

# Effects of Tax Depreciation on Optimal Firm Investments

JACCO WIELHOUSER\*      PETER M. KORT<sup>†</sup>  
ANJA DE WAEGENAERE<sup>‡</sup>

July 5, 1999

## Abstract

This paper studies how the difference between technical depreciation and tax depreciation affects the firm's optimal investment strategy. The objective is maximization of shareholder value. When tax depreciation differs from technical depreciation, an additional investment not only generates value due to the fact that the firm can produce more, but also due to the fact that an additional deferred tax liability arises. Two types of capital stock will therefore affect shareholder value, i.e. the replacement value of the assets and the tax base of the assets. We present a dynamic model of the firm with these two types of capital stock, and study the effects of the tax depreciation rate on the firm's optimal dynamic investment strategy, dividend policy, and long run capital stock level.

**Key Words:** Tax depreciation, technical depreciation, deferred taxation, investments, shareholder value, dynamic optimization.

JEL Codes: C61, E22.

---

\*CentER Accounting research group, Tilburg University, The Netherlands.

<sup>†</sup>Department of Econometrics and CentER, Tilburg University, The Netherlands.

<sup>‡</sup>Department of Econometrics and CentER Accounting research group, Tilburg University, The Netherlands.

# 1 Introduction

There is extensive empirical literature on the effect of corporate taxation on firm's investment and dividend policies. Some recent examples are: Devereux et al. (1994), Faig and Shum (1999), Gentry (1994), Jacobs and Larkins (1998), Newberry (1998), and Pereira (1994). This literature mainly focuses on the effects of the effective tax rate, or asymmetries in corporate taxation, on firm's investment behavior. Our aim here is to study the effect of a difference between technical depreciation and tax depreciation on optimal firm investments.

Since firms are allowed to adopt a number of different depreciation methods for tax purposes, the depreciation method is often chosen in a strategic way (see e.g. Scholes and Wolfson, 1992, for a thorough overview of the different incentives that potentially affect this choice). As a consequence of this strategic behavior, tax depreciation hardly ever equals technical depreciation, which represents the real technical deterioration of the assets. This difference implies that the tax base of the assets generates value to the firm. When, for example, tax depreciation systematically exceeds technical depreciation, a deferred tax liability is generated, and, consequently, shareholder value is lower than the replacement cost of the assets.

Sansing (1998) presents a descriptive model in which firm value is determined under a given, static, investment policy, and provides a formula for the value of the deferred tax liability. Our aim is to determine the *optimal dynamic* investment policy, given the difference in tax depreciation and technical depreciation. The objective of the firm is to maximize its shareholder value. In contrast to the existing literature on dynamic firm investment (see e.g. Van Hilten et al., 1993), we consider a dynamic model that explicitly takes into account the tax base of the assets as well as the level of the capital stock. This allows us to study the effect of tax depreciation on the optimal investment strategy, dividend policy and capital stock level.

The main conclusions are twofold. First, we present results regarding the effects of the tax depreciation rate on the optimal long run capital stock level and investment behavior. They can be summarized as follows. (1) The optimal long run capital stock

level is the unique level at which marginal revenue of investment equals marginal cost, *and* taxable income is positive. (2) A firm that uses more accelerated tax depreciation has a higher optimal long run capital stock level. (3) While maintaining the optimal capital stock level in the long run, taxable income is positive, and tax depreciation converges to technical depreciation.

Second, we study the optimal dynamic investment strategy to reach the optimal long run level. The focus here is on firms that are initially small, so that marginal revenue of investment initially exceeds marginal cost. Here, we find that: (4) The firm should initially grow until it first reaches a level of capital where marginal revenue of investment equals marginal cost. Then, depending on the initial tax base of the assets, and on the difference between tax depreciation and technical depreciation, two situations can occur: (5i) A firm with low initial tax base or highly accelerated tax depreciation will have a positive taxable income by the time it first reaches a level where marginal revenue equals marginal costs, so that it is optimal to stabilize at this level (see 1)). (5ii) In contrast, when the initial tax base of assets is high and tax depreciation is not too accelerated, the firm will still have a negative taxable income by the time it first reaches its optimal long run level. As a consequence, marginal revenue at that time is still higher than marginal cost, so that the firm can benefit from growing further. After some time, however, taxable income will become positive, causing a reduction in marginal revenue. In order to optimally *anticipate* on this future reduction, the firm has to start shrinking *before* its taxable income becomes positive. In this way, it can avoid paying too much taxes in the future. Taxable income will then have become positive by the time the firm has shrunk to its optimal long run level, so that it can stabilize.

The paper is organized as follows. In section 2 we present the dynamic model of the firm and formulate the optimization problem. In section 3 we derive the optimal long run level of the capital stock, investment and dividend strategy, and show how they are affected by the tax depreciation rate. In section 4, the optimal dynamic investment strategies are presented and discussed. Here, we show how the difference between tax depreciation and technical depreciation, as well as the difference be-

tween initial tax base and capital stock level, affect the optimal investment strategy. The paper is concluded in section 5. The dynamic model is solved by path-coupling. The solution of the firm's optimization problem is presented in Appendix A. All other proofs are deferred to Appendix B.

## 2 The Model

The aim here is to study the effect of the tax depreciation rate on the firm's optimal investment strategy when the objective of the firm is to maximize its shareholder value. Shareholder value equals the present value of the dividend stream during the planning period, augmented with the value of the firm at the end of the planning period.

In the sequel, we denote  $I = I(t)$  for the amount of money invested at time  $t$ , and  $D = D(t)$  for the dividends paid to the shareholders at time  $t$ .

In order to be able to study the effect of the tax depreciation rate on the firm's optimal investment strategy, two separate capital stocks need to be distinguished. First, there is the 'real' capital  $K_1 = K_1(t)$  with which the firm produces. Second, there is the tax base of the assets, which we denote  $K_2 = K_2(t)$ .

Investments cause an equal increase in the capital stock and the tax base of assets. On the other hand, both capital stock and tax base decrease due to depreciation. Whereas the decrease of the real capital  $K_1$  depends on the technical depreciation rate  $\beta$ , the decrease of the tax base depends on the tax depreciation rate  $\gamma$  that is chosen by the firm. This implies that the evolution over time of the capital stock and the tax base are given by:

$$\dot{K}_1 = I - \beta K_1, \tag{1}$$

$$\dot{K}_2 = I - \gamma K_2, \tag{2}$$

where  $\dot{K}_i = \frac{\partial K_i(t)}{\partial t}$ , for  $i = 1, 2$ .

We consider tax depreciation rates in the range:

$$\beta \leq \gamma \leq r + \beta, \tag{3}$$

so that tax depreciation at least covers technical deterioration, but does not exceed the deterioration costs plus the time value of money (it will be difficult to get legal support for higher depreciation rates). It is then clear from (1) and (2) that the difference between the technical and the tax depreciation rate implies that the evolution over time of the tax base can deviate from that of the real capital stock. Moreover, whenever  $K_2$  deviates from  $K_1$ , a deferred tax liability (or asset) is generated.

The deferred tax liability can clearly also affect the firm's investment and dividend policies. As stated above, the aim is to study how the investment policy that maximizes shareholder value depends on the choice of the tax depreciation rate  $\gamma$ . Since shareholder value consists of the discounted dividend stream over the planning period, augmented with the firm's value at the horizon date, we first need to specify how, for any given investment policy, the dividend policy and the final value of the firm depend on  $\gamma$ .

Let us start with the dividend policy. Since the firm's objective is to maximize shareholder value it holds no cash, so that dividends consist of gross revenue reduced with investments and tax payments. Producing with capital stock  $K_1$  yields a gross revenue  $C(K_1)$ . The revenue is increasing in  $K_1$  ( $C'(\cdot) > 0$ ), and exhibits decreasing returns to scale ( $C''(\cdot) < 0$ ). A fixed tax rate  $T$  is paid over taxable income, which equals gross revenue  $C(K_1)$  minus tax depreciation  $\gamma K_2$ , if positive. Consequently, the dividend paid to the shareholders equals:<sup>1</sup>

$$D(t) = C(K_1(t)) - I(t) - T(C(K_1(t)) - \gamma K_2(t))^+. \quad (4)$$

In order to determine the value of the firm at the horizon date  $z$ , notice that in general the optimal investment policy consists of roughly two phases: a final phase where the firm carries out only replacement investments to keep its capital stock at a long run optimal level, and an initial phase in which the firm grows or shrinks towards its optimal long run level (see e.g. Van Hilten et al., 1993). We assume

---

<sup>1</sup>Here,  $x^+$  denotes the function that equals  $x$  if  $x > 0$ , and 0 otherwise, so that in (4) dividend is only reduced with tax payments if taxable income is positive.

that the time horizon is long enough so that the firm can reach this optimal size before the final time  $z$ . It can then maintain its optimal size  $K_1(z)$  by engaging in replacement investment, i.e. investments to compensate for technical deterioration. This implies that, for all  $t \geq z$ , one has:

$$I(t) = \beta K_1(z), \tag{5}$$

$$K_1(t) = K_1(z). \tag{6}$$

The value of the firm at time  $z$ , which we denote  $f(K_1(z), K_2(z))$ , equals the discounted future dividend stream. Given that investments and capital stock evolve according to (5) and (6), and since the firm's taxable income is positive in the future<sup>2</sup>, this implies that:

$$f(K_1(z), K_2(z)) = \int_z^\infty e^{-r(t-z)} [(1-T)C(K_1(z)) - \beta K_1(z) + T\gamma K_2(t)] dt.$$

Taking into account the evolution of the tax base of assets after time  $z$ , which can be derived from (2) and (5), one finds that:

**Proposition 2.1** *The value of the firm at time  $z$  equals:*

$$f(K_1(z), K_2(z)) = \frac{1-T}{r} [C(K_1(z)) - \beta K_1(z)] + \frac{\gamma - \beta}{r + \gamma} T K_1(z) + \frac{\gamma}{r + \gamma} T [K_2(z) - K_1(z)], \tag{7}$$

where  $K_1(z)$  is the optimal long run level of the capital stock.

It follows from (7) that the marginal value generated by the tax base of assets at time  $z$ , equals (see also Sansing, 1998):

$$\frac{\partial f}{\partial K_2}(K_1(z), K_2(z)) = \frac{T\gamma}{r + \gamma}.$$

Notice now that the final value of the firm consists of three terms. The first term represents the discounted value generated through production with the capital stock,

---

<sup>2</sup>In Appendix A it is proven that if the optimal investment policy is applied, taxable income will be positive in the final phase.

taking into account the technical depreciation and the replacement investments. The last two terms result from a difference in tax depreciation rate and technical depreciation rate, or a difference in initial tax base and initial capital stock level. This implies that, when both  $\gamma = \beta$ , and  $K_2(0) = K_1(0)$ , then the last two terms vanish. Notice that the third term equals the value of the deferred tax liability derived in Sansing (1998).

Now, for any given technical depreciation rate  $\beta$ , tax depreciation rate  $\gamma$ , initial capital stock level  $K_1(0)$ , initial tax base of the assets  $K_2(0)$ , and investment policy  $I(\cdot)$ , the resulting dividend policy  $D(\cdot)$  and the final value  $f(\cdot, \cdot)$  can be determined. Shareholder value then equals:

$$\int_0^z e^{-rt} D(t) dt + e^{-rz} f(K_1(z), K_2(z)). \quad (8)$$

Then, assuming irreversibility of investment, i.e.  $I(\cdot) \geq 0$ , (for arguments, see e.g. Dixit and Pindyck, 1994), and non-negative dividends, the investment strategy that maximizes shareholder value is the solution of the following optimization problem:

$$\max_{I(t)} \int_0^z e^{-rt} [C(K_1) - I - T(C(K_1) - \gamma K_2)^+] dt + e^{-rz} f(K_1(z), K_2(z)) \quad (9)$$

$$\begin{aligned} s.t. \quad & \dot{K}_1 = I - \beta K_1, \\ & \dot{K}_2 = I - \gamma K_2, \\ & C(K_1) - I - T(C(K_1) - \gamma K_2)^+ \geq 0, \\ & I \geq 0, \end{aligned}$$

where  $r$  denotes the discount rate, i.e. the shareholder time preference rate.

The optimization problem is solved by applying the maximum principle (see e.g. Feichtinger and Hartl, 1986). In the sequel, we will assume that the revenue function

satisfies  $C(0) = 0$ , and:

$$C'(0) \geq \frac{r}{1-T} + \beta, \quad (10)$$

$$C'(\cdot) > \beta, \quad (11)$$

$$C'(\infty) \leq r + \beta, \quad (12)$$

$$C(K_1) > \frac{\beta}{1-T}K_1, \quad \text{for all } K_1 \leq K_1^*, \quad (13)$$

where  $K_1^*$  denotes the optimal long run size of the firm, which will be derived in the next section.

Given that the revenue function has decreasing returns to scale, (10) implies that, at least when  $K_1(\cdot) = 0$ , marginal revenue of production exceeds the user cost of capital, so that it is always worthwhile for the firm to start producing. If (12) were not satisfied, then for firms with strongly accelerated tax depreciation marginal revenue of investment would exceed marginal cost, even at extremely high levels of the capital stock. This would clearly not be a realistic situation. Assumption (13) implies that, as long as the capital stock level is lower than its optimal long run level, a firm that engages in replacement investments will have after tax revenue to pay out to the shareholders.

### 3 The optimal size of the firm

The optimal investment strategy of the firm can be divided in two parts. First, it has to grow or shrink to its optimal long run capital stock level. From there on, it has to invest and pay dividends so as to maintain this optimal level until the horizon date.

In this section we first derive the optimal long run capital stock level, and show how it depends on the depreciation rates, the tax rate, and the discount rate. The following proposition yields an implicit expression for the optimal size of the firm.

**Proposition 3.1** *The optimal long run level of the capital stock  $K_1^*$  satisfies:*

$$C'(K_1^*) = \left( \frac{r}{1-T} + \gamma \right) \left( \frac{r + \beta}{r + \gamma} \right). \quad (14)$$



Moreover,  $K_1^*$  is the unique long run capital stock level at which marginal revenue of investment equals marginal cost, and taxable income is positive.

Notice that the expression in (14) can be rewritten as follows:

$$C'(K_1^*) = \beta + \frac{r}{1-T} - \frac{rT}{(1-T)(r+\gamma)}(\gamma - \beta),$$

which, in its last term, clearly reveals the effect of a difference between  $\gamma$  and  $\beta$  on the marginal revenue of investment at the optimal long run capital stock level.

The above proposition makes clear that the optimal size depends on both the tax depreciation rate and the technical depreciation rate. Furthermore, it also depends on the tax rate and on the time value of money. The following proposition provides more details on the effect of these parameters on the optimal size of the firm.

**Proposition 3.2** *i) Whenever  $r > 0$ ,  $K_1^*$  is increasing in  $\gamma$ , i.e. more accelerated tax depreciation implies a higher optimal long run level of the capital stock. Moreover, for each value of  $\gamma \geq 0$ , one has*

$$C'(K_1^*) \in \left[ r + \beta, \frac{r + \beta}{1-T} \right].$$

*ii)  $K_1^*$  is decreasing in  $\beta$ , i.e. more accelerated technical depreciation implies a lower optimal long run level of the capital stock.*

*iii) Whenever  $r > 0$ , the optimal long run capital stock level is decreasing in both the interest rate  $r$  and the tax rate  $T$ .*

*iv) When the interest rate is zero, the optimal long run capital stock level is independent of both the tax depreciation rate  $\gamma$  and the tax rate  $T$ , and satisfies:*

$$C'(K_1^*) = \beta.$$

Figure 1 provides a graphical exposition of the joint effect of  $\beta$  and  $\gamma$  on the optimal size of the firm, for  $T = 0.35$ ,  $r = 0.20$ , and  $C(x) = 200\sqrt{x}$ .

It is known that, as long as the company's taxable income is never negative, more accelerated depreciation implies that the present value of future tax payments

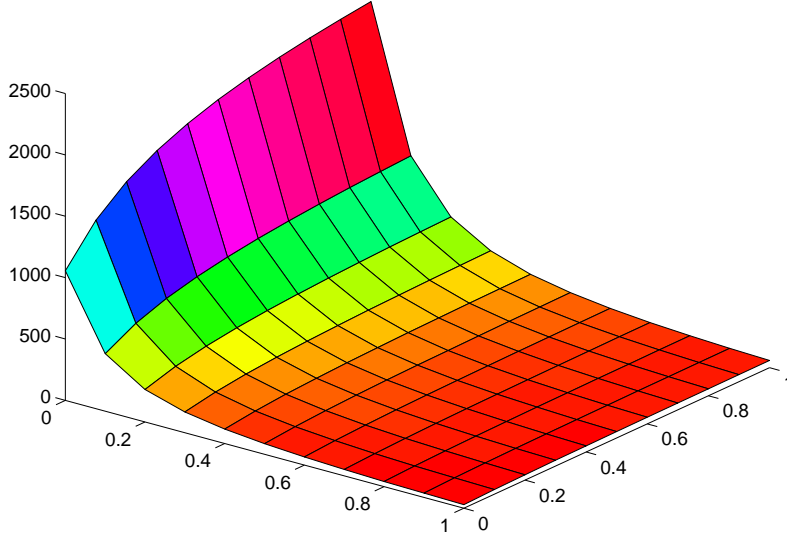


Figure 1:  $K_1^*$  as a function of  $\beta$  and  $\gamma$  (in units of \$ 10.000).

decreases (see e.g. Wakeman, 1980). This is due to the time value of money. The above proposition says that, as a consequence of this effect, a firm that uses more accelerated depreciation should also grow to a higher level of capital stock. When the interest rate is zero, there is no time value of money, and the depreciation policy does not affect the net present value of future tax payments. The optimal capital stock then does not depend on  $\gamma$  or on the tax rate  $T$ .

The following proposition describes the optimal dividend and investment policies once the firm has reached its steady state, as well as the long run behavior of the tax base of assets.

**Proposition 3.3** *Once the steady state is reached, the optimal strategy implies that:*

- i) The firm maintains its optimal level by engaging in replacement investment, i.e.  $I(t) = \beta K_1^*$ .*
- ii) Tax depreciation converges to technical depreciation, i.e.*

$$\lim_{t \rightarrow \infty} \gamma K_2(t) = \beta K_1^*.$$

iii) *Taxable income is positive, i.e.  $C(K_1^*) - \gamma K_2(t) > 0$ .*

iv) *Dividend payments equal:*

$$D(t) = (1 - T)[C(K_1^*) - \beta K_1^*] + e^{-\gamma(t-z)}T[\gamma K_2(z) - \beta K_1^*]. \quad (15)$$

We see from (15) that dividends consist of value generated through production with the capital stock, taking into account the technical depreciation (first term), and value generated through the difference in tax depreciation and technical depreciation (second term). The latter converges to zero in the long run.

## 4 The optimal dynamic investment strategy

In the previous section the optimal long run behavior of the firm is derived. In this section, we present the optimal dynamic strategy the firm should use in order to reach to the optimal long run capital stock level. At each point in time, the firm has to decide on how much it will invest. Given (4), this decision then immediately determines the amount of dividend it will pay to its shareholders. The possible decisions of the firm can therefore be categorized in four different policies:

$P_1$ ) The firm does invest, but not all of its net profits, so that after tax revenues are used for both paying a positive amount of dividend and for investing in the firm.

$P_2$ ) All after tax revenue is used for investments, so that no dividend is paid.

$P_3$ ) The after tax revenue is as a whole paid to the shareholder, so that nothing is invested.

$P_4$ ) The firm neither pays dividend nor invests. Since the firm holds no cash, this only occurs when  $K_1 = 0$ .

Given (10), it is never optimal for the firm to have its capital stock reduced to zero. Therefore, policy 4 will never be part of the optimal strategy.

The optimal dynamic strategy can be described as a sequence of policies used by the firm until it reaches its steady state. The initial state of the firm, and in particular the difference between initial tax base and capital stock, and between tax depreciation and technical depreciation will affect the optimal strategy. We therefore introduce the following terminology.

- The firm has *low initial tax depreciation* if, when constantly investing all its net revenue, it will have a positive taxable income by the time it reaches its optimal long run level  $K_1^*$ .
- The firm has *high initial tax depreciation* if, even when investing all its net revenue, it will still have a negative taxable income by the time it reaches its optimal long run capital stock level. This could be due to a high initial tax base  $K_2(0)$ , or a moderate  $\gamma$ , i.e. not too accelerated depreciation.

The following result holds for both type of firms.

**Proposition 4.1** *If the firm follows its optimal dynamic investment strategy, then, once taxable income has become positive, it will remain positive.*

This result confirms the following intuition. Due to the discounting effect, paying taxes later is preferable to paying them now. Moreover, if it is optimal for a firm to grow, it can grow faster as long as taxable income is zero. Notice finally that, since uncertainty on realized revenue is not modeled explicitly, the evolution of the capital stock and tax base of assets have to be seen as the average trend around which the realized values will evolve. This implies that, although in expectation taxable income will remain positive, fluctuations in realized revenues may cause taxable income to be zero.

We now describe how the initial tax base as well as the tax depreciation rate affect the optimal dynamic strategy of the firm. To focus attention, we consider firms for which the marginal revenue of investment exceeds marginal cost at the start of the planning period.

**Proposition 4.2** *For firms with low initial tax depreciation, the optimal strategy is as follows:*

- 1) *Policy 2 (i.e. invest all after tax revenue) is used until time  $t^*$  where marginal revenue of investment equals marginal cost. At time  $t^*$ , taxable income is positive, and the capital stock level equals the optimal long run level  $K_1^*$ .*
- 2) *From there on, policy 1 is used with replacement investments ( $I = \beta K_1^*$ ).*

The intuition is as follows. Since marginal revenue of investment exceeds marginal cost at time 0, the firm should start investing all its net revenue until the marginal revenue of an additional investment equals its marginal cost. This time instant is denoted  $t^*$ . When the firm uses strongly accelerated depreciation, or when its initial tax base  $K_2(0)$  is sufficiently low, it starts paying taxes before time  $t^*$ , so that it follows from proposition 3.1 and proposition 4.1 that the optimal level of the capital stock is reached at time  $t^*$ . From this moment onwards until the end of the planning horizon, there are only replacement investments, so that  $K_1$  stays at the level  $K_1^*$ . The remaining net revenue (i.e. revenue net from taxes) is paid as dividend to the shareholders.

**Proposition 4.3** *For firms with high initial tax depreciation, the optimal strategy is as follows:*

- 1) *Policy 2 (i.e. invest all after tax revenue) is used until time  $t^*$  where marginal revenue of investment equals marginal cost. At time  $t^*$ , taxable income is negative, and the capital stock level is strictly higher than the optimal long run level  $K_1^*$ .*
- 2) *Then, in order to anticipate on future reduction in marginal revenues due to tax payments, first policy 1 is used with (partial) replacement investments, and subsequently, policy 3 is used, so that the firm shrinks maximally until it again reaches its optimal long run level  $K_1^*$ . Taxable income will become positive during this maximal shrinking phase, and before  $K_1^*$  is reached.*

3) *From there on, policy 1 is used with replacement investments.*

The intuition here is as follows. As before, the firm starts investing all its net revenue and pays no dividends (policy 2), since in the initial state marginal revenue of investment exceeds marginal costs. Due to the high initial tax depreciation, however, the firm's taxable income is still negative by the time it first reaches the optimal long run capital stock level  $K_1^*$ . This implies that, in contrast to the previous case, marginal revenue exceeds marginal cost at the time  $K_1^*$  is reached.

The firm then has to choose between growing further and stabilizing. Both options have their disadvantage. Growing further implies that the capital level will be suboptimally high by the time taxable income becomes positive, since at that time, tax payments will cause a decrease in marginal revenues. Stabilizing at  $K_1^*$  implies that the firm's capital stock level will be suboptimally low as long taxable income is still negative, since marginal revenue is higher than marginal cost.

The optimal strategy is therefore anticipative: first take advantage of the high tax depreciation to grow to a high capital stock level, but anticipate on future reduction in marginal revenue when taxable income will become positive.

More precisely, the firm should first exploit the benefits of high tax depreciation by investing all its net revenue until the time  $t^*$  where marginal cost of an additional investment equals marginal revenue. Consequently, it grows to a level  $K_1(t^*)$  that is strictly higher than the optimal long run level. Maintaining the high level  $K_1(t^*) > K_1^*$  in the long run however implies high tax payments in the future. Indeed, when the firm starts paying taxes, the marginal revenue of investment decreases from  $C'(K_1)$  to  $(1 - T)C'(K_1)$ . Due to tax depreciation, the net reduction in marginal revenue of investment equals  $TC'(K_1) - \gamma T$ . It follows from the proof of proposition 4.3 that in the optimal strategy the gain through tax depreciation ( $\gamma T$ ) cannot fully compensate for the loss in return on investment ( $TC'(K_1)$ ). Therefore, there will be an inefficient period in the optimal solution in which marginal revenue is less than marginal costs. Given that there are decreasing returns to scale, the firm has to anticipate on this by decreasing productive capital stock in order to increase marginal revenue.

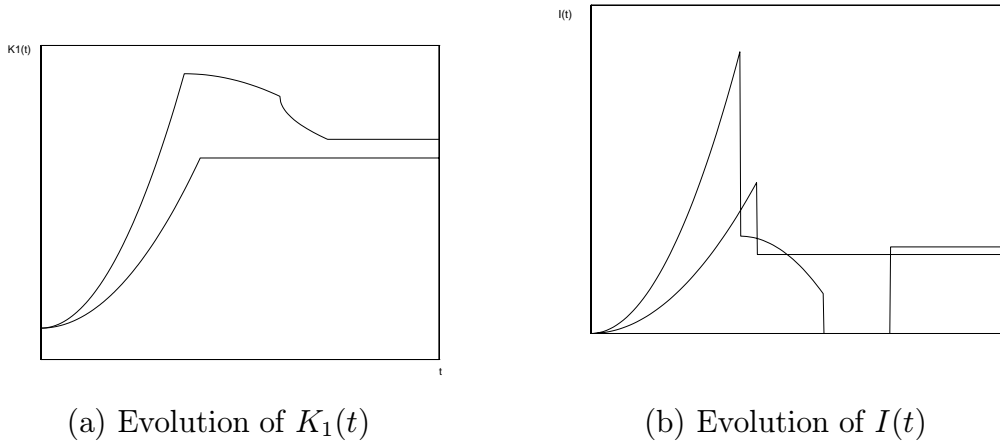


Figure 2: *Comparison of evolution over time of the capital stock  $K_1(t)$  and the investments  $I(t)$  for low and high initial tax depreciation.*

This anticipative time period consists of two phases. First, the firm should stabilize its capital stock by doing only replacement investments if  $\gamma = \beta$ , and should shrink by doing partial replacement if  $\gamma > \beta$  (policy 1), while keeping marginal revenue equal to marginal cost<sup>3</sup>. Then, as the tax base decreases further, it becomes optimal to stop investing (policy 3), until the optimal capital stock level is reached for the second time. From there on, the firm continues with replacement investments in order to maintain the optimal level of capital stock until the end of the planning horizon, i.e invest  $I(t) = \beta K_1^*$ . During policy 3 the firm's taxable income has become positive.

Figure 2 illustrates the development of the capital stock (panel a)) and the investments (panel b)) in the optimal solution for both types of firms. We consider two firms that start with the same capital stock level at date zero. For the first firm (lower curves), the tax depreciation rate  $\gamma$ , and the initial tax base  $K_2(0)$  are such that  $\gamma K_2(0) < C(K_1(0))$ , so that taxable income is positive at time zero. The

---

<sup>3</sup>During policy 1, marginal cost equals marginal revenue. When  $\gamma = \beta$ , the change in marginal cost over time equals the change in marginal revenue. When  $\gamma > \beta$ , marginal revenue becomes lower than marginal cost due to the change in time, so that the firm has to shrink in order to keep marginal revenue equal to marginal cost.

second firm (upper curves) uses a tax depreciation rate  $\gamma > \beta$ , and has an initial tax base  $K_2(0)$  such that  $\gamma K_2(0) > C(K_1(0)) > \beta K_1(0)$ . Its taxable income is therefore negative at time zero. We see that, in the initial stage, the first firm grows slower than the second one, since its taxable income is positive at date zero. The second firm not only initially grows faster, but it also grows to a higher level than the first one. It then shrinks to its optimal long run level, which, due to the higher value for  $\gamma$ , is higher than the long run capital stock level for the first firm.

## 5 Summary of the results

The main conclusion of the paper is that the difference between tax depreciation and technical depreciation affects the optimal size of the firm as well as its optimal dynamic investment and dividend policy. Differences in tax and technical depreciation can be caused by different depreciation rates, and by a difference in initial capital stock level and tax base.

For the optimal size of the firm, we can conclude that a firm that uses a higher tax depreciation rate should also grow to a higher level of capital stock.

For the dynamic dividend and investment policy that leads to the optimal size, we can conclude that:

- i) It is always optimal to grow until the first time  $t^*$  where marginal revenue of investment equals marginal costs. The optimal strategy from there on depends on whether the capital level at time  $t^*$ ,  $K_1(t^*)$ , equals the optimal long run level, or is strictly higher than this level. The latter will occur if the firm's taxable income is still negative at time  $t^*$ , so that marginal revenue is not yet negatively affected by tax payments.
- ii) The optimal investment strategy implies that, once the firm's taxable income has become positive, it will remain positive all over the planning period.

Moreover, the optimal strategy of the firm crucially depends on its initial tax base and initial capital stock level, and on the choice of the tax depreciation rate  $\gamma$ .



The following holds:

- iii) If the firm has low initial tax depreciation its taxable income will be positive by the time it reaches its optimal capital stock level. Therefore, the firm can stabilize from there on.
- iv) Firms with high initial tax depreciation will still have a negative taxable income by the time they first reach their optimal long run size. They therefore maximize shareholder value by initially growing to a higher capital stock level, and shrinking later on to the optimal size of the capital stock.

For firms that use a tax depreciation rate that equals the real technical depreciation, and that have a tax base of assets equal to the capital stock, the optimal strategy is intuitively clear, i.e. grow until the optimal size is reached, and then stabilize.

For firms that start with a high tax base of assets, the dynamic investment strategy is non-trivial in the sense that the firm should first grow, and consequently shrink to its optimal level. The shrinking phase consists of two parts: first do partial replacement investment, and then shrink maximally by not investing at all. The second part of the shrinking phase is an inevitable inefficient, but optimal strategy. It is a consequence of the fact that, for any reasonable value of  $\gamma$ , it is impossible to keep marginal costs and marginal revenue equal at the time taxable income becomes positive. So it is inefficient in the sense that marginal cost cannot be kept equal to marginal revenue, but it is optimal in the sense that it maximizes shareholder value for the firms with high initial tax depreciation.

# Appendices

## A The path-coupling method

In this appendix we show how the path-coupling method is used to prove propositions 4.1, 4.2 and 4.3.

It is structured as follows. In section A.1 we define the Lagrangian of optimization problem (9), and present the corresponding necessary conditions for optimality of an investment strategy. The strategy of a firm can be defined as a sequence of paths. In section A.2 we present the different paths and discuss their dynamics. To obtain the optimal sequence of paths, a formal synthesizing procedure (path coupling) is applied. It determines which path(s) can precede a given path, exploiting the continuity of state- and costate variables, and the necessary conditions for optimality. This procedure is presented in section A.3. The main theorem in that section states that there are two master trajectories. This allows to prove propositions 4.2 and 4.3.

### A.1 The necessary conditions

The current value Lagrangian of problem (9) is as follows:

$$L(K_1, K_2, I, \lambda_1, \lambda_2, \eta, \mu) = (1 + \eta_1) \left\{ C(K_1) - I - T[C(K_1) - \gamma K_2]^+ \right\} \\ + \eta_2 I + \lambda_1 (I - \beta K_1) + \lambda_2 (I - \gamma K_2),$$

in which  $\lambda_1$  and  $\lambda_2$  are the co-state variables corresponding to  $K_1$  and  $K_2$ , respectively, and  $\eta_1$  and  $\eta_2$  are the Lagrange multipliers associated with the two non-negativity constraints. The co-state variables can be interpreted as shadow-prices, i.e.  $\lambda_1(t)$  is the value of an additional unit of  $K_1(t)$  in terms of shareholder value. The additional contribution to the objective function of a unit of  $K_2(t)$  is equal to  $\lambda_2(t)$ .

The necessary conditions for optimality are:

$$\dot{K}_1 = \frac{\partial L}{\partial \lambda_1} = I - \beta K_1, \quad (16)$$

$$\dot{K}_2 = \frac{\partial L}{\partial \lambda_2} = I - \gamma K_2, \quad (17)$$

$$\dot{\lambda}_1 = r\lambda_1 - \frac{\partial L}{\partial K_1} = (r + \beta)\lambda_1 - [1 + \eta_1]C'(K_1)[1 - T \cdot 1_{\{C - \gamma K_2 > 0\}}], \quad (18)$$

$$\dot{\lambda}_2 = r\lambda_2 - \frac{\partial L}{\partial K_2} = (r + \gamma)\lambda_2 - \gamma(1 + \eta_1)T1_{\{C - \gamma K_2 > 0\}}, \quad (19)$$

$$\lambda_1 + \lambda_2 = 1 + \eta_1 - \eta_2, \quad (\Leftarrow \frac{\partial L}{\partial I} = 0), \quad (20)$$

$$\eta_1(C(K_1) - I - T(C(K_1) - \gamma K_2)^+) = 0, \quad \eta_1 \geq 0, \quad (21)$$

$$\eta_2 I = 0, \quad \eta_2 \geq 0, \quad (22)$$

$$0 \leq I \leq C(K_1) - T(C(K_1) - \gamma K_2)^+, \quad (23)$$

$$\lambda_1(z) = \frac{\partial f}{\partial K_1}(K_1^*(z), K_2^*(z)), \quad (24)$$

$$\lambda_2(z) = \frac{\partial f}{\partial K_2}(K_1^*(z), K_2^*(z)). \quad (25)$$

The Lagrangian is not differentiable in the point where  $C(K_1) = \gamma K_2$ . Therefore, in this point, the necessary conditions (18) and (19) have to be replaced by (see e.g. Hartl and Kort, 1996, pp. 257):

$$\dot{\lambda}_1 \in \left[ (r + \beta)\lambda_1 - (1 + \eta_1)C'(K_1), (r + \beta)\lambda_1 - (1 + \eta_1)C'(K_1)(1 - T) \right] \quad (26)$$

$$\dot{\lambda}_2 \in \left[ (r + \gamma)\lambda_2 - \gamma(1 + \eta)T, (r + \gamma)\lambda_2 \right]. \quad (27)$$

so that the changes of the shadow prices can vary within a certain range.

## A.2 Evaluating the paths

The firm has to optimize its investments under the constraint that both investment and dividend payments have to be non-negative. This yields constraints (23). One can now characterize a path followed by the firm at a time instant depending on whether either of the two constraints is binding at that time. The four different paths are therefore characterized by:

	1	2	3	4
$\eta_1$	=	>	=	>
$\eta_2$	=	=	>	>

In terms of policies of the firm, path 1 is the policy of paying dividend and investing a positive amount. With policy 2, the firm does not pay dividend but invests all its after tax revenues. With policies 3 and 4, the firm does not invest. Using policy 3, there is a positive amount of dividend paid to the shareholder, while using policy 4 implies dividend payments are zero.

**Lemma A.1** *Path 4 is not feasible.*

**Proof:** On path 4, one has  $\eta_1 > 0$  and  $\eta_2 > 0$ . (21) and (22) then imply that:

$$I(t) = 0 = C(K_1) - T(C(K_1) - \gamma K_2)^+,$$

which implies that  $K_1 = 0$ . It follows from (10) that it is never optimal to have the capital stock reduced to zero.  $\square$

In the sequel, we will refer to the use of policy  $i$ ,  $i = 1, 2, 3$ , in a situation where the firm's taxable income is positive as policy  $i_A$ , and in a situation where the firm's taxable income is negative as policy  $i_B$ . We now present the dynamics of the states and co-states on each of the six feasible paths.

1<sub>A</sub>) Here, taxable income is positive ( $C(K_1) > \gamma K_2$ ), and  $\eta_1 = \eta_2 = 0$ . Therefore, we find:

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - C'(K_1)(1 - T), \quad (28)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2 - \gamma T, \quad (29)$$

$$\lambda_1 + \lambda_2 = 1. \quad (30)$$

Whether path 1<sub>A</sub> is a shrink- or a growth path will depend on the division of net revenues in dividend and investments. This path can be a steady state path with  $\dot{\lambda}_1$ ,  $\dot{\lambda}_2$  and  $\dot{K}_1$  equal to 0.

1<sub>B</sub>) The fact that taxable income is zero, and  $\eta_1 = \eta_2 = 0$ , imply that:

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - C'(K_1), \quad (31)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2, \quad (32)$$

$$\lambda_1 + \lambda_2 = 1. \quad (33)$$

It is now shown that, if  $\gamma > \beta$ , then  $\dot{K}_1 < 0$ , so that the capital stock decreases on path 1<sub>B</sub>. In contrast, when  $\gamma = \beta$ , one finds that  $\dot{K}_1 = 0$ , so that the capital stock is then constant on this path.

**Proposition A.1** *On Path 1<sub>B</sub>, the firm's capital stock decreases, i.e.  $\dot{K}_1 < 0$ , when  $\gamma > \beta$  and is constant, i.e.  $\dot{K}_1 = 0$  when  $\gamma = \beta$ .*

**Proof:** Since

$$\dot{\lambda}_1 + \dot{\lambda}_2 = 0,$$

the continuity of the co-state variables implies that:

$$\begin{aligned} \frac{\partial \dot{\lambda}_1}{\partial t} &= -\frac{\partial \dot{\lambda}_2}{\partial t} \\ \Rightarrow (r + \beta)\dot{\lambda}_1 - C''(K_1)\dot{K}_1 &= -(r + \gamma)\dot{\lambda}_2 \\ \Rightarrow -(r + \beta)\dot{\lambda}_2 - C''(K_1)\dot{K}_1 &= -(r + \gamma)\dot{\lambda}_2 \\ \Rightarrow C''(K_1)\dot{K}_1 &= (\gamma - \beta)\dot{\lambda}_2 \end{aligned}$$

Now, since  $C''(K_1) < 0$  and  $\dot{\lambda}_2 > 0$ , it is seen immediately that  $\dot{K}_1 < 0$  when  $\gamma > \beta$  and  $\dot{K}_1 = 0$  when  $\gamma = \beta$ .  $\dot{\lambda}_2 > 0$  holds if  $\lambda_2 > 0$ . The path coupling procedure later on in this Appendix shows that  $\lambda_2 > 0$  in an optimal solution. Intuitively  $\lambda_2 > 0$  means that the value/shadow price of an additional unit of  $K_2$  is positive.  $\square$

2<sub>A</sub>) On this path, no dividend is paid, and  $\eta_1 > 0$ , and  $\eta_2 = 0$ , which implies that:

$$I = C(K_1) - T(C(K_1) - \gamma K_2)^+. \quad (34)$$

Since taxable income is positive, one has:

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - (1 + \eta_1)C'(K_1)(1 - T), \quad (35)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2 - (1 + \eta_1)\gamma T, \quad (36)$$

$$\lambda_1 + \lambda_2 = 1 + \eta_1. \quad (37)$$

Now, conditions (16) and (17), combined with (34) imply that:

$$\dot{K}_1 = C(K_1) - T(C(K_1) - \gamma K_2) - \beta K_1,$$

$$\dot{K}_2 = (1 - T)[C(K_1) - \gamma K_2],$$

The fact that  $C(K_1) - \gamma K_2 > 0$  on this path clearly implies that  $\dot{K}_2 > 0$ , so that the tax base increases on this path. Furthermore, (13) implies that  $\dot{K}_1 > 0$ , so that also the capital stock increases for  $K_1 \leq K_1^*$ .

2<sub>B</sub>) Concerning the B-part, we find that:

$$\dot{K}_1 = C(K_1) - \beta K_1,$$

$$\dot{K}_2 = C(K_1) - \gamma K_2,$$

On the B-part, clearly  $K_2$  is decreasing. As for the A-part,  $K_1$  is increasing.

The necessary conditions (18) and (19) of this path are:

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - (1 + \eta_1)C'(K_1), \quad (38)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2. \quad (39)$$

3<sub>A</sub>) On this path investments are zero and dividends are positive, so that  $\eta_1 = 0$  and  $\eta_2 > 0$ . The dynamics of the state are:

$$\dot{K}_1 = -\beta K_1, \quad (40)$$

$$\dot{K}_2 = -\gamma K_2, \quad (41)$$

which clearly implies that both  $K_1$  and  $K_2$  are decreasing. Furthermore,

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - (1 - T)C'(K_1), \quad (42)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2 - \gamma T. \quad (43)$$

These conditions are the same as for path  $1_A$ . The differences are that:

$$D = C(K_1) - T(C(K_1) - \gamma K_2), \quad (44)$$

$$\lambda_1 + \lambda_2 = 1 - \eta_2. \quad (45)$$

3<sub>B</sub>) For the B-part, (45) remains unchanged but the other three necessary conditions become:

$$\dot{\lambda}_1 = (r + \beta)\lambda_1 - C'(K_1), \quad (46)$$

$$\dot{\lambda}_2 = (r + \gamma)\lambda_2, \quad (47)$$

$$D = C(K_1). \quad (48)$$

Since, as in 3<sub>A</sub>,  $\dot{K}_1 = -\beta K_1$ , this also is a shrink-path.

### A.3 Coupling the paths

In order to determine the optimal maximal sequence of policies, the analysis starts at the end of the planning horizon. Necessary transversality conditions ((24) and (25)) yield the optimal policy for the firm at the end of the planning period. For a sufficiently long time-horizon, the firm typically end in a steady state path where dividend is paid to the shareholder. This is shown in various other dynamic models of the firm (see e.g. Van Hilten et al., 1993). One then systematically checks which paths can be coupled before the next path without violating the optimality conditions.

**Proposition A.2** *The final path is path  $1_A$ . This is a steady state path.*

**Proof:** To show that path  $1_A$  can be the final path, the necessary conditions (28) - (30), together with the transversality conditions have to be solved. The transversality conditions imply that  $\lambda_2(z) = \frac{T\gamma}{r+\gamma}$ . Together with (29) one has:

$$\dot{\lambda}_2 = 0$$

(30) implies that  $\dot{\lambda}_1 + \dot{\lambda}_2 = 0$ . So it follows that  $\dot{\lambda}_1 = 0$ . Given (28) and (29) this implies that:

$$r + \beta\lambda_1 + \gamma(1 - \lambda_1) - C'(K_1)(1 - T) - T\gamma = 0,$$

Combined with  $\dot{\lambda}_i = 0$ , this yields the equilibrium:

$$\begin{aligned}\lambda_1 &= \frac{r + (1 - T)\gamma}{r + \gamma}, \\ \lambda_2 &= \frac{T\gamma}{r + \gamma}, \\ C'(K_1) &= \left( \frac{r}{1 - T} + \gamma \right) \left( \frac{r + \beta}{r + \gamma} \right).\end{aligned}$$

Hence, the transversality conditions imply that the co-state variables must be constant on this path, which implies that also  $K_1$  is constant, so this is a steady state path. Remains the question if this steady state can be maintained until infinity. This can be done by investing  $I(t) = \beta K_1^*$ , where  $K_1^*$  satisfies (14). For these investments one has  $C(K_1^*) > I$ , so  $D > 0$ , since  $C'(K_1) > \beta$ .

So one obtains a steady state with  $\dot{K}_1 = \dot{\lambda}_1 = \dot{\lambda}_2 = 0$ .

Moreover, suppose that on path  $1_A$ ,  $\lambda_2(t) < T\gamma/(r + \gamma)$  (resp.  $>$ ). Then (29) implies that  $\lambda_2$  is decreasing (resp. increasing) so that  $\lambda_2(u) < T\gamma/(r + \gamma)$  (resp.  $>$ ) for  $u \geq t$ . A similar argument holds for  $\lambda_1$ , since  $\lambda_1 + \lambda_2 = 1$ . This implies that if the firm is on path  $1_A$ , but not in a steady state (it does not satisfy the transversality conditions), then it cannot satisfy the transversality conditions without leaving path  $1_A$ .

It now remains to show that the other paths cannot be the final paths. For path  $1_B$ , the only equilibrium is:

$$\begin{aligned}\lambda_1 &= 1, \\ \lambda_2 &= 0, \\ C'(K_1) &= r + \beta,\end{aligned}$$

so that the equilibrium conditions imply that  $\dot{K}_1 = 0$ . Consequently, one has

$$\dot{K}_2(t) = \beta K_1^* - \gamma K_2(t).$$



It is verified easily that this implies that:

$$K_2(t) = K_2(t^*)e^{-\gamma(t-t^*)} + (1 - e^{-\gamma(t-t^*)})\frac{\beta}{\gamma}K_1^*,$$

where  $t^*$  denotes the time instant where the steady state is reached. It then follows immediately that:

$$\lim_{t \rightarrow \infty} \gamma K_2(t) = \beta K_1^*,$$

so that in the long run  $C(K_1^*) - \gamma K_2$  will be positive, i.e. the B-path is abandoned.

This argument holds for all paths  $i_B$ , independent of the values of  $\lambda_i$ . The transversality conditions could be satisfied, but when the time-horizon is long enough it is not possible to end in a situation where no taxes are paid.

Finally, notice that on path 2, one has  $\dot{K}_1 > 0$ , and on path 3, one has  $\dot{K}_1 < 0$ . Therefore it is seen that an equilibrium is not possible. The transversality conditions can be satisfied for path 2 (3) only with  $\eta_1 = 0$  ( $\eta_2 = 0$ ). However, then the firm is on path 1 again. We can therefore conclude that the final path is  $1_A$ , and that this is a steady state path.  $\square$

It now remains to determine which paths can be coupled before each other. Proposition 4.1 gives useful information for this path-coupling procedure, since it implies that in the optimal solution, a B-path can never be preceded by an A-path.

**Proof of Proposition 4.1:** We need to show that: If in the optimal solution one has  $C(K_1) - \gamma K_2 \leq 0$  at some time  $\tau$ , then  $C(K_1) - \gamma K_2 < 0$  at any time  $t < \tau$ . This means that zero tax payments at time  $\tau$  imply zero tax payments at any time  $t < \tau$ .

Define the auxiliary function  $g(I) = \frac{\partial}{\partial t}(C(K_1) - \gamma K_2)$ . We find:

$$g(I) = C'(K_1)\dot{K}_1 - \gamma\dot{K}_2 = (C'(K_1) - \gamma)I + \gamma^2 K_2 - C'(K_1)\beta K_1,$$

which is linear in the investments  $I$ . This implies that

$$g(0) \leq g(I) \leq g(C(K_1)) \quad \text{if } C'(K_1) \geq \gamma,$$

$$\text{and } g(0) > g(I) > g(C(K_1)) \quad \text{if } C'(K_1) < \gamma.$$

Given that  $\gamma \geq \beta$ ,  $C(\cdot)$  is concave, and  $C(K_1) \leq \gamma K_2$  (B-path), it holds that:

$$\begin{aligned} g(0) &= \gamma(\gamma K_2) - \beta C'(K_1)K_1, \\ &> \gamma C(K_1) - \beta C(K_1), \\ &\geq 0, \end{aligned}$$

and

$$\begin{aligned} g(C(K_1)) &= C'(K_1)C(K_1) - \gamma C(K_1) + \gamma[\gamma K_2] - C'(K_1)\beta K_1, \\ &\geq C'(K_1)C(K_1) - \gamma C(K_1) + \gamma C(K_1) - C'(K_1)\beta K_1, \\ &= C'(K_1)[C(K_1) - \beta K_1], \\ &> 0. \end{aligned}$$

This implies that  $g(I) > 0$  for all  $I \in [0, C(K_1)]$ , so that  $C(K_1) - \gamma K_2$  is strictly increasing over time when the firm is on a B path.  $\square$

The above proposition, combined with an elaborate path-coupling procedure, allows to prove that there are two master trajectories.

A master trajectory is a maximal sequence of policies that can be applied in order to obtain an optimal dynamic solution. Depending on the initial state of the firm, the optimal solution starts at a certain point in one of the master trajectories. The following definition will be helpful.

**Definition A.1** *We define  $\tilde{\gamma}$  to be the fixed point of the equation  $C'(K_1^*(\gamma)) = \gamma$ .*

*Given (14), this implies that:*

$$\tilde{\gamma} = \frac{1}{2}\beta + \sqrt{\frac{1}{4}\beta^2 + \frac{(r + \beta)r}{1 - T}}.$$

Straightforward calculations yield:

$$\tilde{\gamma} > \frac{1}{2}\beta + \sqrt{\frac{1}{4}(\beta + 2r)^2} = \beta + r,$$

so that  $\tilde{\gamma} \notin [\beta, \beta + r]$ .

We now proceed to determine the master trajectories. The following notation will be used. The instant at which path  $i$  is coupled before path  $j$  will be denoted  $t_{i,j}$

(irrespective of whether they are  $A$  or  $B$  paths). The time instant just before this coupling will be denoted  $t_{i,j}^-$ .

**Proposition A.3** *The following is a master trajectory:*

$$MT_1 : 2_B \rightarrow 2_A \rightarrow 1_A.$$

**Proof:** In order to show that a sequence of paths is a master trajectory, one has to show that the paths can be coupled before each other without violating the optimality conditions, and that no other path can precede the first path in the trajectory. We will therefore subsequently show that

- i) Path  $2_A$  can be coupled before path  $1_A$ .
- ii) Path  $2_B$  can be coupled before the sequence  $2_A \rightarrow 1_A$ .
- iii) Nothing can be coupled before the sequence  $2_B \rightarrow 2_A$ .

[ad *i*)] The fact that  $\lambda_1$  and  $\lambda_2$  are continuous implies that, in order to couple path  $2_A$  before  $1_A$ , one must have  $\eta_1(t_{2,1}) = 0$ .

[ad *ii*)] We first describe the dynamics on path  $2_A$ , given that it is coupled before  $1_A$ .

Path  $2_A$  has  $I = (1 - T)C(K_1) + \gamma TK_2 \geq (1 - T)C(K_1)$ . Condition (13) therefore implies that  $\dot{K}_1 > 0$ . Since  $C(\cdot)$  is concave, this implies that  $C'(K_1)$  is decreasing over time on path  $2_A$ .

$\dot{\lambda}_1, \dot{\lambda}_2$  and therefore  $\dot{\eta}$  are continuous on  $2_A \rightarrow 1_A$ . Furthermore, one has  $\eta_2(t_{2,1}) = 0$ . Since  $\dot{\lambda}_1(t_{2,1}) = 0$ ,  $\dot{\lambda}_1$  is continuous, and  $C'(\cdot)$  decreases over time, it follows that  $\dot{\lambda}_1(t_{2,1}^-) \leq 0$ .

Furthermore, since on path  $2_A$ , one has  $\eta_1 > 0$ , we know that  $\eta_1(t_{2,1}^-) > 0$ . This, together with (36), implies that  $\dot{\lambda}_2(t_{2,1}^-) \leq 0$ . From (37), it now follows that  $\dot{\eta}_1(t_{2,1}^-) = \dot{\lambda}_1(t_{2,1}^-) + \dot{\lambda}_2(t_{2,1}^-) \leq 0$ , and  $C'(K_1)$  decreases over time.

Furthermore,  $\dot{\lambda}_i$  remain non-increasing, since if  $\dot{\lambda}_i$  would be zero at time  $\tau$  because of the change in  $\lambda_i$ ,  $\dot{\lambda}_i(\tau^-) \leq 0$ , since  $\dot{\eta}_1(\tau^-) = \dot{\lambda}_1(\tau^-) + \dot{\lambda}_2(\tau^-) \leq 0$ .

So we can conclude that on path  $2_A$  before the final path,  $\lambda_1, \lambda_2$  and  $\eta_1$  are non-increasing.

The fact that  $\dot{\eta}_1 < 0$  on path  $2_A$  implies that path  $2_B$  can precede  $2_A$ . Since at the coupling point  $t_{2,2}$ , one has  $C(K_1) - \gamma K_2 = 0$ , the costate variables do not have to be differentiable at  $t_{2,2}$ . Instead, (26) and (27) apply.

[ad iii)] In order to determine what can precede path  $2_B$ , we look at the dynamics of the costates. On path  $2_B$ , one has:

$$\dot{\eta}_1^B = (\gamma - \beta)\lambda_2 - (1 + \eta_1)[C'(K_1) - (r + \beta)]. \quad (49)$$

Given that:

- (39) implies that  $\lambda_2$  is increasing on path  $2_B$ ,
- $C'(K_1) > r + \beta$  (in the subsequent paths the firm will grow towards  $K_1^*$ , with  $C'(K_1^*) > r + \beta$ ),
- (13) implies that  $C'(K_1)$  is decreasing on path  $2_B$  ( $\dot{K}_1 > 0$ ),

it follows from (49) that, if  $\dot{\eta}_1^B(\tau) < 0$  at some time  $\tau$ , then  $\dot{\eta}_1^B(t) < 0$  for all  $t < \tau$ . Proposition 4.1 implies that only a B-path can precede path  $2_B$ . Then, in order to couple another path before path  $2_B$ ,  $\eta_1^B$  has to be 0 at that coupling instant  $t_{.,2}$ . The above then implies that  $\dot{\eta}_1^B(t_{2,2})$  has to be positive. Indeed, suppose that  $\dot{\eta}_1^B(t_{2,2}) < 0$ , then  $\dot{\eta}_1^B(t_{.,2}) < 0$ . Now if  $\eta_1^B(t_{.,2}) = 0$ , this is clearly impossible.

To see whether it is possible that  $\dot{\eta}_1^B(t_{2,2}) \geq 0$ , we consider the relation between dynamics of  $\eta_1$  on path  $2_A$  and  $2_B$ . On path  $2_A$  one has:

$$\dot{\eta}_1^A = \dot{\lambda}_1 + \dot{\lambda}_2 = (\gamma - \beta)\lambda_2 - (1 + \eta_1)[(1 - T)C'(K_1) + T\gamma - (r + \beta)] \quad (50)$$

Combined with (49), this implies:

$$\dot{\eta}_1^B = \dot{\eta}_1^A - (1 + \eta_1)T(C'(K_1) - \gamma). \quad (51)$$

Now since  $\dot{\eta}_1^A(t_{2,2}) \leq 0$ ,  $\dot{\eta}_1^B(t_{2,2})$  can only be positive if  $C'(K_1) < \gamma$ .

At path  $2_A$  and  $2_B$  on this master trajectory, one has  $C'(K_1) > C'(K_1^*)$ . Therefore  $\gamma > C'(K_1)$  implies  $\gamma > C'(K_1^*)$ . This only holds when  $\gamma > \tilde{\gamma}$ . Path  $2_B$  can therefore not be preceded by another path for  $\gamma \in [\beta, r + \beta]$ .

This concludes the proof. □

**Proposition A.4** *The following is a master trajectory:*<sup>4</sup>

$$MT_2 : 2_B \rightarrow 1_B \rightarrow 3_B \rightarrow 3_A \rightarrow 1_A.$$

**Proof:** Similarly to the proof of A.3, we will show that:

- i) Path  $3_A$  can be coupled before path  $1_A$ .
- ii) Path  $3_B$  can be coupled before  $3_A \rightarrow 1_A$ .
- iii) Path  $1_B$  can be coupled before path  $3_B - 3_A - 1_A$  if  $C'(K_1) \geq r + \beta$ .
- iv) Path  $2_B$  can be coupled before path  $1_B - 3_B - 3_A - 1_A$ .
- v) Nothing can be coupled before  $2_B \rightarrow 1_B$  or  $2_B \rightarrow 3_B$ .

[ad i)] When coupling  $3_A$  before the terminating path, it holds that  $\eta_2(t_{3,1}) = 0$ .

[ad ii)] We first describe the dynamics on path  $3_A$  before  $1_A$ . Since the dynamics of  $\lambda_2$  are the same on paths  $3_A$  and  $1_A$ , one has  $\dot{\lambda}_2 = 0$  during path  $3_A$  (see (43)). Furthermore, since path  $3_A$  is a shrink-path in capital stock, one has  $\frac{\partial C'(K_1)}{\partial t} = C''(K_1)\dot{K}_1 > 0$ , which, with (42), implies that  $\dot{\lambda}_1 > 0$ . So  $\dot{\eta}_2 < 0$  and only path  $3_B$  can precede path  $3_A$ .

The coupling  $3_B \rightarrow 3_A$  is quite trivial. At the coupling instant, one has  $C(K_1) - \gamma K_2 = 0$ , and  $\eta_2 > 0$ .

---

<sup>4</sup>Strictly speaking, the sequence  $1_B \rightarrow 3_B$  can also be  $1_B \rightarrow 3_B \rightarrow 1_B \rightarrow 3_B \rightarrow \dots \rightarrow 1_B \rightarrow 3_B$ , which is still a shrinking phase. This however can be excluded for functions  $C(\cdot)$  with  $C'''(\cdot) \leq 0$ , e.g. functions in the class  $C(K_1) = aK_1 - bK_1^2$ .

[ad *iii*) and *iv*)] We first show that the sequence  $3_B \rightarrow 3_A$  can only be preceded by another path if  $C'(K_1) \geq r + \beta$ .

At the coupling point  $t_{.,3}$ ,  $\eta_2$  has to be zero, and at a certain point after the coupling instant  $\eta_2$  has to be positive (in order to couple path  $3_B$  before  $3_A$ ). Therefore,  $\lambda_1(t_{.,3}) + \lambda_2(t_{.,3}) = 1$ , and  $\dot{\eta}_2(t_{.,3}) \geq 0$ , which implies for path  $3_B$ :

$$\begin{aligned}\dot{\eta}_2 &= -\dot{\lambda}_1 - \dot{\lambda}_2, \\ &= -(r + \beta)\lambda_1 + C'(K_1) - (r + \gamma)\lambda_2, \\ &\leq -(r + \beta) + C'(K_1).\end{aligned}$$

Therefore,  $\dot{\eta}_2(t_{.,3}) \geq 0$  iff  $C'(K_1) \geq r + \beta$  at time  $t_{.,3}$ .

Therefore, when a path is coupled before path  $3_B$ , one has  $C'(K_1) \geq r + \beta$ .

Proposition 4.1 implies that only a B-path can precede path  $3_B$ . Notice that the coupling  $2_B \rightarrow 3_B$  is as a special case of  $2_B \rightarrow 1_B \rightarrow 3_B$ , with path  $1_B$  followed for an infinitesimal small time period.

It can be verified easily that the coupling of  $1_B$  before  $3_B$  is feasible iff  $C'(K_1) \leq r + \beta$  at time  $t_{1,3}$ . The coupling of  $2_B$  before  $1_B$  is quite trivial, since the continuity of co-states is clearly maintained, and only the investment strategy changes.

[ad *vi*)] We show that when path  $2_B$  is coupled before another B-path, it cannot be preceded by another path.

Preceding paths ( $1_B$  or  $3_B$ ) have  $\dot{\lambda}_1 + \dot{\lambda}_2 \leq 0$ . This implies that  $C'(K_1) \geq r + \beta$ .

We then find from (38), (39) and (37), together with the fact that  $2_B$  is a growth path, that just before the coupling instant, one has  $\dot{\eta}_1(t_{.,2}^-) < 0$ . This implies:

$$\begin{aligned}\dot{\eta}_1(t_{.,2}^-) &= (r + \beta)\lambda_1 - (1 + \eta_1)C'(K_1) + (r + \gamma)\lambda_2, \\ &= (\gamma - \beta)\lambda_2 - (1 + \eta_1)[C'(K_1) - (r + \beta)], \\ &< 0.\end{aligned}$$

Furthermore,

$$\frac{\partial \dot{\eta}_1}{\partial t} = [\gamma - \beta]\dot{\lambda}_2 - [1 + \eta_1]C''(K_1)\dot{K}_1 - \dot{\eta}_1 [C'(K_1) - (r + \beta)]. \quad (52)$$

This is clearly positive, so  $\eta_1$  has increased to its negative value. Therefore, it is negative along the path. This implies that only a path with  $\eta_1 > 0$  can precede path  $2_B$ , but proposition 4.1 implies that path  $2_A$  can not precede path  $2_B$ .

This concludes the proof.  $\square$

In the following proposition, we prove that there are no other master trajectories, so that the optimal investment strategies are determined.

**Proposition A.5** *MT<sub>1</sub> and MT<sub>2</sub> are the only master trajectories of problem (9)*

**Proof:** In order to prove this proposition, it is necessary to show that:

- i) There is no part of  $MT_1$  or  $MT_2$  that can be preceded by paths that are not in these master trajectories, and
- ii) There are no other possible couplings before the final path  $1_A$ .

In the following these two points will be addressed.

[ad i)] In the proofs of propositions A.3 and A.4, it is made clear that the couplings in  $MT_1$  are unique and that the couplings for  $MT_2$  are unique taking into account footnote (4). There were no other couplings possible than the ones that resulted in the master trajectories.

[ad ii)] In order to prove this part, all paths other than  $2_A$  and  $3_A$  must be proven to be not feasible before the final path  $1_A$ .

- *Path  $1_B$  before Path  $1_A$ :* Path  $1_B$  is a shrinkpath. When coupling path  $1_B$  before  $1_A$  at time  $t_{1,1}$ ,  $\dot{\lambda}_1$  and  $\dot{\lambda}_2$  can be discontinuous (see (26) and (27)).  $K_1$  however is continuous. Therefore at time  $t_{1,1}^-$  we have  $C'(K_1) = C'(K_1^*)$ . Given (33), this implies:

$$\begin{aligned}
& \dot{\lambda}_1 + \dot{\lambda}_2 = 0, \\
\Rightarrow & (r + \beta) \frac{r+(1-T)\gamma}{r+\gamma} - \frac{r+\beta}{1-T} \frac{r+(1-T)\gamma}{r+\gamma} + T\gamma = 0, \\
\Rightarrow & -T \frac{r+\beta}{1-T} \frac{r+(1-T)\gamma}{r+\gamma} = -T\gamma, \\
\Rightarrow & \gamma = \tilde{\gamma}.
\end{aligned}$$

Since  $\tilde{\gamma} > r + \beta$ , the proof is complete.  $\square$

- *Path  $2_B$  before  $1_A$ :* Path  $2_B$  can precede path  $1_A$ . Using the fact that  $\dot{\lambda}_1 + \dot{\lambda}_2 = \dot{\eta}_1 < 0$ , the proof is similar to the case  $1_B \rightarrow 1_A$ .

Furthermore we know that nothing can precede  $2_B \rightarrow 1_A$ . This implies that  $2_B \rightarrow 1_A$  could be a master trajectory. This master trajectory is a special case of master trajectory 1, that starts with path  $2_B$  and has path  $2_A$  for an infinitesimal small interval. However it is just one special case, dependent on all starting values and variables, such that full investments result in:

$$\begin{aligned} C(K_1) &= \gamma K_2, \\ C'(K_1) &= C'(K_1^*), \end{aligned}$$

at the coupling instant  $t^*$ . Therefore we neglect this possibility, since it is a boundary case and a special case of  $MT_1$ .

- *Path  $3_B$  before  $1_A$ :* This coupling is not feasible. Using the fact that  $\dot{\lambda}_1 + \dot{\lambda}_2 = \dot{\eta}_1 < 0$ , the proof is similar to the case  $1_B \rightarrow 1_A$ .

This completes the proof  $\square$

The above propositions immediately lead to the following theorem.

**Theorem A.1** *The two master trajectories and therefore optimal solutions are:*

$$MT_1 : 2_B \rightarrow 2_A \rightarrow 1_A,$$

$$MT_2 : 2_B \rightarrow 1_B \rightarrow 3_B \rightarrow 3_A \rightarrow 1_A.$$

**Proof:** Follows immediately from propositions A.3, A.4, and A.5.  $\square$

Depending on the initial state of the firm, its optimal strategy will start at some point in one of the three master trajectories. If for example the firm's taxable income is positive in the initial state, its optimal strategy starts at the  $A$  part of the



master trajectory. For firms that initially have marginal costs larger than marginal revenues, the optimal strategy starts on path 3 of the second master trajectory. If the firm pays taxes it starts on path  $3_A$ , and otherwise on path  $3_B$ .

**Proof of Proposition 4.2:** Follows immediately from Theorem A.1 and the definition of low initial tax depreciation. For firms with low initial tax depreciation master trajectory 1 applies, since there taxable income is positive by the time path 1 is reached.

**Proof of Proposition 4.3:** Follows immediately from Theorem A.1 and the definition of high initial tax depreciation. For firms with high initial tax depreciation master trajectory 2 applies, since there taxable income is negative by the time path 1 is first reached.

## B The optimal size and the value of the firm

This Appendix contains the proofs of the propositions that are stated in sections 2 and 3.

**Proof of Proposition 2.1:** Given that the firm is in the steady state at time  $z$ , we know that  $I(t) = \beta K_1(z)$ , for all  $t \geq z$ . This, together with (2) implies that the evolution of the tax base is given by:

$$K_2(t) = K_2(z)e^{-\gamma(t-z)} + (1 - e^{-\gamma(t-z)})\frac{\beta}{\gamma}K_1(z).$$

The value of the firm at time  $z$  therefore equals:

$$\begin{aligned} f(K_1(z), K_2(z)) &= \int_z^\infty e^{-r(t-z)} \left\{ (1-T)C(K_1(z)) - \beta K_1(z) + T\gamma K_2(t) \right\} dt \\ &= \frac{1-T}{r} C(K_1(z)) - \frac{\beta}{r} K_1(z) \\ &\quad + \gamma T \int_z^\infty \left( e^{-r(t-z)} e^{-\gamma(t-z)} K_2(z) + e^{-r(t-z)} (1 - e^{-\gamma(t-z)}) \frac{\beta}{\gamma} K_1(z) \right) dt \\ &= \frac{1-T}{r} C(K_1(z)) - \frac{\beta}{r} K_1(z) + \frac{\gamma\beta T}{r(r+\gamma)} K_1(z) + \frac{\gamma}{r+\gamma} T K_2(z). \end{aligned}$$

Rearranging the terms leads to (7). □

**Proof of Proposition 3.1:** Follows immediately from the proof of proposition A.2. We now show that  $K_1^*$  is the unique level at which, in a steady state, marginal revenue equals marginal cost and taxable income is positive.

Consider the firm in a steady state at time  $t^*$ . This implies that investments equal  $I(t) = \beta K_1(t^*)$  for all  $t \geq t^*$ . We now determine the marginal value of an additional investment at time  $t^*$ . Due to the extra investment  $x$  at time  $t^*$ , the evolution over time of the capital stock and the tax base after  $t^*$  is given by:

$$\begin{aligned}\tilde{K}_1(t, x) &= K_1(t^*) + xe^{-\beta(t-t^*)}, \\ \tilde{K}_2(t, x) &= K_2(t) + xe^{-\gamma(t-t^*)}.\end{aligned}$$

The marginal value generated by the additional investment is then given by:

$$\begin{aligned}\frac{\partial}{\partial x}|_{x=0} \int_{t^*}^{\infty} e^{-r(t-t^*)} \left[ (1-T)C(K_1(t^*) + xe^{-\beta(t-t^*)}) + T\gamma(K_2(t) + xe^{-\gamma(t-t^*)}) \right] dt \\ = \int_{t^*}^{\infty} e^{-r(t-t^*)} \left[ (1-T)C'(K_1(t^*))e^{-\beta(t-t^*)} + T\gamma e^{-\gamma(t-t^*)} \right] dt \\ = \frac{1}{\beta+r} (1-T)C'(K_1(t^*)) + \gamma T \frac{1}{\gamma+r}.\end{aligned}$$

It now follows that the latter expression equals 1 (the marginal cost of the additional investment), iff

$$C'(K_1(t^*)) = \left( \frac{r}{1-T} + \gamma \right) \left( \frac{r+\beta}{r+\gamma} \right).$$

Since  $C'(\cdot)$  is strictly decreasing, this implies that  $K_1(t^*) = K_1^*$ .  $\square$

**Proof of proposition 3.2:**

*i)* Consider the case where  $r > 0$ . The optimal stock level of the firm satisfies:

$$C'(K_1^*) = \left( \frac{r}{1-T} + \gamma \right) \left( \frac{r+\beta}{r+\gamma} \right).$$

Straightforward calculations yield:

$$\begin{aligned}\frac{\partial}{\partial \gamma} C'(K_1^*) &= \frac{-rT}{(1-T)(r+\gamma)} \frac{(r+\beta)}{(r+\gamma)} \\ &< 0.\end{aligned}$$

Combined with the fact that

$$\frac{\partial}{\partial \gamma} C'(K_1^*) = C''(K_1^*) \frac{\partial}{\partial \gamma} K_1^*,$$

and  $C''(\cdot) < 0$ , we can conclude that  $\frac{\partial K_1^*}{\partial \gamma} > 0$ .

*ii), iii), iv)* These statements can be verified in a similar way. □

## References

- [1] DEVEREUX, M.P., M. KEEN, AND F. SCHIANTARELLI (1994) "Corporation Tax Asymmetries and Investment: Evidence from UK Panel Data", *Journal of Public Economics* 53, 395-418.
- [2] DIXIT, A.K. AND R.S. PINDYCK (1994), *Investment under Uncertainty*, Princeton University Press.
- [3] FAIG, M., AND P. SHUM (1999) "Irreversible Investment and Endogenous Financing: An evaluation of the Corporate Tax Effects", *Journal of Monetary Economics* 43, 143-171.
- [4] FEICHTINGER, G., AND R.F. HARTL (1986), *Optimale Kontrolle ökonomischer Prozesse*, Walter de Gruyter.
- [5] GENTRY, W. (1994) "Taxes, Financial Decisions, and Organizational Form: Evidence from Publicly Traded Partnerships", *Journal of Public Economics* 53, 223-244.
- [6] HARTL, R.F., AND P.M. KORT (1996), "Capital Accumulation of a Firm Facing Environmental Constraints", *Optimal Control: Applications and Methods*, 17, 253-266.
- [7] VAN HILTEN O., P.M. KORT, AND P.J.J.M. VAN LOON (1993), *Dynamic Policies of the Firm: An Optimal Control Approach*, Springer, Berlin.
- [8] JACOBS, F.A., AND E.R. LARKINS (1998) "Export Tax Incentives for Establishing Foreign Markets: An analysis of Marginal Costing Techniques", *Accounting Horizons* 12, 374-396.
- [9] NEWBERRY, R. (1998) "Foreign Tax Credit Limitations and Capital Structure Decisions", *Journal of Accounting Research* 36, 157-166.
- [10] PEREIRA, A.M. (1994) "On the Effects of Investment Tax Credits on Economic Efficiency and Growth", *Journal of Public Economics* 53, 437-461.

- [11] SANSING, R. (1998) "Valuing the Deferred Tax Liability", *Journal of Accounting Research* 36, 357-363.
- [12] SHOLES, M.S., AND M.A. WOLFSON (1992), *Taxes and Business Strategy: A Planning Approach*, Prentice-Hall, New Jersey.
- [13] WAKEMAN, L.M. (1980). "Optimal Tax Depreciation". *Journal of Accounting and Economics* 1, 213-237.