Optimal tax depreciation lives and charges under regulatory constraints

De Waegenaere, A.M.B.; Wielhouwer, J.L.

Published in:
OR Spectrum

Publication date:
2002

OR Spectrum (2002) 24: 151–177

OR Spectrum

Optimal tax depreciation lives and charges under regulatory constraints*

Anja De Waegenaere[1] **and Jacco L. Wielhouwer**[2]

[1] Department of Econometrics and Operations Research, and CentER for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands (e-mail: a.m.b.DeWaegenaere@kub.nl)

[2] CentER for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands (e-mail: J.L.Wielhouwer@kub.nl)

Received: January 9, 2001 / Accepted: Feburary 25, 2002

Abstract. Depreciation is not only a representation of the loss in asset-value over time. It is also a strategic tool for management and can be used to minimize tax payments. In this paper we derive the depreciation scheme that minimizes the expected value of the present value of future tax payments for two types of constraints on the depreciation method. We show how the optimal scheme depends on the discount factor and the cash flow distributions. Moreover, we find the somewhat surprising result that the way in which the optimum is affected by uncertainty depends crucially on the type of regulatory constraint.

Key words: Tax minimization – Depreciation – Discounting – Uncertainty – Dynamic optimization – Path-coupling

1 Introduction

It is well-known that regulation and legislation on corporate taxation leave ample room for strategic behavior of firms. Scholes and Wolfson (1992) provide a thorough overview of the different opportunities for firms to minimize tax expenses through business strategy. An important way to shift income is through depreciation of the firm's assets. Since taxable income consists of cash-flows reduced with depreciation charges, one can shift taxable income from one period to another by depreciating more or less in a certain period, while keeping the total amount to be depreciated over all periods fixed. Consequently, different depreciation schemes can yield a different stream of future taxable income. The decision maker can try to optimize

* We thank Menachem Berg for raising the issue of optimization problems related to tax depreciation.

Correspondence to: A. De Waegenaere

by choosing – among those methods that are accepted by the tax authorities – the depreciation method that minimizes the expected present value of future taxable income.

The development of the research on optimal tax depreciation can be seen as follows. Wakeman (1980) compares accelerated and straight line depreciation and shows that, when taxable income is known and positive, accelerated depreciation dominates straight line depreciation, in the sense that it yields a lower expected value of discounted tax payments for all values of the discount rate. Berg and Moore (1989) consider a 2-period model and show how uncertainty can affect this dominance of accelerated depreciation methods. Berg et al. (2001) provide an analysis of the optimal choice between accelerated and straight line depreciation with uncertain cash-flows and a possibly progressive tax system.

In this paper we do not compare two given methods of tax depreciation, but determine the optimal tax depreciation scheme among those that are accepted by the tax authorities. Within the limitations set by the tax authority, we optimize with respect to both the number of periods the asset should be depreciated in, and the corresponding depreciation charges in each period. We show how this optimal depreciation scheme depends on the discount factor and the probability distributions of future cash-flows. In order to take into account that the tax authority does not accept every possible depreciation scheme, we consider two sets of depreciation schemes. The first set contains all depreciation schemes for which the fraction of the residual value that is depreciated has to lie within certain bounds. The constraints on the depreciation charge in a certain period then clearly depend on the depreciation charges chosen in earlier periods. Commonly used methods that focus at the residual value of the asset are the so-called *Declining Balance* methods, where in each period a given fraction of the residual value is depreciated. The second set contains all depreciation methods for which the amount depreciated in a period has to lie within certain bounds. Here, the constraints in a certain period are clearly independent of decisions made in earlier periods. An example is the *Straight Line* method, where the amount depreciated is equal over all periods. In the sequel the two types of constraints will be referred to as *dynamic constraints* and *static constraints*, respectively.

The paper is organized as follows. Section 2 defines the optimization problems for the two types of constraints described above. In Section 3 we reformulate the optimization problem with dynamic constraints as a dynamic program. We then show that the path-coupling method, which is developed to solve continuous time optimization problems, yields valuable insights when applied to this discrete time optimization problem. We show that a depreciation scheme satisfies the necessary conditions for optimality iff the last non-zero depreciation charge is the unique strictly positive root of a decreasing function, where the depreciation charges in all other periods are specific functions of the last non-zero depreciation charge and its period. Therefore, there are at most N candidate optimal depreciation schemes, where N is the maximal number of periods in which the asset has to be depreciated. The optimal scheme is then found by evaluating all candidate optimal solutions. Section 4 derives the optimal solution in case of static constraints. Also here, one finds at most N candidate optimal solutions by determining the unique root of a decreasing function. As opposed to the case with dynamic constraints however, it can be shown that the optimal depreciation scheme is the candidate optimal scheme in which the number of periods over which the asset is depreciated is maximal or, equivalently, the optimal depreciation scheme is a candidate optimal scheme with the longest depreciation life. There is therefore no need to evaluate all the candidate optimal solutions. Section 5 provides analytical results on the effect of the discount rate and the cash-flow distributions on the optimal scheme. We show that the effect of a change in the cash flow distribution in a certain period depends on whether the firm faces dynamic or static constraints. We present some numerical examples in Section 6. In Section 7 we summarize the differences between the static and dynamic constraints and discuss the implications for the regulator and the firm. The paper is concluded in Section 8.

2 The optimization problems

An asset of value D has to be depreciated over a maximum of N periods. Let d_k denote the amount depreciated in period k. The decision maker has to decide on the number of periods ($\leq N$) that will actually be used to depreciate the value D (i.e. the last k with $d_k > 0$), and the corresponding depreciation charges.

The cash-flow or income in period k (gross revenue before depreciation) is a random variable denoted C_k, with cumulative distribution function $F_k(.)$. We assume that cash-flows are continuously distributed, so that $F_k(.)$ is continuous and strictly increasing.[1]

The decision maker's objective is to minimize the expected present value of future tax payments. With a fixed tax rate T over all taxable income[2], and a discount rate $\alpha \in [0, 1]$, this leads to the following optimization problem.

$$\min_{(d_1,\ldots,d_N)\in\mathcal{D}} T \sum_{k=1}^{N} \alpha^k E\left[(C_k - d_k)^+\right], \tag{1}$$

where $x^+ := \max\{x, 0\}$, and $\mathcal{D}$ is the set of acceptable depreciation methods. One can distinguish two types of depreciation methods :

i) Methods with *dynamic constraints,* i.e. constraints on the depreciation charge as a fraction of the remaining value of the asset, so that

$$\mathcal{D} = \left\{ (d_1, \ldots, d_N) \in I\!R_+^N \;\middle|\; \begin{array}{l} \sum_{k=1}^{N} d_k = D \\ d_k \in [l_k D_{k-1}, u_k D_{k-1}] \end{array} \right\}, \tag{2}$$

[1] Notice that cash flows need not be independent. It is therefore possible that cash flows result from a stochastic process, in which case $F_k(\cdot)$ denotes the marginal distribution function.

[2] The effects of a progressive tax system are studied in Berg et al. (2001) and Wielhouwer et al. (2000). It is shown there that a progressive tax system provides an incentive to smoothen taxable income in order to avoid higher tax brackets. Smoothing taxable income results in a lower marginal tax rate.

with $0 \leq l_k < u_k \leq 1$ for all $k = 1, \ldots, N$. Here, $D_{k-1} = D - \sum_{i=1}^{k-1} d_i$ is the residual value to be depreciated in periods k until N, so that $D_0 = D$.

ii) Methods with *static constraints,* i.e. constraints on the value of the depreciation charges d_k, so that

$$\mathcal{D} = \left\{ (d_1, \ldots, d_N) \in I\!R_+^N \;\middle|\; \begin{array}{l} \sum_{k=1}^N d_k = D \\ d_k \in [\tilde{l}_k, \tilde{u}_k] \end{array} \right\}, \tag{3}$$

with $0 \leq \tilde{l}_k < \tilde{u}_k \leq D$, for all $k = 1, \ldots, N$.

In some cases a solution to problem (1) is found easily. Suppose for example that cash-flows are known with certainty, and that the constraint set equals:

$$\mathcal{D} = \left\{ (d_1, \ldots, d_N) \in I\!R_+^N \mid \sum_{k=1}^N d_k = D \right\}. \tag{4}$$

It is seen immediately that an optimal scheme is given by:

$$\begin{aligned} d_k &= \max\{C_k, 0\}, \text{ if } \quad C_k \leq D_{k-1}, \\ &= D_{k-1}, \qquad\quad \text{if } \quad C_k \geq D_{k-1}. \end{aligned} \tag{5}$$

for $k = 1, \ldots, N-1$, and $d_N = D - \sum_{j=1}^{N-1} d_j$.

Indeed, since due to the discounting effect ($\alpha \leq 1$) paying taxes later is preferable to paying them now, one should depreciate "as much as possible as early as possible", but never more than the actual cash-flow if there is still at least one period to come.

In the more interesting case where future cash-flows are unknown, or where the set of acceptable depreciation schemes is a strict subset of (4), an analytical solution is not found easily. In the next section we present the solution for the case of dynamic constraints.

3 The dynamic constraints

The constraints in (2) imply that the fraction of the residual value to be depreciated is subject to limitations. Commonly used methods that determine the depreciation based on the residual value of the asset are the so-called declining balance methods.

Instead of determining the optimal $(d_1, \ldots, d_N)$, one can then determine the optimal fraction $\gamma_k \in [l_k, u_k]$ of the residual value D_{k-1} to depreciate in period k, so that $d_k = \gamma_k D_{k-1}$, where:

$$D_k = D - \sum_{j=1}^k d_j, \qquad \text{for } k \leq N. \tag{6}$$

Since our aim is also to determine the optimal number of periods in which D is depreciated, we consider the case where $u_k = 1$, so that $\gamma_k \in [l_k, 1]$. It is clear that without loss of generality, we can set $T = 1$. With the expected values written as their corresponding integral, the problem to solve then is:

$$\begin{aligned} &\min_{(\gamma_1,\ldots,\gamma_N)} \sum_{k=1}^{N} \alpha^k \int_{\gamma_k D_{k-1}}^{\infty} (1 - F_k(y))dy \\ &s.t.\ D_k = (1-\gamma_k)D_{k-1}, \\ &\qquad D_0 = D, \\ &\qquad \gamma_k \in [l_k, 1]. \end{aligned} \tag{7}$$

Now, if $(\gamma_1, \ldots, \gamma_N)$ solves (7), the optimal depreciation charges are given by $d_k = \gamma_k D_{k-1}$, and the optimal number of periods used to depreciate the asset equals $J = \min\{k : \gamma_k = 1\}$.

In the sequel we use the *current-value Hamiltonian* and the *path-coupling method* (see e.g. Feichtinger and Hartl, 1986, pp. 504-509; Van Hilten et al., 1993) to determine the solution of problem (7). We proceed as follows. In Section 3.1 we describe the current-value Hamiltonian and the current-value Lagrangian, and state the necessary conditions for optimality. In Section 3.2, we define the paths and describe their dynamics. In Section 3.3, we characterize the set of solutions that satisfy the necessary conditions for optimality, and we show how the optimal solution can be found.

3.1 The necessary conditions

The *current-value Hamiltonian* for problem (7) is given by:

$$H(\delta, \gamma, \lambda, k) = -\int_{\delta\gamma}^{\infty} (1 - F_k(y))dy + \lambda(1-\gamma)\delta, \tag{8}$$

where δ (resp. γ) is the state (resp. control) variable, and λ is the co-state variable. To incorporate the condition $\gamma_k \in [l_k, 1]$, we define the *current-value Lagrangian* of this problem as follows:

$$L(\delta, \gamma, \lambda, \eta^1, \eta^2, k) = H(\delta, \gamma, \lambda, k) + \eta^1(\gamma - l_k) + \eta^2(1-\gamma). \tag{9}$$

Then the necessary conditions for optimality are given by the following system of equations:

$$\lambda_N = 0, D_0 = D, \tag{10}$$

and, for $k = 1, \ldots, N$:[3]

$$(1 - F_k(\gamma_k D_{k-1}))D_{k-1} - \lambda_k D_{k-1} + \eta_k^1 - \eta_k^2 = 0, \tag{11}$$

$$\lambda_{k-1} = \alpha(1 - F_k(\gamma_k D_{k-1}))\gamma_k + \alpha\lambda_k(1-\gamma_k), \tag{12}$$

$$D_k = (1-\gamma_k)D_{k-1}, \tag{13}$$

$$\eta_k^1(\gamma_k - l_k) = 0, \tag{14}$$

$$\eta_k^2(1-\gamma_k) = 0, \tag{15}$$

$$\eta_k^1, \eta_k^2 \geq 0, \gamma_k \in [l_k, 1]. \tag{16}$$

[3] The first two equations express the conditions $\frac{\partial L}{\partial \gamma}(D_{k-1}, \gamma_k, \lambda_k, \eta_k^1, \eta_k^2, k) = 0$, and $\alpha \frac{\partial L}{\partial \delta}(D_{k-1}, \gamma_k, \lambda_k, \eta_k^1, \eta_k^2, k) = \lambda_{k-1}$, respectively.

Since the conditions in (10) -(16) are necessary conditions for an optimum, it is natural to introduce the following definition.

Definition 1. *A depreciation scheme* $(d_1, \ldots, d_N)$ *is a candidate optimal solution if there exist variables* $\gamma_k, \lambda_k, \eta_k^1, \eta_k^2$*, and* D_k *that satisfy (10)-(16), such that for all* $k \leq N$*, one has* $d_k = D_{k-1} - D_k = \gamma_k D_{k-1}$.

We now analyze the set of candidate optimal solutions, using the path-coupling method. To do so, we first define the paths and describe their dynamics.

3.2 The paths

Consider a certain time period k, with a residual depreciable value D_{k-1}. Then there are four different paths that can be followed to the next period, as can be seen in the following table:

Table 1. The paths

	1	2	3	4
η_k^1	0	> 0	0	> 0
η_k^2	0	0	> 0	> 0

(17)

We say that path $i \in \{1, \ldots, 4\}$ is *feasible in period* k if there exists a solution to (10)-(16) in which the values for η_k^1 and η_k^2 satisfy the conditions for path i as given in Table 1. Path 4 is clearly never feasible, since (14) and (15) would then imply that $\gamma_k = l_k = 1$, which is clearly a contradiction. In order to study the other three paths, we use the following lemma.

Lemma 1. *Consider the case where* $D_{k-1} > 0$ *and* $\lambda_k > 0$*, and define:*

$$\tilde{\gamma}_k = \frac{1}{D_{k-1}} F_k^{-1} \left(1 - \lambda_k\right). \tag{18}$$

Then,

- *path 1 is feasible in period* k *iff* $\tilde{\gamma}_k \in [l_k, 1]$,
- *path 2 is feasible in period* k *iff* $\tilde{\gamma}_k < l_k$*, and*
- *path 3 is feasible in period* k *iff* $\tilde{\gamma}_k > 1$.

Proof. By definition, $\tilde{\gamma}_k$ is the unique solution of the equation:

$$\frac{\partial}{\partial \gamma} H(D_{k-1}, ., \lambda_k, k) = 0. \tag{19}$$

It is seen immediately that $\frac{\partial}{\partial \gamma} H(D_{k-1}, \gamma, \lambda_k, k)$ is strictly decreasing in γ. Then, path 1 is not feasible (i.e. the unique solution $\tilde{\gamma}_k$ of (19) is such that $\tilde{\gamma}_k \notin [l_k, 1]$) iff

$$\begin{aligned} &\tfrac{\partial}{\partial \gamma} H(D_{k-1}, l_k, \lambda_k, k) < 0 \;\Leftrightarrow\; \tilde{\gamma}_k < l_k, \\ \text{or } &\tfrac{\partial}{\partial \gamma} H(D_{k-1}, 1, \lambda_k, k) > 0 \;\Leftrightarrow\; \tilde{\gamma}_k > 1. \end{aligned} \tag{20}$$

This concludes the proof. □

We now evaluate the dynamics of the three feasible paths.

– *Path 1:* This path is characterized by $\eta_k^1 = \eta_k^2 = 0$. This is feasible when $D_{k-1} = 0$, or when $\tilde{\gamma}_k \in [l_k, 1]$. When $D_{k-1} \neq 0$, solving (10)-(16) yields that $\gamma_k = \tilde{\gamma}_k$, and

$$\alpha\lambda_k = \lambda_{k-1}. \tag{21}$$

When $D_{k-1} = 0$, there are infinitely many solutions to (10)-(16).

– *Path 2:* This path is characterized by $\eta_k^1 > 0$ and $\eta_k^2 = 0$. This implies that the minimum amount is depreciated in period k, i.e. $\gamma_k = l_k$. It is only feasible when $D_{k-1} > 0, \lambda_k > 0$, and $\tilde{\gamma}_k < l_k$. The dynamics of the co-state are

$$\lambda_{k-1} = \alpha l_k(1 - F_k(l_k D_{k-1})) + \alpha\lambda_k(1 - l_k). \tag{22}$$

– *Path 3:* This path is characterized by $\eta_k^1 = 0$ and $\eta_k^2 > 0$. This implies that everything left in period k is depreciated, i.e. $\gamma_k = 1$. It is feasible when $D_{k-1} > 0, \lambda_k > 0$, and $\tilde{\gamma}_k > 1$, or when $D_{k-1} > 0$ and $\lambda_k \leq 0$. The dynamics of the co-state are

$$\lambda_{k-1} = \alpha(1 - F_k(D_{k-1})). \tag{23}$$

Notice that (11) implies that paths 2 and 3 can only be feasible when $D_{k-1} > 0$. Notice furthermore that, when path 1 is feasible for $\gamma_k = 1$, then the dynamics of the co-state are as in (23).

3.3 The optimal solution

In this section we derive the optimal solution using the path-coupling method. First we characterize the set of candidate optimal depreciation schemes. For any depreciation scheme, we denote J for the last period in which the depreciation charge is non-zero, i.e. $\hat{d} = (d_1, \ldots, d_J, 0, \ldots, 0)$ with $d_J > 0$.

Due to the fact that the objective function in (7) is not strictly convex in $(\gamma_1, \ldots, \gamma_N, D_0, \ldots, D_{N-1})$, there is in general not a unique candidate optimal depreciation scheme. However, in the sequel we show that for any given value of J, there will be at most one candidate optimal solution. This candidate optimal solution equals the optimal depreciation scheme, *given that* exactly J periods are used to depreciate the asset. In general, several values of J will yield a depreciation scheme that satisfies the necessary conditions, but there will be a unique value of J that yields the optimal scheme.

In order to characterize the set of candidate optimal depreciation schemes, we introduce the following definition. Intuitively this definition should be interpreted as the solution of the difference equations for D_k and λ_k, given that the total amount is depreciated in J periods.

Definition 2. *Consider the following recursive definition:*

$$D_{k-1} := \max\Big\{\frac{D_k}{1 - l_k},\ D_k + F_k^{-1}\big(1 - \lambda_k\big)\Big\}, \tag{24}$$

and

$$\lambda_{k-1} := \alpha \min\bigl\{\lambda_k, (1 - F_k(l_k D_{k-1}))l_k + (1 - l_k)\lambda_k\bigr\}. \tag{25}$$

Then, if $D_{J-1} > 0$ and $\lambda_{J-1} > 0$ are given, D_k and λ_k can be determined recursively for all $k = J-2, \dots, 0$. Moreover, we define

$$\Psi_J(d) := D - D_0(d, J), \tag{26}$$

where $D_0(d, J) = D_0$ determined by (24) and (25) with $D_{J-1} = d$ and $\lambda_{J-1} = \alpha\bigl(1 - F_J(d)\bigr)$.

The above definition shows how the candidate optimal solution can be calculated, once the values of D_{J-1} and λ_{J-1} are known.

The following theorem provides necessary and sufficient conditions for $\hat{d} = (d_1, \dots, d_N)$ to be a candidate optimal depreciation scheme.

Theorem 1. *A depreciation scheme $\hat{d} = (d_1, \dots, d_J, 0, \dots, 0)$ with $d_J > 0$ satisfies (10)-(16) iff*

- *$\hat{d}$ satisfies:*

$$\begin{cases} d_J \in \Psi_J^{-1}(0), \\ d_k = \max\bigl\{l_k D_{k-1}\ ,\ F_k^{-1}\bigl(1 - \lambda_k\bigr)\bigr\}, \text{ for all } k \leq J-1. \end{cases} \tag{27}$$

  where D_k and λ_k, for $k = 1, \dots, J-2$, are determined by (24) and (25) with $D_{J-1} = d_J$ and $\lambda_{J-1} = \alpha\bigl(1 - F_J(d_J)\bigr)$, and
- *$d_J \leq F_J^{-1}(1 - \lambda_J^*)$, where*

$$\begin{cases} \lambda_k^* := \alpha \min_{\gamma \in [l_{k+1}, 1]} \left\{\gamma(1 - F_{k+1}(0)) + (1 - \gamma)\lambda_{k+1}^*\right\}, \\ \qquad\qquad k = 1, \dots, N-1, \\ \lambda_N^* := 0 \end{cases} \tag{28}$$

Proof. See Appendix. □

The above theorem implies that all candidate optimal depreciation schemes can be found by solving $\Psi_J(.) = 0$, for $J = 1, \dots, N$. Then, for any J for which $\Psi_J(.)$ has a root $d_J \in (0, D]$ that satisfies $d_J \leq F_J^{-1}(1 - \lambda_J^*)$, there exists a candidate optimal depreciation scheme for which the depreciation charges are given by (27).

The following proposition states that $\Psi_J(.)$ is a decreasing function, so that its root can be found easily. Moreover, the depreciation charge d_J is the *unique* solution of $\Psi_J(.) = 0$. Combined with (27), this yields at most N candidate optimal schemes.

Proposition 1. *The function $\Psi_J(.)$ is decreasing. Moreover, $\Psi_J(.)$ has a non-negative root iff $\Psi_J(\hat{u}_J) \leq 0 \leq \Psi_J(0)$, where $\hat{u}_J = (1-l_1)(1-l_2)\cdots(1-l_{J-1})D$.*

Proof. It is clear that $\Psi_J(.)$ is decreasing iff $D_0(., J)$ is increasing, where $D_0(d, J) = D_0$ determined by (24) and (25) with $D_{J-1} = d$ and $\lambda_{J-1} = \alpha\big(1 - F_J(d)\big)$.

We now show by induction that D_k is increasing in d and that λ_k is decreasing in d for all $k = 0, \ldots, J-1$.

The above statements are trivially satisfied for $k = J-1$. Moreover, it follows immediately from (24) and (25) that, if the statements are satisfied for k, they are also satisfied for $k-1$.

Finally, the fact that the root is less than or equal to $\hat{u}_J$ follows immediately from

$$\begin{aligned} & d = D_{J-1} \\ \Rightarrow\ & d \leq (1 - l_{J-1}) D_{J-2} \\ \Rightarrow\ & d \leq (1 - l_1)(1 - l_2) \cdots (1 - l_{J-1}) D_0(d, J) = \hat{u}_J. \end{aligned}$$

This concludes the proof. □

Theorem 1 and Proposition 1 imply that there are at most N candidate optimal schemes. The following result shows how the set of potential candidates can be further decreased.

Proposition 2. *If a depreciation scheme* $(d_1, \ldots, d_J, 0, \ldots, 0)$ *with* $d_J > 0$ *is optimal, then* d_J *satisfies*

$$d_J \leq F_J^{-1}(1 - \alpha(1 - F_{J+1}(0))), \qquad \textit{if } J \leq N-1. \tag{29}$$

Proof. See Appendix. □

Since the objective function is strictly convex in $(d_1, \ldots, d_N)$, and the constraint set $\mathcal{D}$ is compact, there is a unique optimal scheme. In order to find the unique optimal depreciation scheme, one can proceed as follows. For every $J \in \{1, \ldots, N\}$:

i) Check whether $\Psi_J(u_J^*) \leq 0 \leq \Psi_J(0)$, where $u_N^* = \hat{u}_N$ and $u_J^* = \min\{\hat{u}_J, F_J^{-1}(1 - \alpha(1 - F_{J+1}(0)))\}$ for $J < N$.
ii) If so, calculate $d_J = \Psi_J^{-1}(0)$.
iii) Evaluate the objective function in the resulting depreciation scheme given in (27).

Notice that it is not necessary to calculate λ_J^*, since the condition $d_J \leq F_J^{-1}(1 - \lambda_J^*)$ can be replaced by the stronger condition (29).

4 The static constraints

In this section we determine the optimal depreciation charges in case of static constraints. For ease of notation, we consider the case where $d_k \in [\tilde{l}_k, +\infty)$, and without loss of generality assume that $T = 1$. The problem to be solved is then:

$$\begin{aligned} & \min_{(d_1, \ldots, d_N)} \ \textstyle\sum_{k=1}^N \alpha^k \int_{d_k}^{\infty} (1 - F_k(y)) dy \\ & s.t. \ \textstyle\sum_{k=1}^N d_k = D, \\ & \qquad d_k \geq \tilde{l}_k, \quad \text{for } k = 1, \ldots, N. \end{aligned} \tag{30}$$

For any depreciation scheme $\hat{d}$, we denote J for the last period in which the depreciation charge strictly exceeds the lower bound, i.e. $\hat{d} = (d_1, \ldots, d_J, \tilde{l}_{J+1}, \ldots, \tilde{l}_N)$ with $d_J > \tilde{l}_J$.

Similarly to the case with dynamic constraints, we define the functions $\tilde{d}_k(d, J)$, which can be interpreted as the optimal depreciation charges given that the depreciation charge in period J equals d, and that J is the last period where the lower bound is not binding.[4]

Definition 3. *For all $J \leq N$, and $k \leq J-1$, we define:*

$$\tilde{d}_k(d, J) := \max\left\{\tilde{l}_k\ ,\ F_k^{-1}\left(1 - \alpha^{J-k}(1 - F_J(d))\right)\right\}, \quad k \leq J-1, \tag{31}$$

$$\tilde{\Psi}_J(d) := D - d - \sum_{k=1}^{J-1} \tilde{d}_k(d, J) - \sum_{k=J+1}^{N} \tilde{l}_k, \tag{32}$$

$$\mathcal{P} := \left\{k \in \{1, \ldots, N\} \,\middle|\, \tilde{\Psi}_k(\tilde{l}_k) \geq 0\right\}. \tag{33}$$

In the following theorem we show that in the optimal solution, the last depreciation charge that exceeds the lower bound is the unique root of $\tilde{\Psi}_J(\cdot)$, which is a strictly decreasing function, and all other depreciation charges are given functions of this depreciation charge and its period J. More precisely, we have the following result:

Theorem 2. *The optimal depreciation scheme satisfies:*

$$\begin{cases} d_J \in \tilde{\Psi}_J^{-1}(0), \\ d_k = \tilde{d}_k(d_J, J), \text{ for } k \leq J-1, \\ d_k = \tilde{l}_k, \qquad\quad \text{ for } k \geq J+1. \end{cases} \tag{34}$$

for some $J \in \mathcal{P}$. Moreover, the function $\tilde{\Psi}_J(.)$ is strictly decreasing.

Proof. It is clear that also this problem can be stated as a dynamic problem as in (7), but with the constraints replaced by

$$\gamma_k D_{k-1} \geq \tilde{l}_k, \qquad k = 1, \ldots, N. \tag{35}$$

The necessary conditions for optimality therefore are:

$$\lambda_N = 0, D_0 = D, \tag{36}$$

[4] The reason why the solution of the difference equation for λ_k is not stated in this definition (contrary to the dynamic case), is that we can find a closed form expression for λ_k as a function of J and d so that they do not have to be determined recursively. Therefore, we can immediately state the optimal depreciation charges given period J and its depreciation charge d. This is elaborated upon in the proof of Theorem 2.

and, for $k = 1, \dots, N$:

$$
\begin{aligned}
&(1 - F_k(\gamma_k D_{k-1}))D_{k-1} - \lambda_k D_{k-1} + \eta_k^1 D_{k-1} = 0,\\
&\lambda_{k-1} = \alpha(1 - F_k(\gamma_k D_{k-1}))\gamma_k + \alpha\lambda_k(1 - \gamma_k) + \alpha\eta_k^1\gamma_k,\\
&D_k = (1 - \gamma_k)D_{k-1},\\
&\eta_k^1(\gamma_k D_{k-1} - \tilde{l}_k) = 0,\\
&\gamma_k D_{k-1} \geq \tilde{l}_k,\\
&\eta_k^1 \geq 0, \gamma_k \in [0, 1].
\end{aligned}
$$

Therefore, the proof is similar to the proofs of Theorem 1 and Proposition 1. However, notice that the fact that $d_J = \gamma_J D_{J-1} \geq \tilde{l}_J$ implies that $\eta_J^1 = 0$, so that

$$\lambda_J = 1 - F_J(d_J). \tag{37}$$

Moreover, the dynamics of Path 2 are now equal to those of Path 1. This implies that (25) can now be replaced by

$$\lambda_k = \alpha^{J-k}\lambda_J = \alpha^{J-k}(1 - F_J(d_J)), \qquad k = 1, \dots J - 1. \tag{38}$$

This yields the desired result. □

As opposed to the case with dynamic constraints, it can be shown that out of the set of candidate optimal solutions, the optimal solution is the one in which J is maximal.

We need the following lemma.

Lemma 2. *Let* $(d_1, \dots, d_J, \tilde{l}_{J+1}, \dots, \tilde{l}_N)$ *be a solution that satisfies (34) for some* $J \in \{1, \dots, N\}$. *Then for every* $k \leq J$*, one has:*

i) $\min\{\alpha^k\big(1-F_k(\tilde{l}_k)\big), \alpha^J\big(1-F_J(d_J)\big)\}(d_k-\tilde{l}_k) = \alpha^J\big(1-F_J(d_J)\big)(d_k-\tilde{l}_k)$,

ii) $\min\{\alpha^k(1 - F_k(\tilde{l}_k)), \alpha^J(1 - F_J(d_J))\} = \alpha^k\big(1 - F_k(d_k)\big)$.

Proof. First notice that for any $k, J \leq N$ and $x \in I\!R$, one has:

$$F_k^{-1}\big(1-\alpha^{J-k}(1-F_J(x))\big) \geq \tilde{l}_k \;\Leftrightarrow\; \alpha^k\big(1-F_k(\tilde{l}_k)\big) \geq \alpha^J\big(1-F_J(x)\big). \tag{39}$$

i) Follows from the fact that (31), (34), and (39) imply that if $d_k > \tilde{l}_k$ then

$$\min\{\alpha^k\big(1 - F_k(\tilde{l}_k)\big), \alpha^J\big(1 - F_J(d_J)\big)\} = \alpha^J\big(1 - F_J(d_J)\big). \tag{40}$$

ii) Is trivially satisfied for $k = J$. For all $k < J$, one has:

$$
\begin{aligned}
\alpha^k\big(1 - F_k(d_k)\big) &= \alpha^k\big(1-F_k\left(\max\{\tilde{l}_k, F_k^{-1}\left(1-\alpha^{J-k}(1-F_J(d_J))\right)\}\right)\big)\\
&= \alpha^k\big(1 - \max\{F_k(\tilde{l}_k), 1 - \alpha^{J-k}(1 - F_J(d_J))\}\big)\\
&= \alpha^k \min\{1 - F_k(\tilde{l}_k), \alpha^{J-k}(1 - F_J(d_J))\}\\
&= \min\{\alpha^k(1 - F_k(\tilde{l}_k)), \alpha^J(1 - F_J(d_J))\}.
\end{aligned} \tag{41}
$$

This concludes the proof. □

The following proposition states that the optimal depreciation scheme is the one in which J is maximal. Therefore, as opposed to the case with dynamic constraints, there is no need to evaluate all the candidate optimal solutions.

Proposition 3. *The optimal depreciation scheme satisfies (34) for*

$$J = \max\{k : k \in \mathcal{P}\}. \tag{42}$$

Proof. Notice that $J \in \mathcal{P}$ iff the allocation defined in (34) exists and satisfies $d_J \geq \tilde{l}_J$. It therefore suffices to show that if $J, K \in \mathcal{P}$, and $K < J$, then the allocation as defined in (34) for J yields a lower value of the objective function than the one for K. Let us denote d^J and d^K for the corresponding candidate solutions, i.e.

$$\begin{cases} d_k^J = \max\{\tilde{l}_k, F_k^{-1}\left(1 - \alpha^{J-k}(1 - F_J(d_J^J))\right)\}, & k = 1, \ldots, J, \\ d_k^J = \tilde{l}_k, & k = J+1, \ldots, N, \end{cases} \tag{43}$$

since $d_J^J \equiv F_J^{-1}\left(1 - \alpha^{J-J}(1 - F_J(d_J^J))\right)$, and, equivalently,

$$\begin{cases} d_k^K = \max\{\tilde{l}_k, F_k^{-1}\left(1 - \alpha^{K-k}(1 - F_K(d_K^K))\right)\}, & k = 1, \ldots, K, \\ d_k^K = \tilde{l}_k, & k = K+1, \ldots, N. \end{cases} \tag{44}$$

Then, the *difference* in objective function (expected discounted taxable income) for d^J and d^K is given by::

$$\begin{aligned} &\textstyle\sum_{k=1}^N \alpha^k E\left[\left(C_k - d_k^J\right)^+\right] - \sum_{k=1}^N \alpha^k E\left[\left(C_k - d_k^K\right)^+\right] \\ &= \textstyle\sum_{k=1}^K \alpha^k E\left[\left(C_k - d_k^J\right)^+\right] + \sum_{k=K+1}^J \alpha^k E\left[\left(C_k - d_k^J\right)^+\right] \\ &\qquad - \textstyle\sum_{k=1}^K \alpha^k E\left[\left(C_k - d_k^K\right)^+\right] - \sum_{k=K+1}^J \alpha^k E\left[(C_k - \tilde{l}_k)^+\right] \\ &= \textstyle\sum_{k=1}^K \alpha^k \int_{d_k^J}^{d_k^K} \left(1 - F_k(u)\right) du - \sum_{k=K+1}^J \alpha^k \int_{\tilde{l}_k}^{d_k^J} \left(1 - F_k(u)\right) du \\ &\leq \textstyle\sum_{k=1}^K \alpha^k \left(1 - F_k(d_k^J)\right)\left(d_k^K - d_k^J\right) - \sum_{k=K+1}^J \alpha^k \left(1 - F_k(d_k^J)\right)\left(d_k^J - \tilde{l}_k\right). \end{aligned} \tag{45}$$

Applying lemma 2 ii) to the last expression then yields that:

$$\begin{aligned} &\textstyle\sum_{k=1}^N \alpha^k E\left[\left(C_k - d_k^J\right)^+\right] - \sum_{k=1}^N \alpha^k E\left[\left(C_k - d_k^K\right)^+\right] \\ &< \textstyle\sum_{k=1}^K \min\{\alpha^k\left(1 - F_k(\tilde{l}_k)\right), \alpha^J\left(1 - F_J(d_J^J)\right)\}\left(d_k^K - \tilde{l}_k - (d_k^J - \tilde{l}_k)\right) \\ &\qquad - \textstyle\sum_{k=K+1}^J \min\{\alpha^k\left(1 - F_k(\tilde{l}_k)\right), \alpha^J\left(1 - F_J(d_J^J)\right)\}\left(d_k^J - \tilde{l}_k\right) \\ &\leq \textstyle\sum_{k=1}^K \alpha^J\left(1 - F_J(d_J^J)\right)\left(d_k^K - \tilde{l}_k\right) - \sum_{k=1}^J \alpha^J\left(1 - F_J(d_J^J)\right)\left(d_k^J - \tilde{l}_k\right) \\ &= \alpha^J\left(1 - F_J(d_J^J)\right)\left(\textstyle\sum_{k=1}^K d_k^K + \sum_{k=K+1}^J \tilde{l}_k - \sum_{k=1}^J d_k^J\right) \\ &= 0, \end{aligned} \tag{46}$$

where the second inequality follows from lemma 2 i) and from replacing the minimum by one of its components, and the last equality follows from the fact that d^J and d^K are feasible, and therefore both have components that until period J add up to $D - \sum_{k=J+1}^{N} \tilde{l}_k$. This concludes the proof. □

The above proposition implies that the optimal solution can be found by determining maximal J for which the root of $\tilde{\Psi}_J(.)$ is strictly larger than $\tilde{l}_J$. The optimal depreciation charges are then given by (34).

Notice finally that, whereas in the case of dynamic constraints, J equals the number of periods in which the asset is depreciated, this is no longer necessarily the case here, since d_k is bounded below by $\tilde{l}_k$, so that $d_{J+i} > 0$ if $\tilde{l}_{J+i} > 0$.

5 Effect of distributions and discount rate

In Section 5.1 we study the effect of the discount factor on the optimal depreciation charges. We find that a lower discount factor works in favor of a more accelerated depreciation scheme. In Section 5.2 we study how a change in the distribution function of a cash flow affects the optimal depreciation scheme.

5.1 The effect of the discount rate

To focus on the effect of the discount rate, we assume all cash flows to be equally distributed. Let us denote d_{SL} for the straight line depreciation method, i.e. $d_{SL} = (\frac{D}{N}, \ldots, \frac{D}{N})$, so that the amount to depreciate is divided equally over all periods. In the next theorem we show that d_{SL} is optimal when all cash-flows are equally distributed, there is no discounting, and $d_{SL} \in \mathcal{D}$.

Theorem 3. *If $F_k(.) = F(.)$ for all periods k, $\alpha = 1$, and $d_{SL} \in \mathcal{D}$, then d_{SL} is optimal.*

Proof. Since $d_{SL} \in \mathcal{D}$, it suffices to show that d_{SL} is optimal for problem (30) with $\tilde{l}_1 = \ldots = \tilde{l}_N = 0$.

The fact that $F^{-1}(F(d_N)) = d_N$, implies:

$$\max\{0, F_k^{-1}(F_N(d_N)))\} = d_N = \frac{D}{N}, \qquad \text{for all } k. \tag{47}$$

Therefore, the depreciation charges in d_{SL} satisfy (34) with $J = N$. Given Theorem 2 and Proposition 3, this yields the desired result. □

We now show that, when cash-flows are equally distributed, $\alpha < 1$, and the constraints are such that $l_1 \geq l_2 \geq \cdots \geq l_N$ (resp. $\tilde{l}_1 \geq \tilde{l}_2 \geq \cdots \geq \tilde{l}_N$), then the optimal depreciation method with dynamic (resp. static) constraints is an accelerated depreciation method.

Theorem 4. *When $F_k(.) = F(.)$ for all k, $\alpha < 1$, and $l_1 \geq l_2 \geq \cdots \geq l_N$ (resp. $\tilde{l}_1 \geq \tilde{l}_2 \geq \cdots \geq \tilde{l}_N$), then the optimal depreciation method with dynamic (resp. static) constraints satisfies $d_1 > d_2 > \ldots > d_J$.*

Proof. Consider the case of dynamic constraints. We know from Theorem 1 that the optimal depreciation scheme is such that:

$$d_k = \max\{l_k D_{k-1}, F^{-1}(1-\lambda_k)\},$$
$$\lambda_{k-1} = \min\{\alpha\lambda_k, \alpha(1-F(l_k D_{k-1}))l_k + \alpha(1-l_k)\lambda_k\},$$

for all $k \leq J-1$, and

$$\lambda_{J-1} = \alpha\left(1 - F(D_{J-1})\right). \tag{48}$$

Notice now that $l_k D_{k-1}$ is decreasing in k, and λ_k is strictly increasing in k, due to $\alpha < 1$. This implies that

$$d_{k+1} < d_k, \qquad \text{for all } k = 1, \ldots, J-2. \tag{49}$$

Furthermore, it is seen immediately that $l_J \leq 1$ and $D_{J-1} = d_J$ imply that

$$d_J = \max\left\{l_J D_{J-1}, F^{-1}(1-(1-F(d_J))\right\}. \tag{50}$$

Therefore, since $1 - F(d_J) = 1 - F(D_{J-1}) > \lambda_{J-1}$ it follows that $d_J < d_{J-1}$, so we can conclude that depreciation is accelerated.

In case of static constraints, the proof is similar. □

In conclusion one finds that a higher discount rate (so lower interest rate) implies a less accelerated method. When the discount rate equals one, the optimal depreciation scheme is straight line.

5.2 The effect of the distribution functions

In this section we show that the effect of a change in the cash flow distribution in a certain period depends crucially on the type of regulatory constraints. Therefore, we present two subsections, one for each type of constraint. In both cases, we introduce a change in the cash flow distribution in period κ, and keep the other cash flow distributions equal. For notational convenience, we present results for the case where the new distribution function assigns higher probability to low cash flow levels. It will be clear from the proofs, however, that similar results can be derived for the opposite case.

5.2.1 Dynamic constraints. We consider a candidate optimal solution, $(d_1^*, \ldots, d_J^*, 0, \ldots, 0)$, that depreciates the initial tax base in exactly J periods. The corresponding state and co-state variables are denoted D_k and λ_k respectively. We introduce a change in the cash flow distribution in period κ, and denote $(\hat{d}_1, \ldots, \hat{d}_J, 0, \ldots, 0)$ for the resulting candidate optimal solution.

When the cash flow distribution in a certain period is changed such that there is a higher probability on low income in that period, then one might intuitively expect that the depreciation charge in that period would decrease and the depreciation charges in all other periods would (weakly) increase. In the following proposition we show that this is not necessarily the case. The intuition for this result is given in Section 7.

Proposition 4. *If for a certain period $\kappa < J$ the cash flow distribution function $F_\kappa(\cdot)$ is replaced by $\hat{F}_\kappa(\cdot)$ such that $\hat{F}_\kappa(x) \geq F_\kappa(x)$ for all $x \leq d^*_\kappa$, then the following holds:*

i) If in the current optimal solution path 1 is applied in period κ, and the new distribution function satisfies $\hat{F}_\kappa\left(\frac{l_\kappa D_\kappa}{1-l_\kappa}\right) \leq 1 - \lambda_\kappa$ then:

$$\begin{aligned} &\hat{d}_k \geq d^*_k && \textit{for all } k > \kappa, \\ &\textstyle\sum_{k=1}^{\kappa} \hat{d}_k \leq \sum_{k=1}^{\kappa} d^*_k. \end{aligned} \tag{51}$$

ii) If in the current optimal solution path 2 is applied in period κ, then:

$$\begin{aligned} &\hat{d}_k \leq d^*_k && \textit{for all } k \geq \kappa, \\ &\textstyle\sum_{k=1}^{\kappa-1} \hat{d}_k \geq \sum_{k=1}^{\kappa-1} d^*_k. \end{aligned} \tag{52}$$

iii) If in the current optimal solution path 1 is applied in period κ and $\hat{F}_\kappa\left(\frac{l_\kappa D_\kappa}{1-l_\kappa}\right) > 1 - \lambda_\kappa$, then either (51) or (52) will result.

Proof. Let us denote $d_k(d, J), k = 1, \ldots, J-1$ for the depreciation charges that result from (24), (25) and (27) with $D_{J-1} = d$ and $\lambda_{J-1} = \alpha(1 - F_J(d))$. Similarly, $\hat{d}_k(d, J), k = 1, \ldots, J-1$ denote these depreciation charges when $F_\kappa(\cdot)$ is replaced by $\hat{F}_\kappa(\cdot)$.

Before we can start with the proofs of i), ii), and iii), we make the following three observations.

First, notice that it follows from Theorem 1 that $d^*_k = d_k(d^*_J, J)$ and $\hat{d}_k = \hat{d}_k(\hat{d}_J, J)$ for all $k = 1, \ldots, J-1$. Moreover, by construction

$$\sum_{k=1}^{J-1} \hat{d}_k(\hat{d}_J, J) + \hat{d}_J = D_0 = \sum_{k=1}^{J-1} d_k(d^*_J, J) + d^*_J. \tag{53}$$

Second, notice that $\hat{d}_k(d, J)$ is increasing in d for all $k = 1, \ldots, J-1$. Indeed, take any $\tilde{d} > d$. Then it follows from (24) and (25) that $\tilde{\lambda}_k < \lambda_k$ and $\tilde{D}_k > D_k$ for all $k < J$. Consequently it follows from (27) that $\hat{d}_k(\tilde{d}, J) > \hat{d}_k(d, J)$.

Third, the fact that $\hat{F}_\kappa(x) \geq F_\kappa(x)$ for all $x \leq d^*_\kappa$ implies that $\hat{F}_\kappa^{-1}(y) \leq F_\kappa^{-1}(y)$ for all $y \leq F_\kappa(d^*_\kappa)$. Therefore, since $F_\kappa^{-1}(1-\lambda_\kappa) \leq d^*_\kappa$, it follows that

$$\hat{F}_\kappa^{-1}(1-\lambda_\kappa) \leq F_\kappa^{-1}(1-\lambda_\kappa). \tag{54}$$

The above three observations, combined with (24) and (25) yield that:

i) If the optimal path in period κ is path 1 and $\hat{F}_\kappa\left(\frac{l_\kappa D_\kappa}{1-l_\kappa}\right) \leq 1 - \lambda_\kappa$, then:

$$\begin{aligned} &\hat{d}_k(d^*_J, J) = d_k(d^*_J, J) \text{ for all } k > \kappa, \\ &\hat{d}_k(d^*_J, J) \leq d_k(d^*_J, J) \text{ for all } k \leq \kappa. \end{aligned} \tag{55}$$

This implies that $\sum_{k=1}^{J-1} \hat{d}_k(d_J^*, J) + d_J^* \leq D_0$. Now since $\hat{d}_k(d, J)$ is increasing in d and, by definition, $\sum_{k=1}^{J-1} \hat{d}_k(\hat{d}_J, J) + \hat{d}_J = D_0$, it follows that $\hat{d}_J \geq d_J^*$. Therefore it follows from (55) that $\hat{d}_k(\hat{d}_J, J) \geq d_k(d_J^*, J)$ for all $k > \kappa$. With (53) we conclude that $\sum_{k=1}^{\kappa} \hat{d}_k \leq \sum_{k=1}^{\kappa} d_k^*$.

ii) If the optimal path in period κ is path 2, then:

$$\begin{aligned} \hat{d}_k(d_J^*, J) &= d_k(d_J^*, J) \text{ for all } k \geq \kappa, \\ \hat{d}_k(d_J^*, J) &\geq d_k(d_J^*, J) \text{ for all } k < \kappa. \end{aligned} \tag{56}$$

This implies that $\sum_{k=1}^{J-1} \hat{d}_k(d_J^*, J) + d_J^* \geq D_0$, so that $\hat{d}_J \leq d_J^*$. Therefore one has from (56) that $\hat{d}_k(\hat{d}_J, J) \leq d_k(d_J^*, J)$ for $k \geq \kappa$, which implies (52).

iii) Similar to the proofs of i) and ii).

This completes the proof. □

The above proposition shows that the effect of a change in the distribution function depends crucially on the path that is optimal in that period. A particular change in the distribution function in period κ can lead to either an increase or a decrease of the optimal depreciation charge in any period $k \neq \kappa$. Moreover, even when the lower bound is binding in period κ, the change in the distribution function can affect the optimal solution. In the sequel we see that this is an important difference with the static case. There, a higher probability on low income always implies that the depreciation charge decreases in period κ and increases in all other periods. When the lower bound is binding in period κ, the optimal depreciation charges are unaffected by the change.

5.2.2 Static constraints. This section deals with the effect of a change in the cash flow distribution in period κ in case of static constraints. We show that, as opposed to the dynamic case, a decrease in the marginal contribution to the objective function in a certain period does imply that the depreciation charge in that period decreases and all other depreciation charges increase. As in the dynamic case, the optimal solution before (resp. after) any changes in the distribution function is denoted $(d_1^*, \ldots, d_N^*)$ (resp. $(\hat{d}_1, \ldots, \hat{d}_N)$).

Proposition 5. *If for a certain period $\kappa < J$ the cash flow distribution function $F_\kappa(\cdot)$ is replaced by $\hat{F}_\kappa(\cdot)$ such that $\hat{F}_\kappa(x) \geq F_\kappa(x)$ for all $x \leq d_k^*$, it holds that:*

i) If $d_\kappa^* > \tilde{l}_\kappa$ then

$$\begin{aligned} &\hat{d}_k \geq d_k^* \quad \textit{for all } k \neq \kappa, \\ &\hat{d}_\kappa \leq d_\kappa^*. \end{aligned} \tag{57}$$

ii) If $d_\kappa^* = \tilde{l}_\kappa$ then $\hat{d}_k = d_k^*$ for all $k = 1, ..., N$.

Proof. The Lagrangian for problem (30) is:

$$L = \sum_{k=1}^{N} \alpha^k \int_{d_k}^{\infty} (1 - F_k(y))dy - \mu(D - \sum_{k=1}^{N} d_k) - \sum_{k=1}^{N} \eta_k(d_k - \tilde{l}_k). \tag{58}$$

Let us define the following functions:

$$\begin{aligned} d_k(\mu) &= \max\{\tilde{l}_k, F_k^{-1}(1 - \alpha^{-k}\mu)\}, \text{ for all } k. \\ \hat{d}_\kappa(\mu) &= \max\{\tilde{l}_\kappa, \hat{F}_\kappa^{-1}(1 - \alpha^{-\kappa}\mu)\}, \\ \hat{d}_k(\mu) &= d_k(\mu) \qquad \text{for all } k \neq \kappa. \end{aligned} \tag{59}$$

Then it follows from the necessary Karush-Kuhn-Tucker conditions that there exists a μ^* and a $\hat{\mu}$, such that:

$$\begin{aligned} d_k^* &= d_k(\mu^*) \qquad \text{for all } k. \\ \hat{d}_k &= \hat{d}_k(\hat{\mu}) \qquad \text{for all } k. \end{aligned} \tag{60}$$

Then it follows that:

i) in case $d_\kappa > \tilde{l}_\kappa$:

$$\begin{aligned} \hat{d}_k(\mu^*) &= d_k(\mu^*) \qquad \text{for } k \neq \kappa, \\ \hat{d}_\kappa(\mu^*) &= \hat{F}_\kappa^{-1}(1 - \alpha^{-\kappa}\mu^*) \leq F_\kappa^{-1}(1 - \alpha^{-\kappa}\mu^*) = d_\kappa(\mu^*), \end{aligned} \tag{61}$$

so that $\sum_{k=1}^{N} \hat{d}_k(\mu^*) \leq \sum_{k=1}^{N} d_k(\mu^*) = D_0 = \sum_{k=1}^{N} \hat{d}_k(\hat{\mu})$. Now since clearly, $\hat{d}_k(\mu)$ is decreasing in μ, one can conclude that $\hat{\mu} \leq \mu$ which implies (57).

ii) in case $d_\kappa = \tilde{l}_\kappa$, one sees immediately that the optimum remains unchanged since $\hat{d}_\kappa(\mu^*) = d_\kappa(\mu^*)$ and therefore $\hat{\mu} = \mu^*$.

This completes the proof. □

We conclude this section with a result on the effect of a changed variance on the optimal depreciation charges. It is often assumed that increased variance results in a lower depreciation charge in that specific period, since it implies that there is more uncertainty with respect to income. In the following corollary we show that this is not necessarily true.

Corollary 1. *If $F_\kappa(x) \sim N(\bar{\mu}, \sigma)$ and $d_\kappa^* > \bar{\mu}$ and $\hat{F}_\kappa(x) \sim N(\bar{\mu}, \hat{\sigma})$ with $\hat{\sigma} > \sigma$, then the following holds:*

$$\begin{aligned} &\hat{d}_k \leq d_k^* \quad \textit{for all } k \neq \kappa, \\ &\hat{d}_\kappa \geq d_\kappa^*. \end{aligned} \tag{62}$$

Moreover, a strict inequality holds for all k with $d_k^* > \tilde{l}_k$.

Proof. First notice that $\hat{\sigma} > \sigma$ implies that $\hat{F}_\kappa(x) \leq F_\kappa(x)$ for all $x \geq d^*_\kappa$. Now, the proof is similar to the proof of Proposition 5. □

When the depreciation charge in a period exceeds its expected income, increased variance is an incentive to increase the depreciation charge in that period. The intuition is as follows. For the optimal depreciation charges the probability that income will cover the depreciation charge is 'high enough'. If in the current optimal solution depreciation exceeds expected income, increased variance only increases this probability.

6 Numerical examples

In this section we illustrate our results in numerical examples. In Section 6.1 we show how the optimal depreciation charges can be determined from Theorem 2 and Proposition 3. Moreover, we illustrate that a lower discount factor yields a more accelerated depreciation scheme. In Section 6.2 we illustrate the effect of uncertainty by considering several scenarios for the distribution functions of the cash flows. Finally, in Section 6.3 we show how a change in the cash flow distribution in a certain period affects the optimal outcome.

In all examples, the initial amount to depreciate (D), as well as the maximum number of periods in which the asset can be depreciated (N) are equal to 5.

6.1 The effect of the discount rate

In order to focus on the effect of discounting, we first assume that $l_k = \tilde{l}_k = 0$ for all k.

Given that $l_k = \tilde{l}_k = 0$ for all k, the set of dynamic constraints is equal to the set of static constraints, and Theorem 2 and Proposition 3 imply that in order to find the optimal depreciation scheme, one should find the maximal $J \in \mathcal{P}$, which yields the optimal number of periods in which to depreciate D. The corresponding depreciation charges are given by (34). We now illustrate this procedure in a numerical example.

The future cash-flows have exponential distributions with $E[C_k] = 3$, for all $k = 1, \dots, 5$. The distribution function and inverse distribution function are:

$$\begin{aligned} &F(x) = 1 - e^{-x/3}, \qquad \text{for all } x \geq 0, \\ &F^{-1}(y) = -3ln(1-y), \text{ for all } y \in [0,1). \end{aligned} \tag{63}$$

In order to determine the optimal J, as defined in Theorem 1, we solve $\tilde{\Psi}_5(d) = 0$, i.e.

$$\begin{aligned} &5 - d - \textstyle\sum_{k=1}^4 F^{-1}\left(1 - \alpha^{5-k}(1-F(d))\right) = 0, \\ \Leftrightarrow\ &5 - d + 3\textstyle\sum_{k=1}^4 \ln\left(\alpha^{5-k}e^{-d/3}\right) = 0, \\ \Leftrightarrow\ &5 - d + 3\textstyle\sum_{k=1}^4 ((5-k)\ln(\alpha) - d/3) = 0, \\ \Leftrightarrow\ &d = 1 + 6\ln(\alpha). \end{aligned} \tag{64}$$

Consequently, for all α such that $1 + 6\ln(\alpha) > 0$, i.e. for all $\alpha \in (0.846, 1]$, one has $J = 5 \in \mathcal{P}$, and therefore the optimal depreciation scheme has $J = 5$. Theorem 1 then yields the corresponding depreciation charges:

$$d_5 = D_4 = 1 + 6\ln(\alpha), \tag{65}$$

and, for $k = 1, \ldots, 4$:

$$\begin{aligned} d_k &= F^{-1}\left(1 - \alpha^{J-k}(1 - F(D_4))\right), \\ &= -3(5-k)\ln(\alpha) + D_4. \end{aligned} \tag{66}$$

Straightforward calculations then yield:

$$\begin{cases} d_1 = 1 - 6\ln(\alpha), \\ d_2 = 1 - 3\ln(\alpha), \\ d_3 = 1, \\ d_4 = 1 + 3\ln(\alpha), \\ d_5 = 1 + 6\ln(\alpha). \end{cases} \tag{67}$$

For some values of α the results have been calculated and these are summarized in the following table.

Table 2. Effect of the discount rate

	$\alpha = 1$	$\alpha = 0.95$	$\alpha = 0.9$	$\alpha = 0.85$
d_1	1.000	1.308	1.632	1.975
d_2	1.000	1.154	1.316	1.488
d_3	1.000	1.000	1.000	1.000
d_4	1.000	0.846	0.684	0.512
d_5	1.000	0.692	0.368	0.025

(68)

We see that when α gets smaller, i.e. when the discounting effect gets stronger, the optimal method becomes more accelerated. Notice that, when $\alpha = 1$, the optimal method is the straight line depreciation, as stated in Theorem 3.

Now consider $\alpha \leq 0.846$. Then it follows from the above that the optimal number of periods in which to depreciate the total depreciation charge D is less than 5. Therefore, we solve $\tilde{\Psi}_4(d) = 0$.

$$\begin{aligned} & 5 - d - \textstyle\sum_{k=1}^{3} F^{-1}\left(1 - \alpha^{4-k}(1 - F(d))\right) = 0, \\ \Leftrightarrow\ & 5 - d + 3\textstyle\sum_{k=1}^{3}\left((4-k)\ln(\alpha) - d/3\right) = 0, \\ \Leftrightarrow\ & d = (5 + 18\ln(\alpha))/4. \end{aligned} \tag{69}$$

Consequently, $4 \in \mathcal{P}$ iff

$$d_4 = (5 + 18\ln(\alpha))/4 > 0. \tag{70}$$

So, the optimal depreciation scheme has $J = 4$ for all $\alpha \in (0.757, 0.846]$.

Straightforward calculations then yield:

$$\begin{cases} d_1 = (5 - 18\ln(\alpha))/4, \\ d_2 = (5 - 6\ln(\alpha))/4, \\ d_3 = (5 + 6\ln(\alpha))/4, \\ d_4 = (5 + 18\ln(\alpha))/4, \\ d_5 = 0. \end{cases} \tag{71}$$

As seen before, a lower value of α implies more accelerated depreciation, which in the above case implies that the optimal number of periods in which to depreciate the asset decreases.

6.2 *The effect of uncertainty*

We now illustrate the effect of uncertainty on the optimal depreciation scheme. To focus on the uncertainty effect, we again consider the situation where $l_k = \tilde{l}_k = 0$.

We consider three different scenarios for the cash flow distributions. All cash-flows have normal distributions $C_i \sim N(3, \sigma_i)$, with standard deviations as given in Table 3.

Table 3. Standard deviations for scenarios A, B and C

	Scenario A	Scenario B	Scenario C
σ_1	1	5	5
σ_2	2	4	4
σ_3	3	3	3
σ_4	4	2	3
σ_5	5	1	1

(72)

Whereas scenario A describes a situation where the uncertainty on realized payoffs increases over time, the opposite holds for scenario B. Scenario C is almost equal to scenario B, except for the higher variance in the fourth period. The results are stated in Table 4.

Table 4. Optimal solutions for scenarios A, B and C

	A		B		C	
	$\alpha = 0.8$	$\alpha = 0.90$	$\alpha = 0.8$	$\alpha = 0.90$	$\alpha = 0.8$	$\alpha = 0.90$
d_1	2.902	2.629	2.882	1.343	2.882	1.543
d_2	2.098	1.855	1.601	0.904	1.601	1.096
d_3	0	0.516	0.517	0.706	0.517	0.892
d_4	0	0	0	0.801	0	0
d_5	0	0	0	1.246	0	1.469

(73)

For scenario A, both the discounting effect and the increasing variances over time work in favor of a strongly accelerated method. Scenario's B and C with $\alpha = 0.9$ illustrate that, in contrast to the case where cash-flows are equally distributed (see Theorem 4), the optimal depreciation method is no longer accelerated. The explanation is as follows: The higher variances in the early periods imply that the risk of having a cash-flow that is lower than a given depreciation charge is higher in early periods than in later periods. Therefore there is a trade-off between the discounting effect, which always works in favor of accelerated depreciation, and the decreasing variances, which, since $d_k^* < E[C_k]$, work in favor of the opposite. We see that, whereas the discounting effect still had the upper-hand for $\alpha = 0.8$, this is no longer the case for $\alpha = 0.9$. Scenario C makes clear that increased variance in period 4 can imply that it is optimal not to plan any depreciation charge in that period. The lower bound ($\tilde{l}_4 = 0$) becomes binding.

6.3 *The effect of distribution functions and constraints*

In this section we illustrate that the effect of a change in the cash flow distribution on the optimal depreciation scheme depends crucially on the type of regulatory constraint.

We consider a risky project with high variance cash flow distributions and an expected loss in the second period. The distributions of the cash flows in periods $1, 2, 4$, and 5 are as follows:

$$C_1 \sim N(2,3), \quad C_2 \sim N(-1,3), \quad C_4 \sim N(3,3), \quad C_5 \sim N(3,3).$$

For the cash flow distribution in period 3, we consider the following three scenarios:

- Scenario **A**: $C_3 \sim N(3,3)$.
- Scenario **B**: $C_3 \sim N(3,4)$.
- Scenario **C**: $C_3 \sim N(2,4)$.

Let us first consider dynamic constraints where depreciation is restricted to be at least 40% of the residual tax base in years 2, 3 and 4. The discount rate α equals 0.95. The optimal depreciation charges and fractions for the three scenarios are given in Table 5.

Table 5. Optimal depreciation charges and fractions

Period i	A		B		C	
	d_i	γ_i	d_i	γ_i	d_i	γ_i
1	2.048	0.410	2.116	0.423	2.239	0.448
2	1.181	0.400	1.154	0.400	1.104	0.400
3	1.001	0.565	0.692	0.400	0.663	0.400
4	0.613	0.796	0.735	0.709	0.715	0.720
5	0.157	1.000	0.303	1.000	0.279	1.000

(74)

When comparing scenario **A** and scenario **B**, first notice that in scenario **A**, $d_3^* = 1.001 < 3 = E[C_3]$. Therefore, the higher variance in scenario **B** implies that $F_{3,B}(x) \geq F_{3,A}(x)$ for all $x \leq d_3^*$. Since $\gamma_3 > 0.4$ in scenario **A**, Proposition 4 i) or iii) applies. The numerical example indeed shows that the change in the cash flow distribution in period 3 yields a decrease not only in period 3 but also in period 2. The decrease in period 2 can be explained as follows. Since the expected result in this period is very low, the lower bound is binding. As opposed to the static case, the firm can affect the lower bound on the depreciation charge. More specifically, depreciating more in the first period results in a lower residual tax base, and as a result the lower bounds on d_2 and d_3 decrease. Furthermore, notice that, due to the change in period 3, the lower bound becomes binding there too.

Let us now compare scenario **B** and scenario **C**. The lower expected value in scenario **C**, implies that $F_{3,C}(x) \geq F_{3,B}(x)$ for all x. Now, since path 2 is optimal in scenario **B**, Proposition 4 ii) applies. We see that in this case the depreciation charges in periods 2, 3, 4 and 5 decrease. Again, the increase in the depreciation charge in period 1 allows to decrease the lower bounds that apply to periods 2 and 3.

Let us now consider the static case. The optimal solutions for scenarios **A, B** and **C**, with lower bounds equal to

$$\tilde{l}_1 = 0, \qquad \tilde{l}_2 = 0.7, \qquad \tilde{l}_3 = 0.7, \qquad \tilde{l}_4 = 0.7, \qquad \tilde{l}_5 = 0, \tag{75}$$

are presented in Table 6.

Table 6. Optimal depreciation charges in the static case

Period i	A d_i	B d_i	C d_i
1	1.002	1.097	1.196
2	0.700	0.700	0.700
3	1.440	1.073	0.700
4	1.114	1.241	1.370
5	0.744	0.889	1.034

(76)

The effect of the change in the distribution function on the optimal solution is as expected. More risk (scenario **A** $\rightarrow$ scenario **B**) or a lower expected cash flow (scenario **B** $\rightarrow$ scenario **C**) both imply a lower depreciation charge in period 3, and higher depreciation charges in all other periods. Scenarios with even lower expected values than in scenario **C**, result in the same optimal scheme as found in scenario **C** (see Proposition 5 ii)).

As opposed to the dynamic case, the effect of a change in period 3 does not depend on whether the constraint in period 2 is binding. The optimum is therefore more robust compared to the dynamic case.

7 Comparison static and dynamic constraints

In this section we summarize the differences between the static and dynamic constraints and the implications for the firm and the regulator.

First, there is an important difference in the number of periods in which the tax base can be depreciated. When the firm faces static constraints the regulator implicitly imposes a minimal number of periods in which the tax base is depreciated. When lower bounds are dynamic (so based on the residual tax base), the number of periods used can be fully determined by the firm. As a consequence, the depreciation life is not restricted and the optimal number of periods in which to depreciate follows from the optimization process.

Second, there is an important difference in the effect of a change in the distribution function in a certain period κ. In the static case, a higher probability on cash flows that are not sufficient to cover the depreciation charge in a certain period implies that in the optimal solution the depreciation charge allocated to that period decreases. As a result the depreciation charges in the other periods increase. In the dynamic case however, this is not necessarily true. The depreciation charges in all other periods can then either decrease or increase, depending on the path that is optimal in the initial optimal solution and on the size of the change in the distribution function. The intuition for this difference is as follows. The fact that the lower bounds depend on the earlier depreciation decisions allows the firm to decrease those lower bounds by increasing the earlier depreciation charges. This can imply that depreciation charges in periods other than κ decrease due to the lower residual tax base. If the lower bound is binding in period κ, then, contrary to the case with static constraints, the change still has an effect on the optimal solution.

In conclusion, the optimal solution when facing static constraints is more robust with respect to estimation errors in the cash flow distributions, but leaves less room for strategic tax planning.

8 Conclusion and future research

This paper determines the optimal depreciation scheme given that the objective is to minimize expected discounted future tax payments. Whereas previous research focused on comparing different methods, we determine the *optimal* depreciation scheme given constraints imposed by the tax authority. We consider constraints on the fraction of the initial depreciable value, as well as constraints on the fraction of the remaining depreciable value. This optimization process also yields the optimal depreciation life (the optimal number of periods in which to depreciate the asset). The effects of the discount rate, the cash-flow distributions and the constraints are analyzed. Our results make clear that the degree of uncertainty (e.g. the variance) in future cash-flows largely affects the optimal choice. Decisions based solely on the expected value of future cash-flows can therefore be critically off-mark. With respect to the two types of constraints, we find that the dynamic constraints allow for more strategic behaviour of the firm, but also result in less robust optimal solutions with respect to changes in the cash flow distributions. Finally, we found the surprising result that the effect of an increase in the variance of a cash flow depends on the expected cash flow and on the optimal depreciation charge planned in that period. Contrary to what is often assumed, an increase in the variance may lead to a higher optimal depreciation charge in that period. This is the case when the expected cash flow is less than the planned depreciation charge.

For future research it might be interesting to move to a game-theoretic approach where the tax authority has to set the constraints. Interesting points there are that, due to welfare considerations, the objective of the government is more complex than maximization of tax revenues, and that the information on the cash-flow distributions will be asymmetric. This can possibly be a starting point for the discussion to increase or decrease the freedom of firms in choosing the tax depreciation method.

A Proof of Theorem 1

Define

$$\lambda_k^* := \alpha \min_{\gamma \in [l_{k+1},1]} \left\{\gamma(1 - F_{k+1}(0)) + (1-\gamma)\lambda_{k+1}^*\right\}, \quad k = 1, \ldots, N-1,$$
$$\lambda_N^* := 0,$$
$$\gamma_k^* := \text{argmin}_{\gamma \in [l_k,1]} \left\{\gamma(1 - F_k(0)) + (1-\gamma)\lambda_k^*\right\}, \quad k = 1, \ldots, N.$$

$\Rightarrow$) First suppose that $\hat{d}$ satisfies (10)-(16). We will now show that it satisfies (27) and $d_J \leq F_J^{-1}(1 - \lambda_J^*)$.

First observe that (10)-(16) imply that $\gamma_k < 1$ for all $k = 1, \ldots, J-1$ and $\gamma_J = 1$, or, equivalently, $D_J = 0$ and $D_{J-1} > 0$.

Indeed, if $D_{N-1} > 0$, then since $\lambda_N = 0$, (11), (14) and (15), imply that

$$\begin{aligned} \eta_N^2 &= (1 - F_N(D_{N-1}))D_{N-1} > 0, \\ \gamma_N &= 1, \\ \eta_N^1 &= 0. \end{aligned} \tag{77}$$

It therefore follows that $D_N = 0$, so that $J = N$ in this case. If $D_{N-1} = 0$, obviously the fact that $D_0 = D > 0$, and $D_k \leq D_{k-1}$ for all k, implies that there exists a unique $k < N$ such that $D_k = 0$ and $D_{k-1} > 0$. Since $d_J > 0$, and $d_k = 0$ for all $k \geq J+1$, it follows that $k = J$.

It therefore follows that Path 1 or Path 2 is applied in periods $k = 1, \ldots, J-1$. Consequently, (11) and (12) imply that:

$$\lambda_k := \min\bigl\{\alpha\lambda_{k+1}, \alpha(1 - F_{k+1}(l_{k+1}D_k))l_{k+1} + \alpha(1 - l_{k+1})\lambda_{k+1}\bigr\}, \tag{78}$$

for $k = 0, \ldots, J-2$.

Moreover, $\gamma_J = 1$, implies that in period J either path 3 is followed, or path 1 with $\gamma_J = \tilde{\gamma}_J = 1$. The dynamics in both cases imply that:

$$\lambda_{J-1} = \alpha(1 - F_J(D_{J-1})). \tag{79}$$

Now take an arbitrary $k \leq J-1$. Then $D_{k-1} > 0$ and $\lambda_k > 0$ imply that $\tilde{\gamma}_k$, as defined in lemma 1, exists.

We can now apply lemma 1, which yields that:

- Path 2 is feasible iff $\tilde{\gamma}_k < l_k$, and then $\gamma_k = l_k$.
- Otherwise, path 1 is feasible, and then $\gamma_k = \tilde{\gamma}_k$.

It therefore follows from (11) that:

$$\begin{aligned} d_k &= \gamma_k D_{k-1} = D_{k-1} - D_k, \\ &= \max\big\{ l_k D_{k-1} \ , \ F_k^{-1}\big(1-\lambda_k\big)\big\}, \end{aligned} \tag{80}$$

for all $k \leq J-1$. This implies that

$$\begin{aligned} &D_{k-1} = \max\big\{ \frac{D_k}{1-l_k} \ , \ D_k + F_k^{-1}\left(1-\lambda_k\right)\big\}, \\ &D_{J-1} = d_J, \end{aligned}$$

so that d_k satisfies (27) for all $k \leq J-1$.

Moreover, notice that it follows from (25) that the minimal value of λ_J that can be reached with $\gamma_k \in [l_k, 1]$ for $k \geq J+1$ equals λ_J^*. Therefore, $\eta_J^2 \geq 0$ implies that $1 - F_J(d_J) \geq \lambda_J^*$.

It therefore remains to show that $\Psi_J(d_J) = 0$. This follows immediately from (10), i.e. $D_0(d_J, J) = D_0 = D$.

$\Leftarrow$) Suppose that $(d_1, \ldots, d_J, 0 \ldots, 0)$ satisfies (27) and $d_J \leq F_J^{-1}(1-\lambda_J^*)$. We will show that there exist variables $\gamma_k, \lambda_k, \eta_k^1, \eta_k^2$, and D_k, for $k = 1, \ldots, N$, that satisfy (10)-(16), and lead to depreciation charges as in (27).

Therefore, we define the following variables:

$$\begin{aligned} &D_{k-1} = \max\big\{ \tfrac{D_k}{1-l_k} \ , \ D_k + F_k^{-1}\left(1-\lambda_k\right)\big\}, && k \leq J-2, \\ &D_{J-1} = d_J, \\ &D_k = 0, && k \geq J, \end{aligned} \tag{81}$$

and

$$\begin{aligned} &\lambda_k = \min\{\alpha\lambda_{k+1}, \alpha(1 - F_{k+1}(l_{k+1}D_k))l_{k+1} + \alpha(1-l_{k+1})\lambda_{k+1}\} \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 \leq k \leq J-2, \\ &\lambda_{J-1} = \alpha(1 - F_J(D_{J-1})), \\ &\lambda_k = \lambda_k^*, && J \leq k < N, \\ &\lambda_N = 0. \end{aligned} \tag{82}$$

$$\begin{aligned} &\gamma_k = (D_{k-1} - D_k)/D_{k-1}, \ k \leq J-1, \\ &\gamma_J = 1, \\ &\gamma_k = \gamma_k^*, && k \geq J+1, \end{aligned} \tag{83}$$

$$
\begin{aligned}
&\eta_k^1 = \big(\lambda_k - (1 - F_k(\gamma_k D_{k-1}))\big) D_{k-1}, && k \le J-1,\\
&\eta_k^2 = 0, && k \le J-1,\\
\\
&\eta_J^1 = 0, && \\
&\eta_J^2 = (1 - F_J(D_{J-1}) - \lambda_J^*)\, D_{J-1}, && \\
\\
&\eta_k^1 = \eta_k^2 = 0, && k \ge J+1.
\end{aligned}
\tag{84}
$$

By definition, one has $D_{J-1} > 0$, and consequently, by construction, $D_k > 0$, and $D_{k+1} \le D_k/(1-l_k)$ for $k = 0, \ldots, J-1$. This implies that $\gamma_k \in [l_k, 1]$, and:

$$
\gamma_k = \max\Big\{l_k\ ,\ \frac{1}{D_{k-1}} F_k^{-1}\,(1-\lambda_k)\Big\}, \qquad k = 1, \ldots, J-1. \tag{85}
$$

Now notice that $\gamma_k = l_k$ implies that:

$$
\begin{aligned}
&\tfrac{1}{D_{k-1}} F_k^{-1}\,(1-\lambda_k) \le l_k,\\
&\Rightarrow 1 - F_k(l_k D_{k-1}) \le \lambda_k,\\
&\Rightarrow \lambda_k - (1 - F_k(l_k D_{k-1})) \ge 0,
\end{aligned}
\tag{86}
$$

and $\gamma_k > l_k$ implies that:

$$
\begin{aligned}
&\gamma_k = \tfrac{1}{D_{k-1}} F_k^{-1}\,(1-\lambda_k)\,,\\
&\Rightarrow (1 - F_k(\gamma_k(D_{k-1})) = \lambda_k,\\
&\Rightarrow \lambda_k - (1 - F_k(\gamma_k D_{k-1})) = 0.
\end{aligned}
\tag{87}
$$

This implies that $\eta_k^1 \ge 0$ for all $k \le J-1$, and, by definition, $\eta_k^1 = 0$ for $J \le k \le N$. Obviously, also $\eta_k^2 \ge 0$ for all $k \le N$.

Furthermore, one can check that for all $k \le N$,

$$
\begin{cases}
(1 - F_k(\gamma_k D_{k-1}))D_{k-1} - \lambda_k D_{k-1} + \eta_k^1 - \eta_k^2 = 0,\\
\lambda_{k-1} = \alpha T(1 - F_k(\gamma_k D_{k-1}))\gamma_k + \alpha\lambda_k(1-\gamma_k),\\
\eta_k^1(\gamma_k - l_k) = 0,\\
\eta_k^2(1-\gamma_k) = 0,
\end{cases}
\tag{88}
$$

It therefore only remains to show that $D_0 = D$. This follows immediately from $\Psi_J(D_{J-1}) = 0$. This completes the proof. □

B Proof of Proposition 2

Let us denote $\hat{\gamma}_1, \ldots, \hat{\gamma}_J$ for the fractions that yield the optimal depreciation charges in periods $1, \ldots, J$. Then, since $D_J = 0$, the vector $(\hat{\gamma}_1, \ldots, \hat{\gamma}_J, 1, \ldots, 1)$ must satisfy the necessary conditions for optimality. Now $\gamma_{J+1} = 1$ implies that $\lambda_J = \alpha(1 - F_{J+1}(0))$ so that η_J^2 is non-negative iff (29) is satisfied. □

References

Berg M, De Waegenaere A, Wielhouwer JL (2001) Optimal tax depreciation with uncertain future cash-flows. European Journal of Operational Research 132(1): 185–197

Berg M, Moore G (1989) The choice of depreciation methods under uncertainty. Decision Sciences 20: 643–654

Feichtinger G, Hartl RF (1986) Optimale Kontrolle ökonomischer Prozesse. de Gruyter, Amsterdam

van Hilten O, Kort PM, van Loon PJJM (1993) Dynamic policies of the firm: an optimal control approach. Springer, Berlin Heidelberg New York

Scholes MS, Wolfson, MA (1992) Taxes and business strategy: a planning approach. Prentice-Hall, New York

Wakeman LM (1980) Optimal tax depreciation. Journal of Accounting and Economics 1: 213–237

Wielhouwer JL, De Waegenaere A, Kort PM (2000) Optimal tax depreciation under a progressive tax system. CentER Discussion Paper 2000-51. Journal of Economic Dynamics and Control (forthcoming)