limits. Future research could try to disentangle the effects produced by the tax structures analyzed in this paper from other potential factors.

References

Alstadsæter, A. and Fjærli, E. (2009), Neutral taxation of shareholder income? Corporate responses to an announced dividend tax, International Tax and Public Finance 16, 571-604.

Alstadsæter, A. and Jacob, M. (2012), Income Shifting in Sweden. An Empirical Evaluation of the 3:12 Rules, Report to the expert group on public economics 2012:4, Ministry of Finance, Sweden.

Alstadsæter, A. and Jacob, M. (2013), Payout Policies of Privately Held Firm: Flexibility and the Role of Income Taxes, FAccT Center WP No. 12/2013, WHU - Otto Beisheim School of Management.

Alstadsæter, A. and Jacob, M. (2014), Do Dividend Taxes Affect Corporate Investment, CESifo Working Paper No. 4931.

Auerbach, A. J. (1979), Wealth maximization and the cost of capital, *Quarterly Journal of Economics 94,* 433-436.

Auerbach, A. J. and Hassett, K. A. (2007), The 2003 dividend tax cuts and the value of the firm: An event study, in: Auerbach, A., Hines, J.R. & Slemrod, J. (Eds.), *Taxing Corporate Income in the 21st Century,* Cambridge University Press.

Becker, B., Jacob, M. and Jacob, M. (2012), Payout taxes and allocation for investment, Journal of Financial Economics 107, 1-24.

Bond, S., Devereux, M. P. and Klemm, A. (2007), The effects of dividend taxes on equity prices: A re-examination of the 1997 UK tax reform, Oxford University Centre for Business Taxation WP 07/01.

Bond, S. and Van Reenen, J. (2005), Microeconometric models of Investment and Employment, in Heckman & Leamer (Eds.), Handbook of Econometrics, edition 1, volume 6, chapter 65, Elsevier.

Brown, J. R., Liang, N. and Weisbrenner, S. (2007), Executive financial incentives and payout policy: Firm responses to the 2003 dividend tax cut, Journal of Finance 62, 1935-65.

Bradford, D. F. (1981), The incidence and allocation effects of a tax on corporate distributions, Journal of Public Economics 15, 1-22.

Chetty, R. and Saez, E. (2005), Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut, The Quarterly Journal of Economics 120, 791-833.

Fazzari, S. M., Hubbard, R. G. and Petersen, B. C. (1988), Financing Constraints and Corporate Investment, Brookings Papers on Economic Activity 19, 141-206.

Gordon, R. and Dietz, M. (2008), Dividends and taxes, in: Auerbach, A.J. & Shaviro D.N. (eds.), *Institutional Foundations of Public Finance: Economic and Legal Perspectives,* Harvard University Press, Cambridge, MA.

Gourio, F. and Miao, J. (2010), Firm heterogeneity and the long-run effects of dividend tax reform, American Economic Journal: Macroeconomics 2, 131-168.

Gourio, F. and Miao, J. (2011), Transitional dynamics of dividend and capital gains tax cuts, Review of economic dynamics 14, 368-383.

Harberger, A. C. (1966), Efficiency effects of taxes on income from capital, in: Krzyzaniak, M. (Ed.), Effects of Corporation Income Tax, Wayne State University Press, 107-117.

Harju, J. and Matikka, T. (2013), Entrepreneurs and income shifting: empirical evidence from a Finnish tax reform, VATT Working Papers 43.

Jacob, M. and Jacob, M. (2012), Taxation, dividends and share repurchases: Taking evidence global, Journal of Financial and Quantitative Analysis 48, 1241–1269.

Kari, S. and Karikallio, H. (2007), Tax Treatment of Dividends and Capital Gains and the Dividend Decision under Dual Income Tax, International Tax and Public Finance 14, 427-456.

Kari, S., Karikallio, H. and Pirttilä, J. (2008), Anticipating tax changes: Evidence from the Finnish corporate income tax reform of 2005, *Fiscal Studies 29*, 167-196.

King, M.A. (1974), Taxation and the cost of capital, *Review of Economic Studies 41*, 21-35.

Korinek, A. and Stiglitz, J. E. (2009), Dividend taxation and intertemporal tax arbitrage, *Journal of Public Economics 93*, 142-159.

Lindhe, T., Södersten, J. and Öberg, A. (2002), Economic effects of taxing closed corporations under a dual income tax, ifo Studien 48, 575-610.

Lindhe, T. and Södersten, J. (2009), *Dividend taxation, share repurchases and the equity trap.* CESifo working paper No. 2652.

Lindhe, T. and Södersten, J. (2012), The Norwegian shareholder tax reconsidered, International Tax and Public Finance, 19, 424-441.

Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G. and Poterba, J. (eds.) (2011), Tax by Design: The

Mirrlees Review, Oxford: Oxford University Press for Institute for Fiscal Studies.

OECD (2014), OECD Tax Database. http://www.oecd.org/ctp/tax-policy/tax-database.htm

Poterba, J. M. and Summers, L. H. (1985), The economic effects of dividend taxation. in Altman, E. and Subrahmanyam, M. (eds.), Recent Advances in Corporate Finance, Irwin Publishers, 227-284.

Seierstad, A. and Sydsæter, K. (1987), *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam.

Sinn, H.-W. (1991a), The vanishing Harberger triangle, *Journal of Public Economics 45*, 271-300.

Sinn, H.-W. (1991b), Taxation and the cost of capital: The "old" view, the "new" view, and another view, Tax Policy and the Economy 5, 25-54.

Sørensen, P. B. (2005), Dual Income Tax: Why and How, FinanzArchiv / Public Finance Analysis 61, 559-586.

Weichenrieder, A. (1996), Transfer Pricing, Double Taxation, and the Cost of Capital, Scandinavian Journal of Economics 98, 445-452.

Yagan, D. (2014), Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut, Oxford University Centre for Business Taxation WP 13/22.

Appendix

Proof of Proposition 1

For convenience, we reformulate problem (2) and include the necessary mathematical details: the aim is to find $P \in \mathcal{A}$ which maximizes

$$W(P) = \int_0^\infty (D_1 + \theta D_2 - Q)(t)\,\mathrm{d}\nu(t) - K_0,$$

where we have abbreviated $\mathrm{d}\nu(t) = \mathrm{e}^{-\rho t}\mathrm{d}t$ and a solution $P = (K_0, D, Q)$ belongs to $\mathcal{A}$, the set of admissible policies, if the following properties hold: $K_0 \geq 0$, $D = D_1 + D_2$, the functions D_1, D_2 and Q are almost everywhere continuous on $\mathbb{R}_0^+$ and belong to $L^1(\mathbb{R}_0^+, \nu)$, and there is a continuous and almost everywhere differentiable function $K \in L^1(\mathbb{R}_0^+, \nu)$ such that $F(K) \in L^1(\mathbb{R}_0^+, \nu)$,

$$F(K) + Q = D + \dot{K} \quad \text{on} \quad \Omega,$$

$$K(0) = K_0,$$

$$0 \leq D_1 \leq E, D_2 \geq 0, D \leq F(K), Q \geq 0, K \geq 0,$$

where Ω is the subset of $\mathbb{R}_0^+$ where D_1, D_2 and Q are continuous. Here $L^1(\mathbb{R}_0^+, \nu)$ is the set of integrable functions on $\mathbb{R}_0^+$ with respect to the measure ν. We further assume that D is not identically zero on $\mathbb{R}_0^+$. This problem is an infinite-horizon continuous-time optimal control problem, which we solve by applying standard techniques (see, for example, Seierstad and Sydsæter, 1987).

We first prove part (i) of Proposition 1. Assume that $E \geq F(K^*)$ and let the policy $P = (K_0, D, Q)$ be as in Proposition 1 (i). Then P is clearly admissible. We first show P is maximal, that is, $W(P) \geq W(\tilde{P})$, for any other policy $\tilde{P} = (\tilde{K}_0, \tilde{D}, \tilde{Q}) \in \mathcal{A}$. By using the properties of P, we have

$$\begin{aligned} W(P) - W(\tilde{P}) + K_0 - \tilde{K}_0 &= \int_0^\infty \left(D_1 - \tilde{D}_1 + \theta(D_2 - \tilde{D}_2) - Q + \tilde{Q}\right)\mathrm{d}\nu \\ &= \tau \int_0^\infty \tilde{D}_2\,\mathrm{d}\nu + \int_0^\infty \left(F(K^*) - F(\tilde{K}) + \dot{\tilde{K}}\right)\mathrm{d}\nu =: I_1 + I_2. \end{aligned}$$

Since $\tilde{D}_2 \geq 0$, we have $I_1 \geq 0$. Using the formula for $\dot{K}$ and the properties of F, it is easy to verify that $\tilde{K}(t)\mathrm{e}^{-\rho t/2}$ is bounded from above for all t. Hence $\tilde{K}(t)\mathrm{e}^{-\rho t} \to 0$ as $t \to \infty$. Moreover, by the concavity of F, we have $F(K^*) - F(\tilde{K}) \geq F'(K^*)(K^* - \tilde{K})$ on $\mathbb{R}^+$, so that

$$I_2 \geq \int_0^\infty \left(\rho\left(K^* - \widetilde{K}\right) + \dot{\widetilde{K}}\right) \mathrm{d}v = K^* + \int_0^\infty \frac{\mathrm{d}(\widetilde{K}\mathrm{e}^{-\rho t})}{\mathrm{d}t} \mathrm{d}t = K^* - \widetilde{K}(0)$$
$$= K_0 - \widetilde{K}_0.$$

Hence $W(P) - W(\tilde{P}) \geq 0$. We next show that $W(P) - W(\tilde{P}) > 0$, so that P is a unique optimal solution. By almost everywhere continuity and since $P \neq \tilde{P}$, there exists an open interval $S \subset \mathbb{R}^+$ such that $\dot{\widetilde{K}}$ exists and either $\widetilde{D}_2 > 0$, $\tilde{Q} > 0$ or $\widetilde{K} \neq K^*$ on S. If $\widetilde{D}_2 = 0$ and $\widetilde{K} = K^*$, then $\tilde{Q} > 0$ on S and the formula for $\dot{K}$ gives $\widetilde{D}_1 - \tilde{Q} = F(K^*)$. This contradicts the constraint $\widetilde{D}_1 \leq F(\widetilde{K}) = F(K^*)$. Hence we must have $\widetilde{D}_2 > 0$ or $\widetilde{K} \neq K^*$. In this case either $I_1 > 0$ or $I_2 > K_0 - \widetilde{K}_0$ (for the latter inequality we use the strict concavity of F), so $W(P) - W(\tilde{P}) = I_1 + I_2 - (K_0 - \widetilde{K}_0) > 0$. This proves the claim.

We next prove (ii). Assume that $E < F(K^*)$, let the numbers ζ and t^* be defined as in (2) and (3) and let the $P = (K_0, D, Q)$ be as in Proposition 1 (ii). Then P is admissible and we need to show that $W(P) > W(\tilde{P})$ holds for any other $\tilde{P} = (\widetilde{K}_0, \widetilde{D}, \tilde{Q}) \in \mathcal{A}$. For this aim, define the auxiliary function λ by

$$\lambda(t) = \begin{cases} \dfrac{F(\zeta) - E}{F\big(K(t)\big) - E} \mathrm{e}^{\rho t}, & \text{for} \quad 0 \leq t < t^*, \\ \theta, & \text{for} \quad t^* \leq t. \end{cases}$$

Then $\lambda(0) = 1$, $\theta < \lambda < 1$ on $(0, t^*)$ and $\dot{\lambda} = \big(\rho - F'(K)\big)\lambda$ on $\mathbb{R}^+ \setminus \{t^*\}$. From the properties of P,

$$\begin{aligned} &W(P) - W(\tilde{P}) + K_0 - \widetilde{K}_0 \\ &\qquad = \int_0^\infty (1 - \lambda)\big(E - \widetilde{D}_1\big) \mathrm{d}v + \int_0^{t^*} (\lambda - \theta)\widetilde{D}_2 \,\mathrm{d}v + \int_0^\infty (1 - \lambda)\tilde{Q} \,\mathrm{d}v \\ &\qquad + \int_0^\infty \lambda\left(F(K) - F(\widetilde{K}) - (\dot{K} - \dot{\widetilde{K}})\right) \mathrm{d}v =: I_1 + I_2 + I_3 + I_4. \end{aligned}$$

Since $0 \leq \widetilde{D}_1 \leq E$ and $\widetilde{D}_2, \tilde{Q} \geq 0$, we have $I_1, I_2, I_3 \geq 0$. By the concavity of F, we have $F(K) - F(\widetilde{K}) \geq F'(K)(K - \widetilde{K})$ on Ω. Again, because $(K(t) - \widetilde{K}(t))\mathrm{e}^{-\rho t} \to 0$ as $t \to \infty$, we get

$$I_4 \geq \int_0^\infty \lambda\left(F'(K)(K - \widetilde{K}) - (\dot{K} - \dot{\widetilde{K}})\right) \mathrm{d}v = -\int_0^\infty \frac{\mathrm{d}(\lambda\big(K - \widetilde{K}\big)\mathrm{e}^{-\rho t})}{\mathrm{d}t} \mathrm{d}t$$
$$= K_0 - \widetilde{K}_0,$$

so that $W(P) - W(\tilde{P}) \geq 0$. For the strict inequality, we argue as above that there exists an open interval S such that $\dot{\tilde{K}}$ exists and either $\tilde{D}_1 < E$, $\tilde{Q} > 0$ or $\tilde{K} \neq K$ on S. This implies that $I_1 + I_3 + I_4 > K_0 - \tilde{K}_0$, so $W(P) - W(\tilde{P}) > 0$.

Finally, part (iii) follows by direct computation.

Proof of Proposition 2

For convenience, we reformulate problem (9): The aim is to find $P \in \mathcal{A}'$ which maximizes

$$W(P) = V_0 - K_0,$$

where $W(P)$ is as before and $P = (K_0, D, Q)$ belongs to the set $\mathcal{A}'$ of admissible policies if $K_0 \geq 0$, $D = D_1 + D_2$, the functions D_1, D_2 and Q are almost everywhere continuous on $\mathbb{R}_0^+$ and belong to $L^1(\mathbb{R}_0^+, \nu)$ and there are continuous and almost everywhere differentiable functions $K, R \in L^1(\mathbb{R}_0^+, \nu)$ such that $F(K) \in L^1(\mathbb{R}_0^+, \nu)$,

$$F(K) + Q = D + \dot{K} \quad \text{on} \quad \Omega,$$

$$K(0) = K_0,$$

$$\dot{R} = E - D_1 + \rho R \quad \text{on} \quad \Omega,$$

$$R(0) = 0,$$

$$D_1 \geq 0, D_2 \geq 0, D \leq F(K), Q \geq 0, K \geq 0, R \geq 0,$$

where Ω is the subset of $\mathbb{R}_0^+$ where D_1, D_2 and Q are continuous. We also assume that D is not identically zero on $\mathbb{R}_0^+$. In the model we have assumed that $i = \rho$. We assume that the threshold E is strictly positive. As pointed out in Section 5, the non-negativity of R is equivalent to the condition (10).

The proof of part (i) is similar to the proof of Proposition 1 (i) and part (iv) follows by direct computation so their details will be omitted. We next prove part (ii). Assume that $F(\zeta_0)/\theta < E < F(K^*)$ and let $P = (K_0, D, Q)$ be as in (ii). Then $P \in \mathcal{A}'$. Let $\tilde{P} = (\tilde{K}_0, \tilde{D}, \tilde{Q})$ be any other admissible policy. We need to show that $W(P) > W(\tilde{P})$. Define the function λ by

$$\lambda(t) = \begin{cases} \dfrac{F(\tilde{\zeta}_E)}{F(K(t))} e^{\rho t}, & \text{for} \quad 0 \leq t < \tilde{t}_E^*, \\ \dfrac{F(\tilde{\zeta}_E)}{E}, & \text{for} \quad \tilde{t}_E^* \leq t. \end{cases}$$

We have $\theta < F(\tilde{\zeta}_E)/E < \lambda < 1$ for all t and $\dot{\lambda} = (\rho - F'(K))\lambda$ on $\mathbb{R}^+ \setminus \{\tilde{t}_E^*\}$. By writing $\gamma = 1 - F(\tilde{\zeta}_E)/E$, and using the properties of P, we get

$$\begin{aligned} W(P) - W(\tilde{P}) + K_0 - \widetilde{K}_0 = \\ &= \int_0^\infty \left[(1-\gamma-\lambda)(D_1 - \widetilde{D}_1) + (\theta - \lambda)(D_2 - \widetilde{D}_2)\right. \\ &\left. - (1-\lambda)(Q - \tilde{Q})\right] \mathrm{d}\nu + \int_0^\infty \lambda\left(F(K) - F(\widetilde{K}) - (\dot{K} - \dot{\widetilde{K}})\right) \mathrm{d}\nu \\ &+ \int_0^\infty \gamma\left(\rho(R - \tilde{R}) - (\dot{R} - \dot{\tilde{R}})\right) \mathrm{d}\nu =: I_1 + I_2 + I_3. \end{aligned}$$

By the properties of λ, D_1, D_2 and Q, one has

$$I_1 = \int_0^{\tilde{t}_E^*} \left(\lambda - \frac{F(\tilde{\zeta}_E)}{E}\right) \widetilde{D}_1 \,\mathrm{d}\nu + \int_0^\infty (\lambda - \theta)\widetilde{D}_2 \,\mathrm{d}\nu + \int_0^\infty (1-\lambda)\tilde{Q} \,\mathrm{d}\nu \geq 0$$

and the integral I_2 can be estimated from below by $K_0 - \widetilde{K}_0$ as in the proof of Proposition 1 (ii) (see Appendix A). Finally, to estimate I_3, we solve from the formula for $\dot{R}$ that

$$R(t) = \mathrm{e}^{\rho t}\left(R(t_0)\mathrm{e}^{-\rho t_0} + \int_{t_0}^t (E - D_1(s))\mathrm{e}^{-\rho s} \,\mathrm{d}s\right), \qquad (12)$$

for any $t_0, t \geq 0$. Inserting $t_0 = \tilde{t}_E^*$, $D_1 = 0$ for $t \leq \tilde{t}_E^*$ and $D_1 = F(K^*)$ for $t > \tilde{t}_E^*$ in this equation shows that R takes the constant value $(F(K^*) - E)/\rho$ for $t \geq \tilde{t}_E^*$. Especially, $\lim_{t\to\infty} R(t)\mathrm{e}^{-\rho t} = 0$. Since $\tilde{R}(0) = R(0) = 0$, we get

$$\begin{aligned} I_3 = -\gamma \int_0^\infty \frac{\mathrm{d}\left(\left(R(t) - \tilde{R}(t)\right)\mathrm{e}^{-\rho t}\right)}{\mathrm{d}t} \mathrm{d}t &= \gamma \lim_{t\to\infty} \left(\tilde{R}(t) - R(t)\right) \mathrm{e}^{-\rho t} \\ &= \gamma \lim_{t\to\infty} \tilde{R}(t)\mathrm{e}^{-\rho t} \geq 0. \end{aligned}$$

Thus $W(P) \geq W(\tilde{P})$. The strict inequality follows with a similar argument as in the proof of Proposition 1 (ii) (see Appendix A).

We finally prove part (iii). Assume that $E \leq F(\zeta_0)/\theta$ and let $P = (K_0, D, Q)$ be any policy as in (iii). Then $P \in \mathcal{A}'$. The aim is to show that $W(P) > W(\tilde{P})$, for any policy $\tilde{P} = (\widetilde{K}_0, \widetilde{D}, \tilde{Q}) \in \mathcal{A}'$ not satisfying the properties in (iii). Let

$$\lambda(t) = \begin{cases} \dfrac{F(\zeta_0)}{F(K(t))}\mathrm{e}^{\rho t}, & \text{for } \quad 0 \leq t < t_0^*, \\ \theta, & \text{for } \quad t_0^* \leq t. \end{cases}$$

Then $\theta < \lambda < 1$ on $(0, t_0^*)$ and $\dot{\lambda} = \big(\rho - F'(K)\big)\lambda$ on $\Omega \setminus (\{0\} \cup \{t_0^*\})$. Now

$$\begin{aligned} W(P) - W(\tilde{P}) + K_0 - \widetilde{K}_0 \\ &= \int_0^\infty \Big[(\theta - \lambda)\Big(\big(D_1 - \widetilde{D}_1\big) + \big(D_2 - \widetilde{D}_2\big)\Big) + (\lambda - 1)\big(Q - \tilde{Q}\big)\Big]\, \mathrm{d}\nu \\ &+ \int_0^\infty \lambda\Big(F(K) - F\big(\widetilde{K}\big) - (\dot{K} - \dot{\widetilde{K}})\Big)\, \mathrm{d}\nu \\ &+ \int_0^\infty \tau\Big(\rho(R - \tilde{R}) - (\dot{R} - \dot{\tilde{R}})\Big)\, \mathrm{d}\nu =: I_1 + I_2 + I_3. \end{aligned}$$

By the properties of λ, D_1, D_2 and Q,

$$I_1 = \int_0^{t_0^*} (\lambda - \theta)\big(\widetilde{D}_1 + \widetilde{D}_2\big)\, \mathrm{d}\nu + \int_0^\infty (1 - \lambda)\tilde{Q}\, \mathrm{d}\nu \geq 0$$

and the integral I_2 can be estimated from below by $K_0 - \widetilde{K}_0$ as in the proof of Proposition 1 (ii) (Appendix A). To estimate I_3, note that $\lim_{t\to\infty} R(t)\, \mathrm{e}^{-\rho t} = 0$, by (11) and (12). Since $\tilde{R}(0) = R(0) = 0$, we get

$$I_3 = -\tau \int_0^\infty \frac{\mathrm{d}\Big(\big(R - \tilde{R}\big)\mathrm{e}^{-\rho t}\Big)}{\mathrm{d}t} \mathrm{d}t = \tau \lim_{t\to\infty} \Big(\tilde{R}(t) - R(t)\Big)\, \mathrm{e}^{-\rho t} = \tau \lim_{t\to\infty} \tilde{R}(t)\mathrm{e}^{-\rho t} \geq 0.$$

This shows that $W(P) \geq W(\tilde{P})$. Since $\tilde{P}$ does not satisfy the properties in (iii), either (a) there is an open subinterval of $\mathbb{R}_0^+$ on which $\dot{\widetilde{K}}$ exists and either $\tilde{Q} > 0$ or $\widetilde{K} \neq K$, (b) there is an open subinterval of $[0, t_0^*]$ on which $\widetilde{D} > 0$, or (c) there is an open subinterval of $[t_0^*, \infty)$ on which $\tilde{Q} = 0$ and $\widetilde{K} = K$ (implying $\widetilde{D} = D$) as well as

$$\int_0^\infty \big(E - \widetilde{D}_1(s)\big)\mathrm{e}^{-\rho s}\, \mathrm{d}s > 0.$$

In cases (a) and (b) we get as before that $W(P) > W(\tilde{P})$. In the case (c), $\lim_{t\to\infty} \tilde{R}(t)\mathrm{e}^{-\rho t} > 0$, so that $I_3 > 0$ and $W(P) > W(\tilde{P})$.