"Dynamic risk management: investment, capital structure, and hedging in the presence of financial frictions"

Diego Amaya, Geneviève Gauthier
and Thomas-Olivier Léautier

Dynamic risk management: investment, capital structure, and hedging in the presence of financial frictions

Diego Amaya[*]   Geneviève Gauthier[†]   Thomas-Olivier Léautier[‡][§]

April 12, 2012

Abstract

This paper develops a dynamic risk management model to determine a firm's optimal risk management strategy. The risk management strategy has two elements: first, until leverage is very high, the firm fully hedges its operating cash flow exposure, due to the convexity in its cost of capital. When leverage exceeds a very high threshold, the firm gambles for resurrection and stops hedging. Second, the firm manages its capital structure through dividend distributions and investment. When leverage is very low, the firm fully replaces depreciated assets, fully invests in opportunities if they arise, and distribute dividends to reach its optimal capital structure. As leverage increases, the firm stops paying dividends, while fully investing. After a certain leverage, the firm also reduces investment, until it stop investing completely. The model predictions are consistent with empirical observations.

JEL classification: G32, C61

Keywords: Dynamic programming; risk management; capital structure; hedging.

1 Introduction

Modern risk management theory is grounded in the observation that information asymmetry between managers/insiders and investors/outsiders limits the ability of firms to raise external funds. In a striking reversal of Modigliani and Miller propositions (1958, 1963), not all value-creating firms or projects are financed (Holmstrom and Tirole (2000), Tirole (2006)): profitable – but cash constrained – firms may not be able to refinance themselves after a negative shock to their cash flow, hence may go bankrupt; firms with insufficient internal funds may have to forego profitable investment opportunities, an issue known as the underinvestment problem.

The interaction between costly external financing, underinvestment, and risk management was first modeled in a two-period environment by the two seminal articles Froot *et al.* (1993) and Froot and Stein (1998). The former considers a firm facing random cash flows, random investment opportunities, and convex cost of external financing. At the optimum, the firm fully hedges if cash flows and investment opportunity are uncorrelated, and reduces its hedging as the correlation between both sources of uncertainty increases. The latter introduces capital structure as a risk management device. A marginal increase in equity raises the firm's capacity to pursue risky investments. On the other hand, it generates deadweight costs, arising for example from the tax deductibility of interest payments. The optimal equity level balances these two effects.

A more recent literature has examined this issue in multi-period models. Rochet and Villeneuve (2011) develop an infinite-horizon, continuous-time model, where a constant-size firm faces exogenous cash flow shocks and stringent financial frictions: the firm is liquidated as soon as its cash reserve becomes negative. At each instant, the firm selects its dividend payment and decides its hedging ratio or insurance coverage for discrete risks. Rochet and Villeneuve (2011) restate the risk management problem as an inventory management problem, where the cash reserve is the state variable, and dividend payment and risk transfer decisions are the control variables. They then show that the firm pays dividends if and only if the cash reserve exceeds a threshold, and it fully hedges if the cash reserve is below the threshold. In addition, they show that the firm insures small risks but not large ones.

Bolton *et al.* (2011) extend Rochet and Villeneuve (2011), most notably by including (*i*) investment and growth, and (*ii*) less stringent financial frictions, i.e., refinancing is possible, albeit costly. They first characterize the optimal dividend distribution, investment, and refinancing policies. As Rochet and Villeneuve (2011), Bolton *et al.* (2011) find that the firm optimally distributes dividends if and only if its cash reserve (as a percentage of its size) exceeds a given threshold. The optimal investment policy departs from the Modigliani and Miller rule, which is to equalize the marginal cost of adjusting physical capital to the marginal value of capital (the marginal Tobin's q). Bolton *et al.* (2011) show that the marginal cost of physical capital is set equal to the ratio of the marginal Tobin's q over the marginal cost of financing. Finally, they determine the optimal hedging policy, that balances the marginal benefits and costs of hedging.

While these two last articles are significant conceptual contributions, they do not capture essential real-world features. First, corporate taxes and leverage are absent from these analyses, even though they play an important role in corporate decision making, in particular hedging and capital structure decisions, as found for example in Graham and Rogers (2002)'s econometric analysis of the determinants of hedging. Second, as observed for example in Graham and Harvey (2001)'s survey of financial executives, managers at large, publicly traded firms use the Net Present Value (NPV) of the free cash flows, discounted at the Weighted Average Cost of Capital (WACC), to make capital budgeting decisions, while managers in these analyses maximize the NPV of dividends. Finally, at least a fraction of investment opportunities appears to be stochastic, as was modeled by Froot *et al.* (1993). Firm's growth is shaped by the availability of investment opportunities as well as by real frictions in adding capital.

We are not aware of any model derived from micro-foundations that incorporates these essential features. Furthermore, micro-founded models rely on information asymmetry, whose parameters are by nature difficult to estimate, hence these models' predictions are difficult to test empirically (Graham and Harvey (2001) and Graham and Rogers (2002)).

Therefore, to approximate reality, Léautier *et al.* (2007) have developed a reduced-form model, that aims to represent what firms actually do: managers make capital budgeting and hedging decisions to maximize the NPV of the free cash flows (consistent with Graham and Harvey (2001)'s finding) in a multi-period environment (as in Rochet and Villeneuve (2011) and Bolton *et al.* (2011)), facing uncertainty about both future cash flows and future investment opportunities (as in Froot *et al.* (1993)). Financial frictions are incorporated through the expected return required by investors, which is assumed to be convex in the firm's leverage ratio, and replaces the cash reserve as the state variable.

This article builds upon Léautier et al. (2007), and incorporates two additional features: dividends distribution and the possibility of bankruptcy. This representation of managerial decision-making cannot be derived from micro-foundations. Yet, it provides valuable insights for two reasons. First, as previously mentioned, it describes the behavior of large publicly traded corporations, that make capital budgeting and hedging decisions using the NPV of the free cash flows, recognizing that the discount rate is a U-shaped function of leverage ratio (Graham and Harvey (2001), Pettit (2007), pp. 110-111 and 141-159, and Cohen (2003)). Thus the analysis' implications can be tested on data.

Second, we believe it provides an adequate representation of the value of risk management in the presence of two financial frictions: (i) tax deductibility of interest payments, and (ii) information asymmetry between managers/insiders and investors/outsiders. Most of our results hinge on the average Tobin's q, i.e., the market-to-book value of the firm, being a concave function of leverage with a unique maximum on $[0, b]$, where $b < 1$. This effect is produced in our model by the convexity of the expected return required by investors. We argue later in this article that this shape of the average Tobin's q is consistent with

3

micro-founded models such as Biais *et al.* (2007), and Biais *et al.* (2011), that derive the dynamically optimal financial contracts resolving the underlying repeated moral-hazard problem between managers/insiders and investors/outsiders. In other words, we believe our model constitutes a robust approximation of what firms should do.

Our contribution is therefore to solve analytically for what firms do, which constitutes an approximation of what they should do. We first determine analytically the optimal risk management strategy, i.e., the mix of hedging, dividend distribution, refinancing, and investment policies. Second, we illustrate the optimal strategy for a "representative" industrial firm, using estimates of the main parameters. Finally, we show that the model's predictions are consistent with empirical observations.

The main results are summarized below.

The optimal risk management strategy is surprisingly simple. First, dividend distribution and investment jointly follow four regimes (Proposition 1). For low leverage, the firm enjoys *full financial flexibility*: it fully finances its investment needs, and distributes dividends to reach its optimal leverage ratio, which maximizes its expected continuing value. For intermediate leverage, the firm faces *financial tightness*: its still fully finances its investment needs, but no longer distributes dividends, while leverage increases from one period to the next. For higher leverage, the firm faces a *financial constraint*: it is no longer able to fully finance its investment needs. The portion it finances is determined to reach a target leverage, after which the marginal value of investing becomes negative. Finally, for high leverage, the firm faces *financial hardship*: it is no longer able to finance any of its investment needs, not even depreciation.

Second, full hedging is optimal unless leverage gets higher than some threshold, in which case gambling for resurrection becomes optimal (Proposition 2). This results differs from Rochet and Villeneuve (2011) and Bolton *et al.* (2011), who find that, when the firm's cash reserve (or cash-to-capital ratio in Bolton *et al.* (2011)) is high enough, the firm becomes risk neutral, and, since hedging is costly, stops hedging. In our model, the tax shield drives the concavity of the value function, hence the optimality of full hedging. By choosing leverage as the state variable we are able to capture the tax shield from debt, a real effect, absent from Rochet and Villeneuve (2011) and Bolton *et al.* (2011).

An essential finding of the analysis is that, if the firm's expected profitability is lower than a threshold, function of its investment opportunity, the firm does not exhaust the benefits of the tax shield, rather it keeps an equity cushion (Proposition 3): the optimal leverage target is lower than the static optimum, that minimizes the cost of capital.

The optimal risk management strategy is robust to changes in volatility of operating cash flows and to the imposition of a constraint on the firm's ability to hedge.

This analysis yields a series of testable implications: First, Proposition 1's predictions that, *ceteris paribus*, dividend distribution decreases when firms are less profitable on average or face face higher investment opportunities, are confirmed by Fama and French (2002).

Second, the importance of corporate taxes in the decision to hedge (Proposition 2) is confirmed empirically by Graham and Rogers (2002)). Third, Graham (2000) reports that even profitable firms with low expected cost of financial distress hold an equity cushion, consistent with our model (Proposition 3).

This article is structured as follows. Section 2 presents the model. Section 3 presents the estimation of the parameters used to illustrate the analysis. Section 4 presents the optimal risk management strategy. Section 5 examines the robustness of the optimal risk management strategy to various changes. The concluding Section 6 discusses future research directions. Throughout the article, all proofs are presented in the appendix, while only the results' intuition is presented in the main text.

2   The model

2.1   Timing, decisions, and Free Cash Flow

For $t \geq 0$, period $(t + 1)$ runs from dates t to $(t + 1)$. At date t, the firm's invested capital is I_t, that depreciates at constant rate δ during each period. Depreciation during period $(t+1)$ is therefore δI_t.

At date t, an investment opportunity arises. The magnitude of the opportunity i_t is expressed as a fraction of the invested capital I_t. Therefore, the nominal opportunity is $i_t I_t$. Senior managers do not know in advance when or whether new investment opportunities will arise or materialize, as this occurrence depends on a variety of factors, e.g., regulatory approval, operational limitations. Therefore, investment opportunities $\{i_t : t \in \mathbb{N}\}$ are random variables, assumed to be independent and identically distributed.

At date t, the firm makes four decisions. First, it selects its dividend payout $d_t I_t$ for the next period, i.e., the firm commits at the beginning of a period to a minimum dividend. As will be discussed later, dividend rates are assumed to be non negative, that is,

$$d_t \geq 0. \tag{1}$$

Second, the firm chooses its investment $g_t I_t$ for the next period. The firm has the possibility to replace partially or completely the capital δI_t depreciated throughout the period and to invest partially or completely in the opportunity $i_t I_t$, thus

$$0 \leq g_t \leq i_t + \delta. \tag{2}$$

Taking depreciation and investment into account, invested capital at date $(t + 1)$ is

$$I_{t+1} = (1 + g_t - \delta) I_t.$$

Third, the firm sets its hedging ratio to the underlying source of risk η_t, as described in Section 2.2. Finally, the firm may distribute extraordinary dividends $\tilde{d}_t I_{t-1}$ under conditions described in Section 2.3.

5

Throughout period $(t+1)$, the invested capital I_t generates stochastic Return On Invested Capital (ROIC)[1] x_{t+1}, expressed in percent. The Net Operating Profit after Adjusted Taxes (NOPAT) realized during period t is $\pi_{t+1} = x_{t+1}I_t$. The Free Cash Flow (FCF) is the NOPAT minus the net increase in invested capital $I_{t+1} - I_t = (g_t - \delta) I_t$:

$$FCF_{t+1} = (x_{t+1} - g_t + \delta)I_t.$$

2.2   Hedging technology

The random variable z_{t+1} represents the primitive source of uncertainty that affects the ROIC during period $(t+1)$. For example, for an oil company, z_{t+1} is the wholesale crude oil price (multiplied by its per-period production, less operating costs, taxes and depreciation, divided by invested capital). Returns on invested capital $\{z_t : t \in \mathbb{N}\}$ are assumed to be serially independent and identically distributed through time. This assumption implies that the firm does not diversify its investments.

The firm can costlessly hedge its entire exposure to risk z_{t+1}. This assumption may appear unreasonable. We acknowledge this limitation, and consider this simplifying assumption as a first step towards a more general model. We argue in Section 5 that relaxing this assumption does not significantly modify the structure of the risk management strategy. Finally, we assume that the forward price is equal to the expected spot price. The ROIC x_{t+1} is then:

$$x_{t+1} = \eta_t \mathrm{E}[z] + (1 - \eta_t) z_{t+1}.$$

$\mathrm{E}[x_{t+1}] = \mathrm{E}[z]$ for all values of η_t: hedging does not affect expected profitability. However, as expected, hedging modifies the profitability's volatility:

$$\mathrm{Var}(x_{t+1}) = (1 - \eta_t)^2 \mathrm{Var}(z).$$

We assume that Board of directors, concerned that derivatives are used to speculate, prevent the firm from using derivatives to either (i) increase its exposure, or (ii) become short its underlying exposure. For example, an oil company that buys oil forward, i.e., sets $\eta_t < 0$, increases its exposure, as $(1 - \eta_t) > 1$. On the other hand, an oil company that sells more oil forward than its production, i.e., sets $\eta_t > 1$, becomes short the underlying risk as $(1 - \eta_t) < 0$. Hence, we assume:

$$0 \leq \eta_t \leq 1. \tag{3}$$

2.3   Leverage dynamics

The firm's capital structure at date t is a combination of debt D_t and equity E_t. The firm's leverage ratio is $\lambda_t = \frac{D_t}{I_t}$.

---

[1]See for example Copeland *et al.* (1995)

The firm is assumed to not issue shares, which is consistent with theoretical models as well as empirical evidence. The "pecking order" theory of financing (Myers and Majluf (1984)) suggests that firms finance their growth first by using retained earnings, then by issuing debt, and only when the two other sources are exhausted, by issuing equity. The economic reasoning is that, by issuing shares, managers and existing shareholders, who have superior information on the true value of the firm, signal to outside investors that they believe that the firm is worth less than the offered share price. As a result, investors heavily discount seasoned equity offering. This effect is confirmed by the data. Seasoned equity offerings usually yield a permanent fall in the stock price of about 3% (see Tirole (2006), page 101, and the references therein). As a result, firms are reluctant to issue shares. Rajan and Zingales (2003) report that the fraction of gross fixed-capital formation raised via equity (including initial and seasoned equity offerings) in 1999 was only 12% in the United States, 9% in the United Kingdom and France, 8% in Japan, and 6% in Germany.

Denoting the after-tax cost of debt by $r(\lambda_t)$, the financing flow for period $(t+1)$ is the after-tax interest payments[2] $r(\lambda_t) D_t$, minus changes in financial structure $D_{t+1} - D_t$, plus anticipated and extraordinary dividends paid during the period $\left(d_t + \tilde{d}_{t+1}\right) I_t$:

$$FF_{t+1} = r(\lambda_t) D_t - (D_{t+1} - D_t) + \left(d_t + \tilde{d}_{t+1}\right) I_t.$$

Free cash flow equals financing flow for the period, hence:

$$(x_{t+1} - g_t + \delta)I_t = r(\lambda_t) D_t - (D_{t+1} - D_t) + \left(d_t + \tilde{d}_{t+1}\right) I_t.$$

Dividing by I_t and observing that $I_{t+1} = (1 + g_t - \delta)I_t$ yields:

$$x_{t+1} - g_t + \delta = \mu(\lambda_t) - \lambda_{t+1}(1 + g_t - \delta) + \left(d_t + \tilde{d}_{t+1}\right)$$

where $\mu(\lambda_t) = \lambda_t(1 + r(\lambda_t))$ is the debt and its interest per unit of invested capital. Isolating λ_{t+1} in the previous equation leads to the following definition:

$$\Lambda_{t+1} = \frac{\mu(\lambda_t) + g_t - \delta - x_{t+1} + d_t + \tilde{d}_{t+1}}{1 + g_t - \delta} = y_{t+1} - \sigma \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon_{t+1} + \frac{\tilde{d}_{t+1}}{1 + g_t - \delta}$$

where

$$y_{t+1} = \frac{\mu(\lambda_t) + g_t - \delta + d_t - \mathrm{E}[z]}{1 + g_t - \delta} \tag{4}$$

is the deterministic next period leverage, and $\varepsilon_{t+1} = \frac{z_{t+1} - \mathrm{E}[z]}{\sigma}$ is a random variable of mean zero and standard deviation unity.

---

[2]In fact, since the investment is spread throughout the period, after-tax interest payments are slightly higher than $r(\lambda_t) \cdot D_t$. For example, firms sometimes use the average debt: $r(\lambda_t) \cdot \frac{D_{t+1} + D_t}{2}$. This extension is left for future work.

Extraordinary dividends $\tilde{d}_{t+1}$ are distributed if and only if there still is excess cash flow after interest and all outstanding debt has been paid, that is, extraordinary dividends are

$$\tilde{d}_{t+1} = \max\left(\sigma\left(1 - \eta_t\right)\varepsilon_{t+1} - \left(1 + g_t - \delta\right)y_{t+1}, 0\right)$$

and $\Lambda_{t+1} = \max\left(y_{t+1} - \sigma\frac{1-\eta_t}{1+g_t-\delta}\varepsilon_{t+1}, 0\right)$. Because the firm is considered bankrupt whenever $\Lambda_{t+1} \geq 1$, then λ_{t+1} is set to 1 and

$$\lambda_{t+1} = \min\left(\max\left(y_{t+1} - \sigma\frac{1-\eta_t}{1+g_t-\delta}\varepsilon_{t+1}, 0\right), 1\right). \tag{5}$$

Equations (4) and (5) shows that λ_{t+1} increases with λ_t and d_t, and decreases with ε_{t+1}. The effect of a higher investment level g_t is mixed: it increases (decreases) the leverage when $(1 - \lambda_t + x_{t+1} + \delta)$ is larger (smaller) than the dividends plus interest payments per unit of invested capital $\left(d_t + \tilde{d}_{t+1} + \lambda_t r\left(\lambda_t\right)\right)$.

2.4   Managers' objective function

This article assumes that managers of large publicly traded firms maximize V_t, the total value of the firm at date t, i.e., the Net Present Value of the free cash flows.

   This assumption differs from Rochet and Villeneuve (2011) and Bolton *et al.* (2011), who assume managers maximize the value to shareholders, i.e., the Net Present Value of dividends, and may appear contrary to managers' fiduciary duties. However, it is consistent with managerial practices: Graham and Harvey (2001) report that 75% of the CFOs they surveyed use always or almost always the Net Present Value (NPV) as the primary capital budgeting method, and that this effect is stronger for larger firms. This objective function is also consistent with *ex ante* optimization of the value of the firm, as discussed by Welch (2008, pages 452-454): if a management team announces *ex ante* it will maximize the value of shareholders at the expense of other stakeholders, another team that commits to maximizing the value for all stakeholders can raise more capital to purchase the assets, hence replace the former.

   Firms use a two-step approach for valuation: cash flows are explicitly computed for a first period (usually five years), and a continuing value is estimated for cash flows arising at the end of that period. Denote $(T + 1)$ the date where the continuation value is determined. For $t \leq T$, V_t, the value of the firm at date t, is

$$V_t = \mathrm{E}_t\left[\sum_{s=t+1}^{T+1}\left(\prod_{k=t}^{s-1}D\left(\lambda_k\right)\right)FCF_s + \left(\prod_{k=t}^{T}D\left(\lambda_k\right)\right)V_{T+1}\right], \tag{6}$$

where V_{T+1} is the continuing value, $D\left(\lambda_t\right) = \left(1 + w(\lambda_t)\right)^{-1}$ is the discount factor for the period $]t, t + 1]$, assumed to be a function of the leverage ratio λ_t, and the conditional

expectation E_t is taken with respect to the information[3] available at date t. Financial frictions are incorporated in the value of the firm through the cost of capital $w(\lambda_t)$. Expression (6) is equivalent to the recursive definition:

$$V_t = D(\lambda_t)(E_t[FCF_{t+1} + V_{t+1}]).\qquad(7)$$

2.4.1 Discount rate

The discount rate $w(\lambda_t)$ is the expected rate of return required by shareholders. Absent financial frictions, $w(\lambda)$ is constant: the required rate of return (or equivalently, the value of the firm) is independent of the leverage ratio. This is the Modigliani and Miller (1958) irrelevance proposition.

Financial frictions alter this result. First, the tax deductibility of interest payments generates a tax shield, that increases as leverage increases. For example, if leverage is constant and the debt is risk-free, this friction can be represented by a linearly decreasing $w(\lambda)$ (see for example Welch (2008), and Copeland *et al.* (2005)).

If $w(\lambda)$ was decreasing for all values of λ, firms would be all debt financed, which is of course inconsistent with reality. Information asymmetries between insiders/managers and outsiders/investors imply that, as leverage increase, the latter require a higher expected return than under the Modigliani and Miller hypotheses. Tirole (2006), chapter 3, derives this result using a simple and general moral-hazard model where investors/outsiders must leave rents to a manager/insider to elicit effort. External financing available to the manager cannot exceed a multiple d of her initial net worth A, i.e., the cost of expected return required by investors is constant until the leverage reaches $\frac{d}{d+1}$ (taxes are ignored), then becomes infinite afterwards. Welch (2008) provides a clear discussion of these frictions and illustrates their impact on the discount rate $w(\lambda)$. Indeed, its Figure 18.1 shows how including the tax deductibility of interest payments leads to a decreasing $w(\lambda)$ while its Figure 19.1 illustrates how including information asymmetries leads to convex $w(\lambda)$.

Copeland *et al.* (2005) confirm empirically the convexity of $w(\lambda)$. They estimate $w(\lambda)$ for a set of companies using the classical WACC formula:

$$w(\lambda) = (1 - \lambda)k^e(\lambda) + \lambda(1 - \tau)k^d(\lambda)$$

where $k^e(\lambda)$ is the expected return on equity, and $k^d(\lambda)$ is the expected yield on debt. First, they start from the yield curve for bonds, that is, the relationship between the yield to maturity of 10-year bonds and their ratings. They notice a strong yield increase between BBB3, the last investment grade rating, and BB1, the first speculative grade rating. Second, they estimate the statistical relationship between debt rating (dependent variable) and leverage (independent variable). This provides them with an empirical estimate of $k^d(\lambda)$. Third, they compute the cost of equity $k^e(\lambda)$ using the Capital Asset Pricing Model. Combining these

---

[3]The filtration $\{\mathcal{F}_t : t = 0, ..., T\}$ is generated by the random terms, that is, $\mathcal{F}_t = \sigma\{i_u, z_u : u = 1, 2, ..., t\}$. The conditional expectation with respect to $\mathcal{F}_t$ is denoted $E_t[\cdot]$.

expressions, they compute $w(\lambda)$ for each company in their sample. Their Figure 15.16 shows that $w(\lambda)$ is convex, and reaches a minimum around the leverage corresponding to a BBB rating, which contradicts empirically the Modigliani and Miller hypotheses.

Firms use a method similar to Copeland *et al.* (2005) to estimate their cost of capital, and recognize that the discount rate is a U - shaped function of leverage ratio (Pettit (2007), pp. 110-111 and 141-159, and Cohen (2003)).

Following the above discussion, we assume that:

Assumption 1 (*Shape of w and r*)
1. $w(\cdot)$ is a continuous, differentiable and convex function on $[0,1]$ reaching a minimum at λ^*.
2.   The discount factor $D(\cdot) = (1 + w(\lambda))^{-1}$ is a convave function of λ on $[0,1]$.
3. $\lim_{\lambda \to 1} w(\lambda) = \infty$.
4. $\lim_{\lambda \to 1} \frac{w'(\lambda)}{(w(\lambda))^2} = 0$.
5.   The interest rate $r(\cdot)$ *is a continuous, increasing, convex and twice differentiable function of λ with $\lim_{\lambda \to 1} r(\lambda) = \infty$.*
6. $(D\mu)'' - D''\mu$ is a positive function[4] of λ on $[0,1]$ where $\mu(\lambda) = \lambda(1 + r(\lambda))$ representes the debt and its interest per unit of invested capital.

Points 1 and 2 imply that the discount factor $D(\cdot)$ is a continuous function, increasing to the left of λ^* and decreasing afterward. Point 3 indicates that, when leverage reaches 1, the firm is no longer able to access financial markets and has to declare bankruptcy. To ensure continuity of the value function, the value to existing investors (shareholders and creditors) is then assumed to be equal to zero. Alternatively, we could have assumed the market for corporate control provides an exogenous liquidation value, for example book value. This extension is discussed in Section 5. The same comment also applies to the assumption that $\lim_{\lambda \to 1} r(\lambda) = +\infty$. Point 4 implies[5] that $\frac{\partial J_t^*}{\partial \lambda_t}$ converges to 0 as λ_t tends to 1: for highly indebded firms, an slight increase in its leverage has almost no impact of the firm relative value which is already very small. Point 5 implies that $\mu(\cdot)$ is a non-negative, increasing and convex function of λ on $[0,1]$. Point 6 is a sufficient technical condition for the concavity of the relative firm value on some interval $[0,b] \subseteq [0,1]$.

2.4.2   Continuing value

To estimate the continuing value, the firm assumes that from date $(T+1)$ onwards, all values remain constant:

---

[4] f'' stands for the second derivative of f.
[5] The proof is available at Section F of the Technical Report.

- the long-term growth rate of the invested capital is g,[6]

- the ROIC is $\mathrm{E}[z]$

- the leverage ratio is λ_{T+1}, hence the cost of capital is $w(\lambda_{T+1})$

Assumption 2 (*Continuing value parameters*)

$$g - \delta < w\left(\lambda^*\right) < \mathrm{E}\left[z\right].$$

Since for all $s \geq T+1$, $x_{s+1} = \mathrm{E}[z]$, $g_s = g$, $w(\lambda_s) = w(\lambda_{T+1})$, and $I_s = (1+g-\delta)^{s-(T+1)}I_{T+1}$, the continuing value is equivalent to the value of a growing perpetuity of the expected free cash flow at date $(T+1)$, discounted by the cost of capital of the firm at date $(T+1)$:

$$V_{T+1}(\lambda_{T+1}) = \frac{\mathrm{E}[z] - g + \delta}{1 + w\left(\lambda_{T+1}\right)} \sum_{s=T+2}^{\infty} \left(\frac{1 + g - \delta}{1 + w(\lambda_{T+1})}\right)^{s-2-T} = \frac{\mathrm{E}[z] - g + \delta}{w(\lambda_{T+1}) - g + \delta}. \tag{8}$$

Since $w(\lambda_{T+1}) \geq w\left(\lambda^*\right)$, Point 1 of Assumption 2 guarantees that the continuing value is finite and positive.

Assumptions 1 and 2 are sufficient to derive the risk management strategy presented in Sections 4 and 5. They are met for the parameters we estimate in Section 3.

2.5 Dynamic Programming Formulation

From Equation (6), and using the identity

$$I_{t+s} = I_t \prod_{k=0}^{s-1} \left(1 + g_{t+k} - \delta\right), s \geq 1,$$

the average Tobin's q $v_t = \frac{V_t}{I_t}$ is

$$v_t = \mathrm{E}_t \left[\sum_{s=t+1}^{T+1} \left(\prod_{u=t}^{s-2} (1 + g_u - \delta) \right) \left(\prod_{k=t+1}^{s} D\left(\lambda_{k-1}\right) \right) (x_s - g_{s-1} + \delta) \right] \tag{9}$$

$$+ \mathrm{E}_t \left[+ \left(\prod_{s=t}^{T} (1 + g_s - \delta) \right) \left(\prod_{s=t}^{T} D\left(\lambda_s\right) \right) v_{T+1} \right]$$

---

[6] g is not necessarily the growth rate during the last period g_T, since the latter may not be representative of the firm's long-term growth potential.

with the convention that $\prod_{u=t}^{t-1} (1 + g_u - \delta) = 1$.

Because I_t does not depend on future control variables, maximizing firm's value V_t from date t onwards is equivalent to maximizing the average Tobin's q $v_t = V_t/I_t$. This can be cast as a dynamic program. The state of the system at date t is described by the leverage λ_t, and the available investment opportunities i_t. The decision variables or controls are the hedging ratio η_t, the dividend payout ratio d_t, the investment level g_t, and if necessary the exceptional dividend $\tilde{d}_t$.

Conditional on the state of the firm at date t, the relative value of the firm at date t is:

$$J_t^*(\lambda_t, i_t) = \max_{\substack{\{g_s, \eta_s, d_s\}_{s \geq t} \\ s.t.\ (1),\ (2),\ (3),\ (5)}} v_t. \tag{10}$$

Using the continuing value at time $T + 1$ and taking $t = T$ as the last decision point, Appendix B shows that the solution to equation (10) verifies:

$$J_T^*(\lambda_T, i_T) = \max_{\substack{g_T, \eta_T, d_T \\ s.t.\ (1),\ (2),\ (3),\ (5)}} D(\lambda_T) \left(\mathrm{E}[z] - g_T + \delta + (1 + g_T - \delta)\mathrm{E}_T \left[\frac{\mathrm{E}[z] - g + \delta}{w(\lambda_{T+1}) - g + \delta} \right] \right), \tag{11a}$$

$$J_t^*(\lambda_t, i_t) = \max_{\substack{g_t, \eta_t, d_t \\ s.t.\ (1),\ (2),\ (3),\ (5)}} \Theta_t(g_t, \eta_t, d_t; \lambda_t), \quad t = 1, ..., T - 1 \tag{11b}$$

where

$$\Theta_t(g_t, \eta_t, d_t; \lambda_t) = D(\lambda_t) \left(\mathrm{E}[z] - g_t + \delta + (1 + g_t - \delta)\mathrm{E}_t \left[J_{t+1}^*(\lambda_{t+1}(\lambda_t, g_t, d_t, \eta_t), i_{t+1}) \right] \right). \tag{12}$$

As illustrated on Equations (11a), (11b), and (12), leverage interacts with refinancing in two ways. First, it impacts the expected return investors require to hold their investment in the firm, hence the discount rate $D(\lambda_t)$. If the firm's leverage is low, increasing the leverage increases the tax shield, hence reduce the expected return required by investor, for the same free cash flows. If leverage is high, this tax shield effect is overshadowed by a concern that managers/insiders may protect their interests before those of investors/outsiders, hence the latter require a higher expected return.

Second, leverage impacts the incremental cost of financing: as leverage increases, incremental refinancing becomes more expensive. Since it does not change the free cash flows, this does not impact directly the value of the firm. It does, however, impact the value indirectly, through next period's leverage: as leverage λ_t increases, so does the incremental cost of refinancing $r(\lambda_t)$, hence *ceteris paribus* next period leverage λ_{t+1}. The firm may then need more external refinancing. This is the "death spiral", observed in practice.

3   Data and Estimation

We derive in Section 4 the optimal risk management strategy for any functions $w(\cdot)$ and $r(\cdot)$ and parameters $\mathbb{E}[z]$, g, and δ that satisfy assumptions 1 and 2. To illustrate the analysis

Table 1: Descriptive statistics for ROIC, Depreciation and Leverage

	Median	Mean	Std. Dev.	P1	P99
$E[z]$	8.4%	7.9%	11.4%	-35.6%	27.5%
σ	5.8%	8.9%	22.5%	0.9%	69.8%
δ	12.0%	14.4%	8.7%	8.7%	46.8%

The median, mean, standard deviation (Std), and percentiles 1% (P1) and 99% (P99) are based on a sample of 854 industrial firms from annual Compustat files. For each firm, the return on invested capital ($E[z]$), its standard deviation (σ), and the depreciation (δ) are computed for the period 1990 to 2009. The return on invested capital (ROIC) and depreciation are computed according to the procedure described in Section 3.

on a "representative" industrial firm, we estimate these functions and parameters using a balanced panel of 854 industrial firms (first two digits of The North America Industry Classification System (NAICS) between 20 and 39) with annual data covering the period $1990 - 2009$. In total there are $17,080$ firm-year observations. Firm data are taken from Compustat.

3.1   Source of uncertainty

Since the underlying source of uncertainty is not directly observable, we employ the Return On Invested Capital (ROIC) as a proxy for this variable. The ROIC for firm n at the end of the fiscal year t is given by

$$ROIC_{n,t} = \frac{OIADP_{n,t} * (1 - \tau)}{ICAPT_{n,t}}, \tag{13}$$

where OIADP (operating income after depreciation) and ICAPT (invested capital - total) are items from Compustat, and τ is the tax rate, assumed to be constant across firms and years. Following Hennessy and Whited (2007), we select $\tau = 40\%$, corresponding to the average combined federal and state tax for firms in the top tax bracket.

For each firm n, we estimate the average ROIC over the period and its standard deviation across time. Then, to limit the impact of outliers, we use the median values of the resulting distribution for the analysis, as seen on Table 1. ROIC is then assumed to be normally distributed, with expectation $E[z] = 8.4\%$ and standard deviation $\sigma = 5.8\%$.

The standard deviation of the observed ROIC is probably slightly lower than the standard deviation of the underlying source of uncertainty, as the former includes firms' risk management strategies. We show in Section 5 that doubling σ does not change the structure of the optimal risk management strategy.

Table 2: Investment Opportunity and Long-Term Growth Estimates

	$g - \delta$	p	m	σ_H
Point estimate (%)	2.1	21.2	12.6	6.4
Standard deviation (%)	(0.32)	(0.14)	(0.3)	(3.3)

Estimation of Equation (14) is based on a sample of 854 industrial firms from annual Compustat files for the period 1990 to 2009.

3.2 Depreciation

The depreciation rate δ is taken as the median value of δ_t, the cross-firm average annual depreciation rate for year t. Following Eberly *et al.* (2009), δ_t is computed using the double-declining balance method, reflecting accelerated depreciation in the early years of an asset, that is, $\delta_t = 2/L_t$ where L_t is the average useful life of capital goods for year t (in years),

$$L_t = \frac{1}{N} \sum_{n=1}^{N} \frac{PPE_{n,t-1} + DEPR_{n,t-1} + CE_{n,t}}{DEPR_{n,t}},$$

N is the number of firms in the sample, $PPE_{n,t-1}$ is the previous year's book value of property, plant, and equipment (PPEGT item in Compustat), $DEPR_{n,t-1}$ is the depreciation expense (DP item in Compustat), and $CE_{n,t}$ is the capital expenditure incurred in the present year (CAPX item in Compustat). As reported in Table 1, the depreciation rate is estimated at $\delta = 12\%$.

3.3 Investment opportunity and long-term growth

Firms' invested capital increase in every period depends on the arrival (or not) of an investment opportunity. Therefore, we assume that at each period, invested capital grows at an average constant rate g (adjusted for depreciation δ) and, with probability p, increases by a quantity m beyond its normal growth rate. This yields the following panel model for the invested capital $I_{n,t+1}$ of the firm n at time $t+1$:

$$I_{n,t+1} = I_{n,t} \left(1 + g - \delta + m B_{n,t+1} + H_{n,t+1}\right), \tag{14}$$

where $\{B_{n,t}\}_{n,t}$ are i.i.d. Bernoulli variables of parameter p, and $\{H_{n,t}\}_{n,t}$ are i.i.d. centered normal variables of standard deviation σ_H. The parameters are estimated using Maximum Likelihood Estimation (MLE). The MLE estimates, with standard errors in parentheses, of the panel model are given in Table 2.

Equation (14) allows us to identify $(g - \delta)$ only. We use our previous estimate $\delta = 12\%$ to estimate $g = 14.1\%$. Assuming the firms in our sample always invest to the maximum, comparing equations $I_{t+1} = (1 + g_t - \delta) I_t$ and (14) yields $i = m + g - \delta = 14.7\%$.

3.4 Cost of capital and cost of debt

We compute the leverage of the firm as the ratio between the book value of liabilities and the total value of the firm as measured by the book value of its assets and liabilities, that is:

$$\lambda_{n,t} = \frac{D_{n,t}}{D_{n,t} + A_{n,t}}, \tag{15}$$

where $D_{n,t}$ is the sum of debt in current liabilities (DLC in *Compustat*) and long-term debt (DLTT in *Compustat*), and $A_{n,t}$ the total book value of assets (AT in Compustat).

The cost of capital $w(\lambda)$ is computed in a two-step process. In the first step, the cost of capital (WACC) is computed for each firm and for each period using the procedure described in Copeland et al. (2005), that is, the cost of capital for firm n and year t is estimated using the formula

$$WACC_{nt} = (1 - \lambda_{nt})k_{nt}^e + \lambda_{nt}(1 - \tau)k_{nt}^d, \tag{16}$$

where k_{nt}^e and k_{nt}^d are the cost of equity and the cost of debt for firm n and year t, and τ is the tax-rate (40%).

The cost of equity k_{nt}^e is computed following the Capital Asset Pricing Model (CAPM). Stock monthly-return data come from the Center for Research in Security Prices (CRSP). The Treasury bill rate, taken as risk free rate, and the expected market return come from the Fama-French data base available through the Wharton Research Data Services (WRDS). The firms' beta at any point in time is estimated by regressing the firm's stock return on that of the market's return based on the previous five years of monthly prices.

The expected cost of debt k_{nt}^d is obtained as the sum of the ten-year zero-coupon yield plus the expected yield spread according to the firm's rating. Expected corporate yield spreads based on the company's rating come from Chen et al. (2007), while the 10-year risk free zero-coupon yield is computed following the methodology described in Gurkaynak et al. (2007). The historical rating for a firm is obtained from the long-term rating of S&P available in WRDS. Since not all firms have a rating in the S&P rating data base, the resulting dataset is decreased to 340 firms.

The cost of capital is thus computed firm-by-firm on a year basis. Given that yield spreads in Chen et al. (2007) are computed for the period 1995 to 2003, the estimate for the firm's cost of capital, $WACC_n$, is computed as an average of the WACC's year estimates over this period.

The second step is to estimate the function $w(\lambda)$ using $WACC_n$. The functional form proposed for $w(\lambda)$ is

$$w(\lambda) = \frac{(\beta_0 + \beta_1\lambda + \beta_2\lambda^2)}{1 - \lambda}, \tag{17}$$

where β_0, β_1, and β_2 are coefficients to estimate. As required, this functional form for the cost of capital is convex in λ provided that $\beta_0 + \beta_1 + \beta_2 > 0$ and has an asymptote toward ∞ at $\lambda = 1$ if $\beta_2 > 0$. The time-aggregated means of Equations (16) and (15) across

Table 3: Estimates for the WACC Function

	β_0	β_1	β_2
Point estimate	0.1287	−0.4159	0.5389
Standard deviation	(0.008)	(0.071)	(0.139)

Estimation for Equation (17) is based on a sample of 340 firms during the period of 1995 to 2003.

Table 4: Estimates for the Cost of Debt Function

	α_0	α_1
Point estimate	0.0770	−0.0248
Standard deviation	(0.008)	(0.071)

Estimation for Equation (18) is based on a sample of 340 firms during the period of 1995 to 2003.

firms is then used to estimate β_0, β_1, and β_2 in Equation (17) with the Quasi-Maximum Likelihood Estimator (QMLE). The resulting estimates are presented in Table 3. With these parameters, the cost of capital is minimized for a leverage value of $\lambda^* = 31.6\%$ and has a convex shape as shown in Figure 1.

[Insert Figure 1]

The promised yield on debt k^d is assumed to follow

$$k^d\left(\lambda\right) = \frac{\alpha_0 + \alpha_1 \lambda}{1 - \lambda},$$

where α_0, and α_1 are coefficients to estimate and τ is the tax rate. This function is increasing and monotonic in λ provided that $\alpha_0 + \alpha_1 > 0$, which reflects the fact that debt yield increases with leverage. α_0 and α_1 are estimated using the procedure previously described and presented in Table 4. The after tax promised yield used in the analysis is then:

$$r\left(\lambda\right) = \left(1 - \tau\right) k^d\left(\lambda\right) = \left(1 - \tau\right) \frac{\left(\alpha_0 + \alpha_1 \lambda\right)}{1 - \lambda} \tag{18}$$

4   Optimal risk management strategy

In this section, the optimal hedging, investment and dividends policies are presented. We first start with a numerical example which provides intuition for the general case.

The value function $J_t^*(\lambda_t, i_t)$ is represented on Figure 2 for both $i_t = 0$ and $i_t = i$. For consistency, all figures are presented for a date t selected "far enough" from the last period

T that the value function is numerically stationary, that is,

$$\left\| J_t^* - J_{t+1}^* \right\| = \max_{i_t} \left[\max_{\lambda_t} \left| J_t^*(\lambda_t, i_t) - J_{t+1}^*(\lambda_t, i_t) \right| \right] \leq 10^{-4}.$$

This criterion translates into $(T - t) \cong 20$.

[Insert Figure 2]

As seen on Figure 2, $J_t^*(\lambda_t, i_t)$ admits a unique maximum: for small values of λ_t, increasing leverage increases the tax shield, hence the relative value of the firm, whereas for large values of λ_t, increasing leverage decreases the relative value as financial frictions increase the expected return required by investors. We also observe that $J_t^*(\lambda_t, i) \geq J_{t+1}^*(\lambda_t, 0)$: when λ_t is small, the relative value of the firm is larger whenever an investment opportunity is present, while when λ_t is large, the presence of an investment opportunity has no impact on the value of the firm. Figure 2 also illustrates that there exists critical value of λ_t, denoted b_t, such that $J_t^*(\lambda_t, i_t)$ is concave for $\lambda_t \leq b_t$ and convex for $\lambda_t \geq b_t$. In our numerical implementation, the leverage for which the relative value reaches a maximum is 29% when investment opportunity is present, and 31% when the investment opportunity is not present. The change in convexity occurs at leverage ratio of 74%. The shape of the value function is responsible for the optimal risk management strategy discussed below.

4.1 Optimal investment and dividend payments in the full hedging case

We first examine the full hedging case, that is, $\eta_t = 1$. To maximize the value function $\Theta(\lambda_t, g_t, d_t, \eta_t)$ given by expression (12), we compute the first partial derivatives of Θ, hence find critical values that are candidates for the optimal solution. General expressions are presented at Equations (23)-(26) of Appendix A.1. The full hedging case provides simpler expressions for the partial derivatives:

$$\frac{\partial \Theta_t}{\partial d_t} (\lambda_t, g_t, d_t, \eta_t = 1) = D(\lambda_t) \varphi'_{t+1}(y_{t+1}), \tag{19a}$$

$$\frac{\partial \Theta_t}{\partial g_t} (\lambda_t, g_t, d_t, \eta_t = 1) = D(\lambda_t) \phi_{t+1}(y_{t+1}) \tag{19b}$$

where $y_{t+1}(\lambda_t, g_t, d_t)$ is the expected leverage defined at Equation (4), which is equal to λ_{t+1} since $\eta_t = 1$,

$$\varphi_{t+1}(y) = \mathrm{E}_t \left[J_{t+1}^*(y, i_{t+1}) \right]$$

is the firm's relative expected value, viewed as a function of next period's expected leverage, and

$$\phi_{t+1}(y) = \varphi_{t+1}(y) - 1 + (1 - y) \varphi'_{t+1}(y).$$

17

The risk management strategy is derived using a backward induction argument on the shape of φ_{t+1}. Suppose that φ_{t+1} is concave on $[0, b_{t+1}]$ (Induction Hypothesis 1) and reaches a unique maximum denoted $\overline{\lambda}_{t+1} \in (0, b_{t+1})$ such that $\varphi_{t+1}\left(\overline{\lambda}_{t+1}\right) > 1$ (Induction Hypothesis 2). Under these assumptions, we derive the optimal controls (g_t^*, d_t^*, η_t^*). Finally, we prove that, given these controls, φ_t satisfies Induction Hypotheses 1 and 2 on $[0, b_t]$. The intuition for the optimal controls is derived in the main text, while the verification of Induction Hypotheses 1 and 2 is presented in Appendix A.

Suppose φ_{t+1} satisfies Induction Hypotheses 1 and 2 on $[0, b_{t+1}]$. Corollary 2 of Appendix A.2 shows that there exists a unique $\widehat{\lambda}_{t+1} \in \left(\overline{\lambda}_{t+1}, b_{t+1}\right)$ such that ϕ_{t+1} is positive if and only if $y < \widehat{\lambda}_{t+1}$. These critical values $\overline{\lambda}_{t+1} < \widehat{\lambda}_{t+1} < b_{t+1}$ then lead to different thresholds which determine hedging, investment and dividends policies.

Definition 1 *Define* $\lambda_t^{(1)}(i_t) \leq \lambda_t^{(2)}(i_t) \leq \lambda_t^{(3)} \leq \lambda_t^{(4)} \leq b_t \leq b < 1$ *as the unique solution of the following equations:*

$$
\begin{aligned}
\mu\left(\lambda_t^{(1)}(i_t)\right) &= \mathrm{E}[z] + \overline{\lambda}_{t+1}(1 + i_t) - i_t, \\
\mu\left(\lambda_t^{(2)}(i_t)\right) &= \mathrm{E}[z] + \widehat{\lambda}_{t+1}(1 + i_t) - i_t, \\
\mu\left(\lambda_t^{(3)}\right) &= \mathrm{E}[z] + \widehat{\lambda}_{t+1}(1 - \delta) + \delta, \\
\mu\left(\lambda_t^{(4)}\right) &= \mathrm{E}[z] + b_{t+1}(1 - \delta) + \delta,
\end{aligned}
$$

and b *is the unique solution on* $(0, 1)$ *of the fixed point problem*

$$
\mu(b) = \mathrm{E}[z] + b(1 - \delta) + \delta.
$$

From Assumption 1, the function $\mu(\cdot)$ is increasing and convex with $\mu(0) = 0$ and $\lim_{\lambda \to 1} \mu(\lambda) \to \infty$. Consequently, $\lambda_t^{(1)}(0)$ and $\lambda_t^{(2)}(0)$ always exist. If the magnitude of the investment opportunity i is very large, the right hand side of the corresponding equality may be negative and, therefore, $\lambda_t^{(1)}(i)$ or $\lambda_t^{(2)}(i)$ are ill defined. In such cases, they are set to zero. Assumption 1 guarantees existence and unicity in $(0, 1)$ of $\lambda_t^{(3)}\left(\widehat{\lambda}_{t+1}\right)$, $\lambda_t^{(4)}(b_{t+1})$, and b.

$\lambda_t^{(1)}(i_t) \leq \lambda_t^{(2)}(i_t) \leq \lambda_t^{(3)} \leq \lambda_t^{(4)}$ where the first inequality arises because $\overline{\lambda}_{t+1} \leq \widehat{\lambda}_{t+1}$, the second one is a consequence of $\delta > 0$ and the third one is justified because $\widehat{\lambda}_{t+1} < b_{t+1}$. Lemma 3 in Appendix A shows that $\lambda_t^{(4)} \leq b_t \leq b$. Moreover, as the right hand side decreases with the investment opportunity, $\lambda_t^{(1)}(i) \leq \lambda_t^{(1)}(0)$ and $\lambda_t^{(2)}(i) \leq \lambda_t^{(2)}(0)$.

Proposition 1 (Investment and dividends regimes in the full hedging case) *Investment and dividend payout decisions follow four regimes:*

18

1.   For $\lambda_t \in \left[0, \lambda_t^{(1)}(i_t)\right]$, the firm exhibits full financial flexibility: *it fully finances its investment needs, and distributes dividends to reach its optimal leverage ratio.*

2.   For $\lambda_t \in \left(\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)\right]$, the firm faces financial tightness: *its still fully finances its investment needs, but no longer distributes dividends.*

3.   For $\lambda_t \in \left(\lambda_t^{(2)}(i_t), \lambda_t^{(3)}\right]$, the firm faces a financial constraint: *it is no longer able to fully finance its investment needs. The portion it finances is determined to reach its target leverage $\hat{\lambda}_{t+1}$.*

4.   Finally, for $\lambda_t \in \left(\lambda_t^{(3)}, 1\right]$, the firm faces financial hardship: *it is no longer able to finance any of its investment needs, not even depreciation. Its debt ratio increases.*

[Insert Figure 3]

Proof and illustration. In the following $y_{t+1}^* = y_{t+1}(\lambda_t, g_t^*, d_t^*)$ where $g_t^*(\lambda_t)$ and $d_t^*(\lambda_t)$ are the optimal investment level and dividends rate in the full hedging case. In the neighborhood of $y_t = 0$, the partial derivative (19b) is positive (Corollary 2 of Appendix A.2). Therefore, the investment is set to its maximal value, $g_t^* = i_t + \delta$, and the dividends are chosen such that the next period leverage maximizes the firm's relative value:

$$\overline{\lambda}_{t+1} = y_{t+1}^* = \frac{\mu(\lambda_t) + g_t^* - \delta + d_t^* - \mathrm{E}[z]}{1 + g_t^* - \delta} = \frac{\mu(\lambda_t) + i_t + d_t^* - \mathrm{E}[z]}{1 + i_t}$$

which implies that

$$d_t^* = \mathrm{E}[z] + \overline{\lambda}_{t+1}(1 + i_t) - i_t - \mu(\lambda_t) = \mathrm{E}[z] + \overline{\lambda}_{t+1} - \left(1 - \overline{\lambda}_{t+1}\right)i_t - \mu(\lambda_t). \qquad (20)$$

The right hand side is a decreasing function of λ_t. This strategy lasts until the dividend rate is null, that is, for all $\lambda_t \leq \lambda_t^{(1)}(i_t)$. This produces the almost linear decay of the optimal dividends policy on the bottom right panel of Figure 2 and the expected leverage plateau on the left of Figure 3.

For λ_t slightly larger than $\lambda_t^{(1)}(i_t)$, the firm is still financially constraint since $\frac{\partial \Theta_t}{\partial g_t} = D(\lambda_t)\phi(y_{t+1}) > 0$: the optimal control is still $g_t^* = i_t + \delta$ as can be seen on the right panels of Figure 2. In that case,

$$y_{t+1}^*(\lambda_t) = \frac{\mu(\lambda_t) + g_t^* - \delta + d_t^* - \mathrm{E}[z]}{1 + g_t^* - \delta} = \frac{\mu(\lambda_t) + i_t - \mathrm{E}[z]}{1 + i_t}$$

increases almost linearly, as shown in Figure 3. Since next period leverage is an increasing function of λ_t, this situation lasts until $\frac{\partial \Theta_t}{\partial g_t} = D(\lambda_t)\phi(y_{t+1})$ reaches zero, that is, until $\lambda_t \leq \lambda_t^{(2)}(i_t)$ where $\lambda_{t+1}^*\left(\lambda_t^{(2)}(i_t)\right) = \hat{\lambda}_{t+1}$.

Once $\lambda_t > \lambda_t^{(2)}(i_t)$, the optimal investment g_t^* is chosen according to the first order condition $\frac{\partial \Theta_t}{\partial g_t} = D(\lambda_t) \phi_{t+1}\left(y_{t+1}^*\right) = 0$. This happens if $y_{t+1}^* = \widehat{\lambda}_{t+1}$ which produces the second plateau of Figure 3. Therefore,

$$\widehat{\lambda}_{t+1} = \frac{\mu(\lambda_t) + g_t^* - \delta - \mathrm{E}[z]}{1 + g_t^* - \delta},$$

which implies that

$$g_t^* = \frac{\mathrm{E}[z] + \widehat{\lambda}_{t+1}(1 - \delta) + \delta - \mu(\lambda_t)}{1 - \widehat{\lambda}_{t+1}}.$$

This lasts until the optimal investment g_t^* reaches 0. Let $\lambda_t^{(3)}$ be the unique solution to $g_t^*(\lambda_t) = 0$. The decay of the optimal investment policy may be observed at top right panel of Figure 2.

For $\lambda_t > \lambda_t^{(3)}$, the firm is so indebted that it can not invest anymore. Consequently, $J_t^*(\lambda_t, i) = J_t^*(\lambda_t, 0)$. Next period leverage λ_{t+1}^* increases with λ_t. ∎

Equation (20) shows that, ceteris paribus, when dividends are distributed, they increase with expected profits, and decrease with the magnitude of the potential investment opportunity and leverage. Dividend policy is a complex topic, that includes multiple theoretical considerations, ranging from the tax impact of payments to signalling (see for example Welch (2008)). One key finding of all CFOs surveys concerning dividend distribution is managers' extremely reluctance to cut dividends (Lintner (1956), Brav, Graham, Harvey, and Michaely (2005)). This model does not claim to fully explain dividend distribution. However, predictions from the model are consistent with the data.

Fama and French (2002) find that ceteris paribus, when dividends are distributed, they increase with expected profits, and decrease with the magnitude of the potential investment opportunity. Our unreported regression of dividends on leverage and the explanatory variables used in Fama and French (2001) confirm that dividend decrease with leverage. The coefficient of leverage is negative and statistically significant. However, the estimate is much smaller than predicted by the model, due to the stickiness of dividend distributions.

4.2   Optimal hedging

Proposition 2 (Hedging policy) *For any $t \leq T$,*

1.   *The firm fully hedges for $\lambda_t < \lambda_t^{(4)}$*

2.   *The firm gambles for resurrection when $\lambda_t > \lambda_t^{(4)}$.*

Proof. We prove in Appendix A that (i) $\left.\frac{\partial \Theta}{\partial \eta_t}\right|_{\eta_t=1} = 0$ (Equation (26)), hence, $\eta_t = 1$ is always a candidate for the optimal strategy, and (ii)

$$\left.\frac{\partial^2 \Theta}{\partial \eta_t^2}\right|_{\eta_t=1} = D(\lambda_t) \sigma_Z^2 (1 + g_t - \delta)^{-1} \varphi_{t+1}''(y_{t+1}(\lambda_t, g_t, d_t))$$

(Lemma 1). $\eta_t = 1$ is the optimal strategy if and only if $\varphi_{t+1}\left(y^*_{t+1}\right)$ is concave. Using Induction Hypothesis 1, this yields:

$$\eta^*_t = 1 \Leftrightarrow y^*_{t+1} \in [0, b_{t+1}) .$$

From the optimal controls derived in Proposition 1, $y^*_{t+1}\left(\lambda_t\right)$ is a non-decreasing function of λ_t such that

$$y^*_{t+1}\left(\lambda_t\right) \in [0, b_{t+1}) \Leftrightarrow \lambda_t < \lambda^{(4)}_t$$

(a formal proof is provided in Lemma 2 in Appendix A.4). Consequently

$$\eta^*_t = 1 \Leftrightarrow \lambda_t < \lambda^{(4)}_t .$$

For $\lambda_t > \lambda^{(4)}_t$, $\eta_t = 1$ is not optimal. Lemma 1 shows that, in that case, $\eta^*_t = 0$. ∎

Corollary 1 *Investment and dividends policies described at Proposition 1 are optimal for leverage λ_t smaller than $\lambda^{(4)}_t$.*

The hedging decision at date t is driven by the concavity of $\varphi_{t+1}\left(y\right) = \mathrm{E}_t\left[J^*_{t+1}\left(y, i_{t+1}\right)\right]$. For $\lambda_t < \lambda^{(4)}_t$, the firm fully hedges. This result differs from Rochet and Villeneuve (2011) and Bolton *et al.* (2011), who find that, when the firm's cash reserve (or cash-to-capital ratio in Bolton *et al.* (2011)) is high enough, the firm becomes risk neutral, and, since hedging is costly, stops hedging. In our model, the tax shield drives the concavity of the value function, hence the optimality of full hedging. By choosing leverage as the state variable we are able to capture the tax shield from debt, a real effect, absent from Rochet and Villeneuve (2011) and Bolton *et al.* (2011).

For $\lambda_t > \lambda^{(4)}_t$, the firm gambles for resurrection. Since we impose the constraint $\eta_t \geq 0$, it selects $\eta^*_t = 0$. In unreported simulations, if the constraint $\eta_t \geq 0$ is relaxed, the firm buys its risk forward to increase its risk exposure, i.e., chooses $\eta^*_t < 0$. The intuition is that, if $\lambda_t > b$, then $y^*_{t+1}\left(\lambda_t\right) > \lambda_t$: hedging leads to certain bankruptcy. Gambling for resurrection is then optimal. A similar effect can be found in Rochet and Villeneuve (2011), where the firm chooses not to hedge a discrete and large risk.

4.3 Equity cushion

Proposition 3 (Equity cushion) *For any $t \leq T$,*
1.   If the expected ROIC is lower than interest payments (per unit) and the equity fraction of the investment (per unit),

$$\mathrm{E}\left[z\right] < \lambda^* r\left(\lambda^*\right) + \left(1 - \lambda^*\right) i, \tag{21}$$

then for any t, $\bar{\lambda}_t < \lambda^*$, that is, the leverage that maximizes the expected firm relative value φ_t is smaller than the leverage that minimizes the cost of capital w. The difference $\lambda^* - \bar{\lambda}_t > 0$

is the equity cushion.

2.   If $\mathrm{E}\left[z\right] \geq \lambda^* r\left(\lambda^*\right) + \left(1 - \lambda^*\right) i$, *then* $\bar{\lambda}_t = \lambda^*$, *that is, the firm minimizes the static cost of capital.*

Proof. See Appendix A.7.

We show in Appendix A.7 that

$$\varphi_t'\left(\lambda_t\right) = \frac{D'\left(\lambda_t\right)}{D\left(\lambda_t\right)}\varphi_t\left(\lambda_t\right) + pD\left(\lambda_t\right)\left(1 + i\right)\varphi_{t+1}'\left(y_{t+1}^*\left(\lambda_t\right)\right)\frac{\partial y_{t+1}^*}{\partial \lambda_t}\left(\lambda_t, i\right). \tag{22}$$

Equation (22) illustrates why minimizing the cost of capital cannot be optimal. If the firm minimizes its cost of capital, i.e., sets $\lambda_t = \lambda^*$, the first term in Equation (22) is equal to zero. However, if an investment opportunity arises at date t, the firm can no longer optimize its capital structure at $(t + 1)$, and selects $y_{t+1}^*\left(\lambda_t\right) > \bar{\lambda}_{t+1}$, which reduces the continuing value of the firm:

$$\varphi_t'\left(\lambda^*\right) = pD\left(\lambda^*\right)\left(1 + i\right)\varphi_{t+1}'\left(y_{t+1}^*\left(\lambda^*\right)\right)\frac{\partial y_{t+1}^*}{\partial \lambda_t}\left(\lambda^*, i\right) < 0.$$

The firm therefore maintains an equity cushion to protect against that occurrence.

The estimated values of the parameters are such that Inequality (21) holds. Hence $\bar{\lambda}_t < \lambda^*$ for all $t \leq T$: in order to maximize value, the firm does not minimize the cost of capital, i.e., does not exhaust the benefits of the tax shield. The firm pays the static cost $\Delta = w\left(\bar{\lambda}_t\right) - w\left(\lambda^*\right) > 0$ to protect financial flexibility. Δ can be interpreted as an insurance cost, or the cost of financial flexibility. This result is consistent with Rochet and Villeneuve (2011) and Bolton *et al.* (2009), who find that firms optimally hold cash reserves: they accept to pay the opportunity cost of holding these reserves in order to protect their financial flexibility.

This result complements the common wisdom found in the practitioners' literature. First, it confirms that the optimal capital structure is not simply the one that minimizes the cost of capital, as was sometimes incorrectly argued (for example Cohen (2003)). Second, it refines qualitative insights. For example, Pettit (2007) argues (pp. 154-155) that the optimal capital structure depends on the growth prospect of the firm: a high growth firm, e.g., a tech firm, needs lower leverage than a mature, low-growth firm.

This result is consistent with Graham (2000)'s observation that firms do not exhaust the tax advantages of debt. One striking finding in Graham's study is that even profitable, large (hence diversified), liquid firms with low *ex ante* distress cost use debt conservatively, as long as they have growth options. This model explains that observation: condition (21) does not depend on the size of the firm, nor on the volatility of returns. So, for equally profitable firms, the main driver of debt conservativeness is the magnitude of the anticipated investment opportunity, consistent with Graham (2000).

22

For equally profitable firms, the main driver of debt conservativeness is the magnitude of the anticipated investment opportunity, consistent with Graham (2000).

Condition (21) is reversed if $\mathrm{E}[z]$ is much higher and/or i much smaller: highly profitable firms with small investment opportunities do not maintain an equity cushion, rather minimize static cost of capital.

5 Robustness of the results

5.1 Higher volatility

As previously mentioned, we use the volatility of the actual ROICs, which are post risk management activities by the firms, to estimate of the volatility of z_{t+1}. Therefore, we probably underestimate the true underlying volatility. To correct for this bias, we double the volatility of ROIC.

As presented on Figure 4, the results are almost unchanged. For $\lambda_t < \lambda_t^{(4)}$, the firm optimally hedges fully, hence the underlying volatility does not matter. For $\lambda_t > \lambda_t^{(4)}$, the firm's expected relative value J_t^* is higher when the volatility is higher. This is a consequence of the gambling for resurrection: the probability of achieving leverage lower than b, hence the value of the unhedged firm, increases with the underlying volatility.

[Insert Figure 4]

5.2 No Hedging

Suppose now that for an exogenous reason, the firms does not hedge, i.e., set $\eta_t = 0$. We have verified empirically that the shape of value function and the optimal investment and dividend payments strategies are qualitatively unchanged for $\eta_t = \eta^*$ and $\eta_t = 0$. The intuition is as follows: the logic applied to $\lambda_{t+1} = y_{t+1}$, next period's leverage when firms hedge, applies to y_{t+1}, the expected next period leverage when firms do not hedge, as long as truncation is not an issue.

The leverage thresholds, however, are slightly lower, as illustrated on Table 5. Comparison of $\lambda_t^{(1)}$ in Table 5 shows that the presence of hedging allows the firms to pay dividends for higher leverage. Similarly, hedging implies that full investment remains optimal for higher leverage (see $\lambda_t^{(2)}$) and that the firm continues to invest at higher leverage (see $\lambda_t^{(3)}$).

The firm's optimal leverage ratio $\bar{\lambda}_t$ is higher, or equivalently, the precautionary savings is lower by about 2%, when hedging is allowed. This translates into higher relative firm value: the average "hedging premium", defined as

$$\frac{J_t^*(\lambda_t, i_t)_{\eta_t=\eta^*} - J_t^*(\lambda_t, i_t)_{\eta_t=0}}{J_t^*(\lambda_t, i_t)_{\eta_t=0}},$$

23

Table 5: investment and dividends thresholds in the no hedging case

Leverage thresholds (%)	$\bar{\lambda}_t$	$\hat{\lambda}_t$	$\lambda_t^{(1)}(i)$	$\lambda_t^{(1)}(0)$	$\lambda_t^{(2)}(i)$	$\lambda_t^{(2)}(0)$	$\lambda_t^{(3)}$
Hedging	31	40	27	36	37	45	51
No hedging	29	37	25	35	33	42	48

In both cases, the parameters are as in Tables 2, 3 and 4. Critical values are for a selected t "far enough" from the last period T that the value function is numerically stationary, that is, $\left\| J_t^* - J_{t+1}^* \right\| = \max_{i_t} \left[\max_{\lambda_t} \left| J_t^*(\lambda_t, i_t) - J_{t+1}^*(\lambda_t, i_t) \right| \right] \leq 10^{-4}$. This criterion translates into $(T - t) \cong 20$. The cost of capital $w(\cdot)$ is minimized at $\lambda^* = 31.7\%$.

is 7.1% if an investment opportunity is present, 6.7% otherwise, and the maximum premium is 12%. As the optimization of Equations (11a) and (11b) is performed over less control variables, the relative firm value in the no hedging case is smaller or equal to the firm value when hedging is permitted.

Finally, since the optimal leverage ratio is higher, Equation (20) indicates that dividends are also higher when the firm is allowed to hedge.

This analysis suggests that firms hedge to take advantage of the tax shield from debt, not of the convexity of the tax schedule, as was suggested for example by Smith and Stulz (1985). This confirms the widely held view among practitioners that "hedging is tax advantaged equity". This observation is confirmed empirically by Graham and Rogers (2002): firms hedge to increase their leverage, hence the tax shield, hence their value. On their sample, hedging firms increase leverage by 3%, and value by 1.1%. The magnitude of the "hedging premium" is also consistent with Allayanis and Weston (2001), who estimate that hedging firms' average Tobin's q is 5.7% higher than non hedging firms. Since we do not include hedging cost, our estimate is higher than theirs. Taking this effect into account, our model is remarkably close to reality. The value we obtain is larger than Bolton *et al.* (2011), who find only a 1% increase in value arising from hedging. Two reasons explain this difference: first, taxes are included in our analysis. Hedging enables the firm to increase leverage, hence capture tax savings, an effect not included in Bolton *et al.* (2011). Second, we exclude hedging costs.

5.3 Changes in the model's specification

5.3.1 Value function concave with a unique maximum in leverage for $\lambda_t \leq b$

Most of the results hinge on $J_t^*(\lambda_t, i_t)$ being concave with a unique maximum for $0 \leq \lambda_t \leq b < 1$. In our model, this is caused by the convexity of the the expected return required by

investors, i.e., it is produced by the denominator of the relative firm value. We believe this effect is robust to different specifications that incorporate financial frictions in the numerator of the firm value.

Consider first the impact of the tax shield. Suppose for example the tax shield is included in the numerator of the value function as opposed to its denominator, an approach known as the Adjusted Present Value. While there is no consensus on the appropriate discount rate for the tax shield (Welch (2008), pp. 504-507), all agree that the value of the tax shield, hence the value of the firm, increases with the leverage.

Multiple models can then justify why the value function that decreases with leverage. The first example is the inclusion of business disruption costs, developed by Leland (1994) and Leland and Toft (1996). Combined with the tax-shield effect, it produces a value function concave in leverage with a unique maximum.

The following heuristic argument justifies why micro-founded models also produce concave value functions. As mentioned in the introduction, a series of articles derive the dynamically optimal financial contracts resolving the underlying repeated moral-hazard problem between managers/insiders and investors/outsiders, and explore their implementation through standard securities (Biais *et al.* (2011) develop a synthetic model and review the litterature). Biais *et al.* (2007) introduce cash reserve as a state variable, since the limited liability constraints of the manager/insider and financier/outsider impose that negative operating cash flow cannot be financed by the agents. In our model, we allow intermediate refinancing: when the firm's free cash flow is negative, it can borrow to finance the shortfall. Hence, the cash reserve in Biais *et al.* (2007) is replaced in our model by the borrowing capacity $(1 - \lambda)$, held outside of the firm, not inside.

Biais *et al.* (2007) show that the market-to-book ratio of equity is a concave function of cash reserves, with a unique maximum[7]. Furthermore, the market value of stocks and debt goes to zero when the cash reserve goes to zero: the firm has to be liquidated to maintain incentive compatibility. With our re-interpretation as $(1 - \lambda)$ being equivalent to the cash reserve, this translates into the market-to-book ratio being a concave function of λ, with a unique maximum, and $J_t^*(1, i_t) = 0$.

5.3.2 Equity issuance and asset sales

We expect including these possibilities would not significantly alter the optimal strategy, rather would create an upper bound on feasible leverage.

Consider first equity issuance. As is well documented (e.g., Rajan and Zingales (2003)), equity issuance is associated with a significant reduction in share value, most likely due to signalling effect. The firm would then issue equity if and only if this loss is compensated by a relative value increase due to lower leverage. Given the shape of the value function,

---

[7]Biais *et al.* (2007) prove that the book-to-market ratio of equity is a $U-$shaped function of cash reserves, which is equivalent.

this would likely mean that the firm issues equity if and only if leverage exceeds a given threshold.

The same argument applies to asset sales. If the external asset price is higher than the internal value, the firm sells assets. This is simply arbitrage. The more relevant case is when the external asset price is lower than the internal value. The firm would sell assets and use the proceeds to retire debt if and only if the value loss is compensated by a relative value increase due to lower leverage. Given the shape of the value function, this would again translate into a maximum leverage decision rule.

5.3.3 Hedging technology

Our analysis assumes that the firm can (i) hedge all its exposure, and (ii) do so costlessly. (i) is of course not met in practice: no firm can hedge all of its risk. Even if an oil company fully hedges its oil price risk, its operational risk remains. (ii) is not as irrealistic as it seems. Transaction fees, the variable costs of hedging, are small and not essential to the decision of large industrial firms. In general, these do not buy and sell derivatives repeatedly, hence pay transaction fees only once. The collateral firms are required to post when hedging, that is included for example in Bolton *et al.* (2011), constitutes an important cost not included here. Finally, the fixed costs of hedging are significant: hiring of the traders, systems, control, accounting, etc. However, these do not depend on the transaction volume, so they have a limited impact on the hedging ratio, once the decision to set up a hedging group has been made.

The analysis of the no hedging case yields some preliminary insights into the optimal solution in a more general hedging environment, where (i) the firm is able to hedge only a portion of its underlying risk, and (ii) hedging is costly. We expect the firm's risk management strategy will remain qualitatively unchanged, even though the critical values will change. First, the no hedging case shows that the structure of the dividend/investment strategy is unchanged. Second, the hedging strategy will be altered. The optimal hedging ratio balances the marginal cost of hedging against its marginal benefit. The latter arises from the concavity of the objective function, hence hedging will be more valuable as the function becomes more concave. If we assume that the main variable cost is the cost of collateral that need to be posted, the cost should be very low for firms with low leverage (a highly rated firm is required to post less collateral), and increase as firms become more levered. Including costs should marginally reduce the hedging ratio for low-leverage firms, and progressively reduce the optimal hedging ratio as firms are more levered. It will lead to gambling for resurrection earlier, as in Rochet and Villeneuve (2011).

26

6 Concluding remarks

This paper develops a dynamic model to determine a firm's optimal risk management strategy in the presence of two financial frictions: tax deductibility of interest payments and information asymmetry between managers/insiders and investors/outsiders. These frictions are incorporated through the required rate of return, assumed to be a convex function of leverage. The risk management strategy has two elements: first, the firm manages its capital structure through dividend distributions and investment. When leverage is very low, the firm fully replaces depreciated assets, fully invests in opportunities if they arise, and distribute dividends to reach its optimal capital structure. As leverage increases, the firm stops paying dividends, while fully investing. After a certain leverage, the firm also reduces investment, until it stop investing completely. Second, until leverage is very high, the firm fully hedges its operating cash flow exposure, due to the convexity in its cost of capital. When leverage exceeds a very high threshold, around 74% with the parameters we have estimated, the firm gambles for resurrection. The model predictions are consistent with empirical observations.

This analysis can be enriched by several additions. First, a more complex capital structure, including equity issuance could be examined. Currently, financing needs are met by debt issuance, and excess cash is used to pay down debt, until all debt is repaid. Similarly, sale of assets as a means to reduce leverage could be introduced.

Second, we would like to examine richer hedging environment. Multiple risk factors, not all hedgeable, could be included. Except for commodity producers, very few firms face a single risk factor. Determining the overall hedging strategy and, in particular, the trade-offs between hedging the different risks would be very valuable. Similarly, non-linear hedging strategies could be examined. Senior managers and Boards are often reluctant to hedge their risk using forward contracts, as it deprives them of the potential upside, should the output price go up (or the input price go down). Investing in options might prove more acceptable to them, if the value was clearly identified.

Third, we would like to introduce cash reserves as another state variable. In this model, firms are assumed to always be able to finance themselves on external markets, albeit at an increasing cost. Holding cash reserves would enable firms to lower the debt increase required to finance their investment.

These extensions would render the analysis closer to the reality and choices faced by firms, hence would provide clear and practical guidance as firms strive to define their risk management strategy.

Finally, we would like to refine the empirical analysis by testing the predictions of these richer models on firms of varying characteristics (e.g., growth rate and intrinsic volatility).

References

[1] Eberly, J., S. Rebelo, and N. Vincent, 2009, Investment and value: a neoclassical benchmark, CIRPEE working paper, march.

[2] Allayanis, Y. and J. Weston, 2001, The Use of Foreign Currency Derivatives and Firm Market Value, *Review of Financial Studies* 14, 243-276.

[3] Biais, Mariotti, Plantin, and Rochet, 2007, Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications, Review of Economic Studies 74, 345-390.

[4] Biais, Mariotti, and Rochet, 2011, Dynamic Financial Contracting, working paper, Toulouse School of Economics.

[5] Bolton, P., H. Chen, N. Wang, 2011, A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management, Journal of finance, 66, 1545-1578.

[6] Brav, A., J. R. Graham, C. R. Harvey and R. Michaely, 2005, Payout Policy In The 21st Century, *Journal of Financial Economics* 77, 483-527.

[7] Chen, L., D.A. Lesmond, and J. Wei, 2007, Corporate Yield Spreads and Bond Liquidity, The Journal of Finance 62, 119-149.

[8] Cohen, R., 2003, An Analytical Process for Generating the WACC Curve and Locating the Optimal Capital Structure, Wilmott magazine, 86-95.

[9] Copeland, T., T. Koller, and J. Murrin, 1995, *Valuation: Measuring and Managing the Value of Companies,* John Wiley and Sons, Inc. 558 pages.

[10] Copeland, T. F. Weston and K. Shastri, 2005, Financial Theory and Corporate Policy, Fourth edition, Pearson International Edition, 1000 pages.

[11] Fama, E. F. and K. R. French, 2001, Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay? *Journal of Financial Economics* 60, 3–43.

[12] Fama, E. F. and K. R. French, 2002, Testing Trade-Off and Pecking Order Predictions about Dividends and Debt, *Review of Financial Studies*, 15, 1-33.

[13] Froot, K., D. Sharfstein, and J. Stein, 1993, Risk Management: Coordinating Corporate Investment and Financing Policies, Journal of Finance 48, 1629-1658.

[14] Froot, K. and J. Stein, 1998, Risk Management, Capital Budgeting, and Capital Structure Policy for Financial Institutions: an Integrated Approach, Journal of Financial Economics 47, 55-82.

[15] Graham, J. R., 2000, How Big Are the Tax Benefits of Debt? *The Journal of Finance* 55, 1901–1941.

[16] Graham, J. R and C. R. Harvey, 2001, The Theory and Practice of Corporate Finance: Evidence from the Field, *Journal of Financial Economics* 60, 187–243.

[17] Graham, J. R. and D. A. Rogers, 2002, Do Firms Hedge in Response to Tax Incentives?, Journal of Finance 57, 815-839.

[18] Gurkaynak, R., Sack, B. and J. Wright, 2007, The U.S. Treasury Yield curve: 1961 to the Present, Journal of Monetary Economics 54, 2291-2304.

[19] Henessey, C. A. and T. M. Whited 2007, How costly is external financing? Evidence from a structural estimation, The Journal of Finance, (42) 4, pages 705-745.

[20] Holmström, B. and J. Tirole, 2000, Liquidity and Risk Management, Journal of Money, Credit and Banking 32, 295-319.

[21] Léautier, T. O., J.-C. Rochet, and S. Villeneuve, 2007, Defining Risk Appetite, Working paper, Toulouse School of Economics.

[22] Leland, H. E, 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, Journal of Finance 49, 1213-52.

[23] Leland, H. E. et K. Toft, 1996, Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads, *The Journal of Finance* 51, 987-1019.

[24] Lintner, J., 1956, Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes, *The American Economic Review,* 46, 97-113.

[25] Miller, M. and F. Modigliani, 1958, The Cost of Capital, Corporation Finance and the Theory of Investment, The American Economic Review 48, 261-297.

[26] Miller, M. and F. Modigliani, 1963, Corporate Income Taxes and the Cost of Capital: A Correction, The American Economic Review 53, 433-92.

[27] Myers, S. C. and N. S. Majluf, 1984, Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not Have, *Journal of Financial Economics*, 13, 187-221.

[28] Pettit, J., 2007, Strategic corporate finance, applications in valuation & capital structure, Wiley Finance, 284 pages

[29] Rajan, R. and L. Zingales, 2003, The Great Reversals: The Politics of Financial Development in the 20th Century, Journal of Financial Economics 69, 5-50.

[30] Rochet, J.-C. and S. Villeneuve, 2011, Liquidity Risk and Corporate Demand for Hedging and Insurance, Journal of Financial Intermediation 3, 300-323.

[31] Smith, C. W. and R. Stulz, 1985, The Determinants of Firms' Hedging Policies, Journal of Financial and Quantitative Analysis 20, 391-405.

[32] Tirole, J., 2006, *The Theory of Corporate Finance*, Princeton University Press. 644 pages.

[33] Welch, I. 2008, Corporate Finance: An Introduction, Prentice-Hall.

A   Proofs of Propositions 1, 2 and 3

The first-order conditions are derived in Appendix A.1. The induction hypotheses on the shape of φ_{t+1} are formally defined in Appendix A.2. The optimal controls, given the induction hypotheses, are derived in Appendices A.3 and A.4. The induction hypotheses for φ_t are derived in Appendices A.5 and A.6. The equity cushion is derived in Appendix A.7.

A.1   The first order conditions

To optimize the firm relative value, the partial derivatives of Equation (12) with respect to each of the control variables are required. Assuming the independence between z_{t+1} and i_{t+1},

$$\frac{\partial \Theta}{\partial d_t} (\lambda_t, g_t, d_t, \eta_t) = D(\lambda_t) \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (\lambda_{t+1}, i_{t+1}) \right] \tag{23}$$

$$- D(\lambda_t) \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (0, i_{t+1}) \right] \Phi(-B_t),$$

$$\frac{\partial \Theta}{\partial g_t} (\lambda_t, g_t, d_t, \eta_t) = D(\lambda_t) \left(\mathrm{E}_t \left[J_{t+1}^* (\lambda_{t+1}, i_{t+1}) \right] - 1 + \mathrm{E}_t \left[(1 - \lambda_{t+1}) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (\lambda_{t+1}, i_{t+1}) \right] \right) \tag{24}$$

$$- D(\lambda_t) \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (0, i_{t+1}) \right] \Phi(-B_t),$$

$$\frac{\partial \Theta}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t \neq 1) = -D(\lambda_t) \frac{1 + g_t - \delta}{1 - \eta_t} \mathrm{E}_t \left[(\lambda_{t+1} - y_{t+1}) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (\lambda_{t+1}, i_{t+1}) \right] \tag{25}$$

$$- D(\lambda_t) \frac{1 + g_t - \delta}{1 - \eta_t} y_{t+1} \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (0, i_{t+1}) \right] \Phi(-B_t),$$

$$\frac{\partial \Theta}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t = 1) = 0 \text{ if } 0 < y_{t+1} < 1, \tag{26}$$

where Φ represents the cumulative distribution function of a standard normal random variable and $B_t (y_{t+1}, g_t, d_t, \eta_t) = y_{t+1} (\lambda_t, g_t, d_t) \frac{1 + g_t - \delta}{\sigma_Z (1 - \eta_t)}$.

30

Sketch of the proof[8]. Conditional on the available information at time t, λ_{t+1} is a truncated Gaussian random variable, that is $0 \leq \Lambda_{t+1} \leq 1$ if and only if $A_t(\lambda_t, d_t, \eta_t) \leq \varepsilon_{t+1} \leq B_t(\lambda_t, g_t, d_t, \eta_t)$ where

$$A_t(y_{t+1}, g_t, d_t, \eta_t) = \frac{\mu(\lambda_t) + d_t - 1 - \mathrm{E}[z]}{\sigma_Z(1 - \eta_t)} = (y_{t+1}(\lambda_t, g_t, d_t) - 1)\frac{1 + g_t - \delta}{\sigma_Z(1 - \eta_t)},$$

$$\text{and } B_t(y_{t+1}, g_t, d_t, \eta_t) = \frac{\mu(\lambda_t) + g_t - \delta + d_t - \mathrm{E}[z]}{\sigma_Z(1 - \eta_t)} = y_{t+1}(\lambda_t, g_t, d_t)\frac{1 + g_t - \delta}{\sigma_Z(1 - \eta_t)}.$$

Expressing the conditional expectation as an integral,

$$\mathrm{E}_t\left[J_{t+1}^*(\lambda_{t+1}, i_{t+1})\right] = \mathrm{E}_t\left[\int_{-\infty}^{A_t(g_t, d_t, \eta_t)} J_{t+1}^*(1, i_{t+1})\Phi'(\varepsilon)d\varepsilon\right]$$

$$+\mathrm{E}_t\left[\int_{A_t(g_t, d_t, \eta_t)}^{B_t(g_t, d_t, \eta_t)} J_{t+1}^*(\mu_t + \sigma_t\varepsilon, i_{t+1})\Phi'(\varepsilon)d\varepsilon\right] + \mathrm{E}_t\left[\int_{B_t(g_t, d_t, \eta_t)}^{\infty} J_{t+1}^*(0, i_{t+1})\Phi'(\varepsilon)d\varepsilon\right]$$

where the bounds A_t and B_t as well as the conditional mean and standard deviation of λ_{t+1}, are functions of the control variables. Partial derivatives of $\mathrm{E}_t\left[J_{t+1}^*(\lambda_{t+1})\right]$ are obtained from the application of Leibniz integration rule to each of the three integrals. ∎

In the following, we refer to the first order conditions (FOC) the solution to $\frac{\partial \Theta}{\partial u_t}(\lambda_t, g_t^*, d_t^*, \eta_t^*) = 0$, u_t being one of the three control variables.

A.2   Value function at time $t + 1$

The induction hypothesis in broken in two parts:

Induction hypothesis 1 φ_{t+1} *is concave on* $[0, b_{t+1})$.

Induction hypothesis 2 $\varphi_{t+1}(\overline{\lambda}_{t+1}) > 1$ *and* $\varphi_{t+1}\left(\lambda_{t+1}^{(4)}\right) < 1$.

The following mild technical assumption is required in the initiation of the recursion.

Assumption 3 *Shape of the continuing value*
1. $w\left(\lambda_T^{(4)}\right) > \mathrm{E}[z]$.
2. $J_{T+1}(\cdot)$ *is concave on* $[a_{T+1}, b_{T+1}]$ *where* $0 \leq a_{T+1} < \lambda^* < \widehat{\lambda}_{T+1} \leq b_{T+1} \leq 1$.

---

[8]A detailed proof is available at Section C of the technical report.

Point 1 is used at Appendix A.6. Intuitively, if the cost of capital is too low compared to the expected ROIC, the firm has no incentive to control its indebtedness and invests as much as it can. Consequently, we assume there exists a leverage above which the expected return on investment is smaller than the cost of capital, that is $w(\lambda) > \mathrm{E}[z]$ for all sufficiently large λ. Point 2 starts the convexity recursion about the firm's relative expected value.

Corollary 2 *There is a unique $\widehat{\lambda}_{t+1} \in \left[\overline{\lambda}_{t+1}, b_{t+1}\right]$ such that $\phi_{t+1} > 0$ on $\left[0, \widehat{\lambda}_{t+1}\right)$, $\phi_{t+1}\left(\widehat{\lambda}_{t+1}\right) = 0$ and $\phi_{t+1} < 0$ on $\left(\widehat{\lambda}_{t+1}, b_{t+1}\right]$.*

Proof. The function $\phi_{t+1}(\lambda) = \varphi_{t+1}(\lambda) - 1 + (1 - \lambda)\varphi'_{t+1}(\lambda)$ is decreasing on $[0, b_{t+1}]$. Indeed, φ_{t+1} being concave, $\varphi''_{t+1} < 0$

$$\phi'_{t+1}(\lambda) = (1 - \lambda)\varphi''_{t+1}(\lambda) < 0.$$

Moreover,

$$
\begin{aligned}
\phi\left(\overline{\lambda}_{t+1}\right) &= \varphi\left(\overline{\lambda}_{t+1}\right) - 1 > 0, \\
\phi_{t+1}\left(\lambda_{t+1}^{(4)}\right) &= \varphi_{t+1}\left(\lambda_{t+1}^{(4)}\right) - 1 + \left(1 - \lambda_{t+1}^{(4)}\right)\varphi'_{t+1}\left(\lambda_{t+1}^{(4)}\right) < \varphi_{t+1}\left(\lambda_{t+1}^{(4)}\right) - 1 < 0,
\end{aligned}
$$

where the last inequality is justified by Induction Hypothesis 2. Therefore, $\phi(\lambda)$ is positive on $\left[0, \overline{\lambda}_{t+1}\right]$, remains positive for a while before it reaches 0 and becomes negative. ■

Verification of Induction Hypotheses 1 and 2 at time T is similar to the case $t < T$ presented in Appendices A.5 and A.6. Therefore, we refer to Appendix G of the technical report for a detailed proof.

A.3 Investment and dividends policies in the full hedging case given properties of value function at date $(t + 1)$

See Section 4.1.

A.4 Hedging policy given properties of value function at date $(t + 1)$

Lemma 1 *Let φ''_{t+1} stands for the second derivative of φ_{t+1}, that is, $\varphi''_{t+1}(y) = \mathrm{E}_t\left[\frac{\partial^2 J^*_{t+1}}{\partial \lambda^2_{t+1}}(y, i_{t+1})\right]$.*

1.   *If $\varphi''_{t+1}(y_{t+1}(\lambda)) > 0$ then the optimal hedging decision is $\eta^*_t(\lambda) = 0$.*
2.   *If $\varphi''_{t+1}(y_{t+1}(\lambda)) < 0$ then the optimal hedging decision is $\eta^*_t(\lambda) = 1$.*

Proof of Lemma 1. Equation (26) implies that full hedging always satisfies the first order condition. Equation (25) goes to zero as $\eta_t \to -\infty$. As the hedging parameter is constrainted to be positive, it is not an admissible solution. Hence, there will be some cases where the optimal solution will stand on limit of the constraint $\eta_t \geq 0$, that is $\eta_t^* (\lambda_t) = 0$. To determine whether the full hedging decision is optimal, we consider the second derivative[9]:

$$
\begin{aligned}
\frac{\partial^2 \Theta_t}{\partial \eta_t^2} (\lambda_t, g_t, d_t, \eta_t = 1) &= D(\lambda_t) \frac{\sigma_Z^2}{1 + g_t - \delta} \mathrm{E}_t \left[\frac{\partial^2 J_{t+1}^*}{\partial \lambda_{t+1}^2} (y_{t+1}(\lambda_t, g_t, d_t), i_{t+1}) \right] \\
&= D(\lambda_t) \frac{\sigma_Z^2}{1 + g_t - \delta} \varphi_{t+1}'' (y_{t+1}(\lambda_t, g_t, d_t)).
\end{aligned}
$$

Since $\frac{1}{1+w(\lambda_t)} \frac{\sigma_Z^2}{1+g_t-\delta} > 0$, the sign of φ_{t+1}'' determines the sign of $\frac{\partial^2 \Theta_t}{\partial \eta_t^2}$. If $\varphi_{t+1}'' > 0$, then $\eta_t = 1$ minimizes Θ and cannot be the optimal control. Therefore, $\varphi_{t+1}''(y_{t+1}) > 0$ implies that $\eta_t^*(\lambda_t) = 0$. If $\varphi_{t+1}''(y_{t+1}) < 0$, then $\eta_t(\lambda_t) = 1$ is a local maximum. As there is no other root, it is the optimal decision. ∎

Lemma 2 *The deterministic part of the next period leverage in the full hedging case,*

$$
y_{t+1}^* (\lambda_t) = \frac{\mu(\lambda_t) + g_t^*(\lambda_t) - \delta + d_t^*(\lambda_t) - \mathrm{E}[z]}{1 + g_t^*(\lambda_t) - \delta},
$$

is a continuous and non-decreasing function of λ_t.

Proof. According to the optimal management policies of Proposition 1,

$$
y_{t+1}^* (\lambda_t) = \begin{cases} \overline{\lambda}_t & \text{if } \lambda_t \in \left[0, \lambda_t^{(1)}(i_t)\right] \\ \frac{\mu(\lambda_t) + i_t - \mathrm{E}[z]}{1 + i_t} & \text{if } \lambda_t \in \left(\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)\right] \\ \widehat{\lambda}_{t+1} & \text{if } \lambda_t \in \left(\lambda_t^{(2)}(i_t), \lambda_t^{(3)}\right] \\ \frac{\mu(\lambda_t) - \delta - \mathrm{E}[z]}{1 - \delta} & \text{if } \lambda_t \in \left(\lambda_t^{(3)}, \lambda_t^{(4)}\right] \end{cases} \tag{27}
$$

Continuity arises from the definitions of $\lambda_t^{(1)}(i_t)$, $\lambda_t^{(2)}(i_t)$ and $\lambda_t^{(3)}$. Because μ is an increasing function, y_{t+1}^* is a non-decreasing function of λ_t. ∎

---

[9] Proof at Section E of the technical report.

A.5 Verification of Induction Hypothesis 1 at time t (Concavity of J_t^* over $[0, b_t]$)

Replacing the optimal control variables in Equation (12), the firm relative value at time t becomes $J_t^*(\lambda_t, i_t) =$

$$
\begin{cases}
D(\lambda_t)\left(\mathrm{E}[z] - i_t + (1+i_t)\varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right) & \text{if } \lambda_t \in \left[0, \lambda_t^{(1)}(i_t)\right] \\[2mm]
D(\lambda_t)\left(\mathrm{E}[z] - i_t + (1+i_t)\varphi_{t+1}\left(\frac{\mu(\lambda_t)+i_t-\mathrm{E}[z]}{1+i_t}\right)\right) & \text{if } \lambda_t \in \left(\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)\right] \\[2mm]
D(\lambda_t)\left(\mathrm{E}[z] + \left(\varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right)-1\right)\frac{\mathrm{E}[z]+\widehat{\lambda}_{t+1}-\mu(\lambda_t)}{1-\widehat{\lambda}_{t+1}} + \varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right)\right) & \text{if } \lambda_t \in \left(\lambda_t^{(2)}(i_t), \lambda_t^{(3)}\right] \\[2mm]
D(\lambda_t)\left(\mathrm{E}[z] + \delta + (1-\delta)\varphi_{t+1}\left(\frac{\mu(\lambda_t)-\delta-\mathrm{E}[z]}{1-\delta}\right)\right) & \text{if } \lambda_t \in \left(\lambda_t^{(3)}, \lambda_t^{(4)}\right]
\end{cases}
\tag{28}
$$

For $\lambda_t \in \left[0, \lambda_t^{(1)}(i_t)\right]$, $J_t^*(\lambda_t, i_t) = D(\lambda_t)\left(\mathrm{E}[z] - i_t + (1+i_t)\varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)$ which is concave on $\left[0, \lambda_t^{(1)}(i_t)\right]$ because the discount factor is.

For $\lambda_t \in \left(\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)\right]$, the concavity is verified by looking at the second derivative of J_t^* which must be negative. Equation (28) leads to

$$
\begin{aligned}
\frac{\partial^2 J_t^*}{\partial \lambda_t^2}(\lambda_t, i_t) &= \frac{D''(\lambda_t)}{D(\lambda_t)} J_t^*(\lambda_t, i_t) + \left(2D'(\lambda_t)\mu'(\lambda_t) + D(\lambda_t)\mu''(\lambda_t)\right)\varphi_{t+1}'\left(y_{t+1}^*(\lambda_t)\right) \\
&\quad + D(\lambda_t)\varphi_{t+1}''\left(y_{t+1}^*(\lambda_t)\right)\frac{\left(\mu'(\lambda_t)\right)^2}{1+i_t}.
\end{aligned}
$$

$D'' < 0$ by Assumption 1 making the first term negative. $\lambda_t \in (\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)]$ implies that $\overline{\lambda}_{t+1} \leq y_{t+1}^*(\lambda_t) \leq \widehat{\lambda}_{t+1}$. Therefore, $\varphi_{t+1}'\left(y_{t+1}^*(\lambda_t)\right) < 0$. Consequently, the second term is negative provided that $2D'\mu' + D\mu'' > 0$. $(D\mu)'' = D''\mu + 2D'\mu' + D\mu''$ implies that $2D'\mu' + D\mu'' = (D\mu)'' - D''\mu > 0$ by Assumption 1. Induction Hypothesis 1 implies that $\varphi_{t+1}''\left(y_{t+1}^*(\lambda_t)\right) < 0$, therefore, the third term is negative.

Consider $\lambda_t \in (\lambda_t^{(2)}(i_t), \lambda_t^{(3)}]$. First, note that $\varphi_{t+1}(\widehat{\lambda}_{t+1}) > 1$. Indeed,

$$
0 = \phi_{t+1}\left(\widehat{\lambda}_{t+1}\right) = \varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) - 1 + \left(1 - \widehat{\lambda}_{t+1}\right)\varphi_{t+1}'\left(\widehat{\lambda}_{t+1}\right) < \varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) - 1
$$

where the last inequality holds because $\widehat{\lambda}_{t+1} > \overline{\lambda}_{t+1}$ implies that φ_{t+1} is decreasing at $\widehat{\lambda}_{t+1}$. Second, the concavity is obtained by means of the second derivative. Indeed, applying the derivative operator to Equation (28) leads to

$$
\frac{\partial^2 J_t^*}{\partial \lambda_t^2}(\lambda_t, i_t) = \frac{D''(\lambda_t)}{D(\lambda_t)} J_t^*(\lambda_t, i_t) - \left(2D'(\lambda_t)\mu'(\lambda_t) + D(\lambda_t)\mu''(\lambda_t)\right)\frac{\varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) - 1}{1 - \widehat{\lambda}_{t+1}}.
$$

$D'' < 0$ implies the first term is negative. Since $\varphi_{t+1}(\widehat{\lambda}_{t+1}) > 1$, then $\frac{\varphi_{t+1}(\widehat{\lambda}_{t+1})-1}{1-\widehat{\lambda}_{t+1}}$ is positive. $2D'\mu' + D\mu'' = (D\mu)'' - D''\mu > 0$ by Assumption 1.

Finally, let $\lambda_t \in (\lambda_t^{(3)}, \lambda_t^{(4)}]$. As there is no investment, the value function remains the same whenever an investment opportunity is present or not, that is to say, $J_t^*(\lambda_t, i) = J_t^*(\lambda_t, 0)$.

$$
\frac{\partial^2 J_t^*}{\partial \lambda_t^2}(\lambda_t) = \frac{D''(\lambda_t)}{D(\lambda_t)} J_t^*(\lambda_t) + (2D'(\lambda_t)\mu'(\lambda_t) + D(\lambda_t)\mu''(\lambda_t))\varphi_{t+1}'(y_{t+1}^*(\lambda_t))
$$
$$
+ D(\lambda_t)\mu'(\lambda_t)\varphi_{t+1}''(y_{t+1}^*(\lambda_t)).
$$

$\lambda_t \in (\lambda_t^{(3)}, \lambda_t^{(4)}]$ implies that $y_{t+1}^*(\lambda_t) = \frac{\mu(\lambda_t)-\delta-\mathrm{E}[z]}{1-\delta} > \widehat{\lambda}_{t+1}$. Hence, $\varphi_{t+1}'(y_{t+1}^*(\lambda_t)) < 0$. $D'' < 0$ implies that the first term of $\frac{\partial^2 J_t^*}{\partial \lambda_t^2}$ is negative. The second term is negative since $2D'\mu' + D\mu'' = (D\mu)'' - D''\mu > 0$ by Assumption 1. Finally, Induction Hypothesis 1 implies that φ_{t+1} is concave on $[0, b_{t+1}]$. Hence $\varphi_{t+1}''(y_{t+1}^*(\lambda_t))$ and, consequently, the third term are negative for $y_{t+1}^*(\lambda_t) = \frac{\mu(\lambda_t)-\delta-\mathrm{E}[z]}{1-\delta} \leq b_{t+1}$, that is, for $\lambda_t \leq \lambda_t^{(4)}$.

Since the first two terms are negative and the third equal to zero, $\frac{\partial^2 J_t^*}{\partial \lambda_t^2}\left(\lambda_t^{(4)}\right) < 0$ hence $\frac{\partial^2 J_t^*}{\partial \lambda_t^2} < 0$ on $(\lambda_t^{(3)}, \lambda_t^{(4)}]$. Then $\lambda_t^{(4)} \leq b_t$ defined by $\varphi_t''(b_t) = 0$.∎

A.6 Verification of the Induction Hypothesis 2 at time t

A.6.1 $\varphi_t(\overline{\lambda}_t) \geq 1$

We will show that on $[0, \lambda_t^{(3)}]$, $D^{-1}(\lambda_t) J_t^*(\lambda_t, i_t) \geq 1 + \mathrm{E}[z]$. Hence, since $\lambda^* \in [0, \lambda_t^{(3)}]$, then $J_t^*(\lambda^*, i_t) \geq \frac{1+\mathrm{E}[z]}{1+w(\lambda^*)} > 1$. Therefore,

$$
\varphi_t(\overline{\lambda}_t) \geq \varphi_t(\lambda^*) = J_t^*(\lambda^*, i)p + J_t^*(\lambda^*, 0)(1-p) \geq 1.
$$

On $[0, \lambda_t^{(1)}(i_t)]$, the Induction Hypothesis 2 implies that $\varphi_{t+1}(\overline{\lambda}_{t+1}) > 1$. Therefore, starting from Equation (28),

$$
D^{-1}(\lambda_t) J_t^*(\lambda_t, i_t) = \mathrm{E}[z] - i_t + (1+i_t)\varphi_{t+1}(\overline{\lambda}_{t+1}) \geq \mathrm{E}[z] + 1.
$$

Because $\lambda_t \in (\lambda_t^{(1)}(i_t), \lambda_t^{(2)}(i_t)]$ implies that $\overline{\lambda}_{t+1} \leq \lambda_{t+1}^*(\lambda_t) \leq \widehat{\lambda}_{t+1}$, then

$$
\varphi_{t+1}(\lambda_{t+1}^*(\lambda_t)) \geq \varphi_{t+1}\left(\lambda_{t+1}^*\left(\lambda_t^{(2)}(i_t)\right)\right) = \varphi_{t+1}(\widehat{\lambda}_{t+1}) > 1.
$$

Therefore, Equation (28) leads to

$$
D^{-1}(\lambda_t) J_t^*(\lambda_t, i_t) = \mathrm{E}[z] - i_t + (1+i_t)\varphi_{t+1}(\lambda_{t+1}^*(\lambda_t)) \geq \mathrm{E}[z] + 1.
$$

35

On the interval $\lambda_t \in (\lambda_t^{(2)}(i_t), \lambda_t^{(3)}]$,

$$\frac{\mathrm{E}[z] + \widehat{\lambda}_{t+1} - \mu(\lambda_t)}{1 - \widehat{\lambda}_{t+1}} \geq \frac{\mathrm{E}[z] + \widehat{\lambda}_{t+1} - \mu\left(\lambda_t^{(3)}\right)}{1 - \widehat{\lambda}_{t+1}} = -\delta$$

Therefore, starting from Equation (28),

$$
\begin{aligned}
D^{-1}(\lambda_t) J_t^*(\lambda_t, i_t) &= \mathrm{E}[z] + \left(\varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) - 1\right) \frac{\mathrm{E}[z] + \widehat{\lambda}_{t+1} - \mu(\lambda_t)}{1 - \widehat{\lambda}_{t+1}} + \varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) \\
&\geq \mathrm{E}[z] - \delta\left(\varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) - 1\right) + \varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) \\
&= \mathrm{E}[z] + \delta + (1 - \delta)\,\varphi_{t+1}\left(\widehat{\lambda}_{t+1}\right) \\
&\geq \mathrm{E}[z] + \delta + (1 - \delta) \\
&= \mathrm{E}[z] + 1. \quad \blacksquare
\end{aligned}
$$

A.6.2 $\varphi_t\left(\lambda_t^{(4)}\right) < 1$

Lemma 3 *For all $t \leq T$,*

$$b_{t+1} \leq \lambda_t^{(4)} \leq b_t \leq b$$

Proof of Lemma 3. As we have seen from the analysis of the concavity of $J_t^*(\lambda_t, i_t)$, $\lambda_t^{(4)} \leq b_t$. We claim that $\lambda_t^{(4)} \leq b$. Indeed, because b is the fixed point of $y_{t+1}^*(b) = b$, $y_{t+1}^*(\lambda_t) \leq \lambda_t \Leftrightarrow \lambda_t \leq b$. It implies that if $\lambda_t > b$, then $y_{t+1}^*(\lambda_t) > \lambda_t$, hence the fully-hedging firm becomes bankrupt with probability one. Therefore, it cannot be optimal for the firm to fully hedge for $\lambda_t > b$, implying that $\lambda_t^{(4)} \leq b$. From this last inequality,

$$b_{t+1} = y_{t+1}^*\left(\lambda_t^{(4)}\right) \leq \lambda_t^{(4)} \leq b_t$$

and this is verified for all $t \leq T$.$\blacksquare$

Lemma 3 implies that $\lambda_{t+1}^4 \leq b_{t+1} = y_{t+1}^*\left(\lambda_t^{(4)}\right)$. Since φ_{t+1} is decreasing on the right of $\overline{\lambda}_{t+1}$, then $\varphi_{t+1}\left(y_{t+1}^*\left(\lambda_t^{(4)}\right)\right) \leq \varphi_{t+1}\left(\lambda_{t+1}^{(4)}\right)$. Hence,

$$
\begin{aligned}
\varphi_t\left(\lambda_t^{(4)}\right) &= D\left(\lambda_t^{(4)}\right)\left(\mathrm{E}[z] - \delta + (1 - \delta)\varphi_{t+1}\left(y_{t+1}^*\left(\lambda_t^{(4)}\right)\right)\right) \quad \text{(Equation 28)} \\
&\leq D\left(\lambda_t^{(4)}\right)\left(\mathrm{E}[z] - \delta + (1 - \delta)\varphi_{t+1}\left(\lambda_{t+1}^{(4)}\right)\right) \\
&\leq D\left(\lambda_t^{(4)}\right)(\mathrm{E}[z] - \delta + (1 - \delta)) \quad \text{(Induction Hypothesis 2)} \\
&\leq \frac{1 + \mathrm{E}[z]}{1 + w\left(\lambda_t^{(4)}\right)}
\end{aligned}
$$

which is strictly smaller than one provided that $w\left(\lambda_t^{(4)}\right) > \mathrm{E}[z]$. But w being increasing to the right of $\overline{\lambda}_{t+1}$, $w\left(\lambda_t^{(4)}\right) \geq w\left(\lambda_T^{(4)}\right) \geq \mathrm{E}[z]$ where the last inequality comes from Assumption 3.∎

A.7 Equity cushion

The proof proceeds in four steps. First, it is shown that if $\lambda_t^1(i) < \lambda^*$, then $\lambda_t^1(i) < \overline{\lambda}_t < \lambda^*$. Second, backward induction is used to show that if $\lambda_T^1(i) < \lambda^*$, then $\lambda_t^1(i) < \lambda^*$ for all t. Third, a necessary and sufficient condition for $\lambda_T^1(i) < \lambda^*$ is given, ending the induction. The fourth step examines the case $\lambda_T^1(i) \geq \lambda^*$.

Part I. Suppose $\lambda_t^{(1)}(i) < \lambda^* \leq \lambda_t^{(1)}(0)$ and consider $\lambda_t \in \left(\lambda_t^{(1)}(i), \lambda^*\right]$. Because $\lambda_t < \lambda_t^{(4)}$, full hedging is optimal. If $i_t = 0$, then $g^*(\lambda_t, i_t) = \delta$ and the optimal dividend rate is chosen such that $y_{t+1}(\lambda_t, g_t^*, d_t^*) = \overline{\lambda}_{t+1}$. If $i_t = i$, then $d^*(\lambda_t, i_t) = 0$ and the investment is chosen such that $y_{t+1}^*(\lambda_t, i) = y_{t+1}(\lambda_t, g_t^*, d_t^*) > \overline{\lambda}_{t+1}$. Therefore

$$
\begin{aligned}
\varphi_t(\lambda_t) &= J_{t+1}^*\left(y_{t+1}^*(\lambda_t, i), i\right)p + J_{t+1}^*\left(\overline{\lambda}_{t+1}, 0\right)(1-p) \\
&= D(\lambda_t)\left(\mathrm{E}[z] - i + (1+i)\varphi_{t+1}\left(y_{t+1}^*(\lambda_t)\right)\right)p + D(\lambda_t)\left(\mathrm{E}[z] + \varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)(1-p)
\end{aligned}
$$

and

$$
\varphi_t'(\lambda_t) = \frac{D'(\lambda_t)}{D(\lambda_t)}\varphi_t(\lambda_t) + D(\lambda_t)(1+i)\varphi_{t+1}'\left(y_{t+1}^*(\lambda_t)\right)\frac{\partial y_{t+1}^*}{\partial \lambda_t}(\lambda_t, i)p.
$$

$\varphi_t'\left(\lambda_t^{(1)}(i)\right) > 0$ since (i) $y_{t+1}^*\left(\lambda_t^{(1)}(i)\right) = \overline{\lambda}_{t+1}$, (ii) $\varphi_{t+1}'\left(\overline{\lambda}_{t+1}\right) = 0$ and (iii) $D'(\lambda_t) > 0$, λ_t being smaller that λ^*. Moreover, since $D'(\lambda^*) = 0$,

$$
\varphi_t'(\lambda^*) = D(\lambda^*)(1+i)\varphi_{t+1}'\left(y_{t+1}^*(\lambda^*)\right)\frac{\partial y_{t+1}^*}{\partial \lambda_t}(\lambda^*, i)p
$$

which is negative because $\frac{\partial y_{t+1}^*}{\partial \lambda_t} > 0$ and $y_{t+1}^*(\lambda^*) \geq \overline{\lambda}_{t+1}$ implies that $\varphi_{t+1}'\left(y_{t+1}^*(\lambda^*)\right) \leq 0$. By continuity of φ_t', since $\varphi_t'\left(\lambda_t^{(1)}(i)\right) > 0$ and $\varphi_t'(\lambda^*) \leq 0$, there is $\overline{\lambda}_t \in \left(\lambda_t^{(1)}(i), \lambda^*\right]$ such that $\varphi_t'\left(\overline{\lambda}_t\right) = 0$.

Part II. Assume that $\lambda_T^{(1)}(i) < \lambda^*$. A backward induction argument is used to show that $\lambda_t^{(1)}(i) < \lambda^*$. Indeed, from Part I, if $\lambda_{t+1}^{(1)}(i) < \lambda^*$, then $\overline{\lambda}_{t+1} \in \left(\lambda_{t+1}^{(1)}(i), \lambda^*\right]$. Therefore, because $\mu\left(\lambda_t^{(1)}(i_t)\right) = \mathrm{E}[z] + \overline{\lambda}_{t+1} - \left(1 - \overline{\lambda}_{t+1}\right)i_t$

$$
\mu\left(\lambda_T^{(1)}(i)\right) - \mu\left(\lambda_t^{(1)}(i)\right) = \left(\lambda^* - \overline{\lambda}_{t+1}\right)(1+i) > 0.
$$

Hence, as μ is an increasing function, $\lambda_t^{(1)}(i) < \lambda_T^{(1)}(i) \leq \lambda^*$.

Part III. $\lambda_T^{(1)}(i) < \lambda^*$ if and only if

$$
\begin{aligned}
0 \; &< \; \mu(\lambda^*) - \mu\left(\lambda_T^{(1)}(i)\right) \\
&= \; \lambda^*(1 + r(\lambda^*)) - \mathrm{E}[z] - \lambda^* + (1 - \lambda^*)i \\
&= \; \lambda^* r(\lambda^*) - \mathrm{E}[z] + (1 - \lambda^*)i
\end{aligned}
$$

that is, the expected ROIC is lower than interest payments (per unit) and the equity fraction of the investment (per unit):

$$
\mathrm{E}[z] < \lambda^* r(\lambda^*) + (1 - \lambda^*)i.
$$

Part IV. What happens if $\lambda_T^{(1)}(i) \geq \lambda^*$? One can show by induction that $\overline{\lambda}_t = \lambda^*$ for any t. Indeed, if $\overline{\lambda}_{t+1} = \lambda^*$, then

$$
\mu\left(\lambda_T^{(1)}(i)\right) - \mu\left(\lambda_t^{(1)}(i)\right) = \left(\lambda^* - \overline{\lambda}_{t+1}\right)(1 + i) = 0,
$$

implying that $\lambda_t^{(1)}(i) = \lambda_T^{(1)}(i) \geq \lambda^*$. But, for any $\lambda_t \leq \lambda^* \leq \lambda_t^{(1)}(i)$,

$$
\varphi_t(\lambda_t) = D(\lambda_t)\left(\mathrm{E}[z] - i + (1 + i)\varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)p + D(\lambda_t)\left(\mathrm{E}[z] + \varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)(1 - p)
$$

and

$$
\varphi_t'(\lambda_t) = D'(\lambda_t)\left(\mathrm{E}[z] - i + (1 + i)\varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)p + D'(\lambda_t)\left(\mathrm{E}[z] + \varphi_{t+1}\left(\overline{\lambda}_{t+1}\right)\right)(1 - p).
$$

Because $D'(\lambda^*) = 0$, $\varphi_t'(\lambda^*) = 0$ and $\varphi_t(\lambda_t)$ is maximized at λ^*, that is $\overline{\lambda}_t = \lambda^*$. $\blacksquare$

Figure 1: Cost of capital vs. Leverage
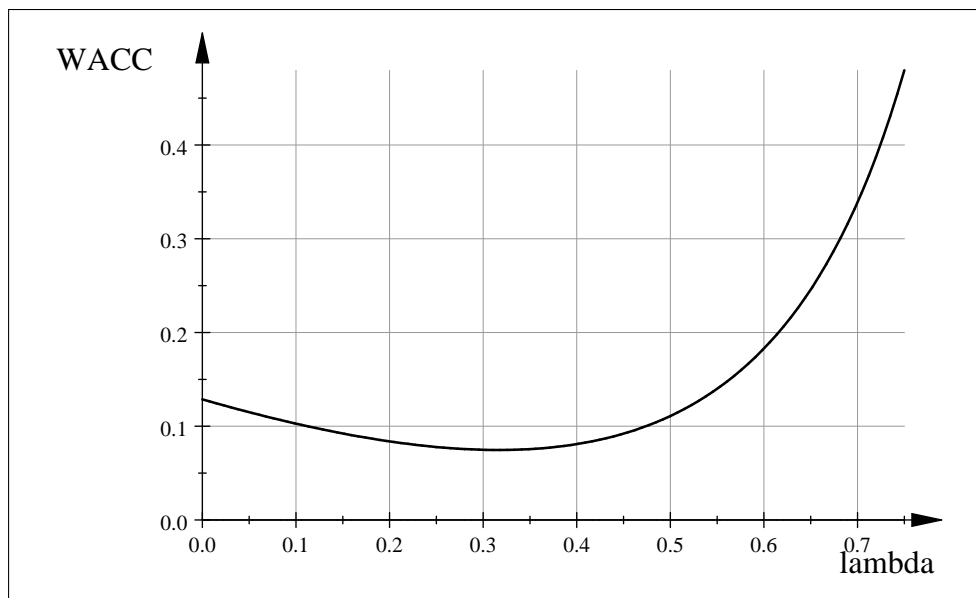
This figure presents the cost of capital as a function of the firm's leverage (Equation (17)). The parameters presented in Table 3 are obtained from a sample of 340 firms during the period of 1995 to 2003.

Figure 2: Firm's Relative Value and Optimal Policies – Multi-period Model
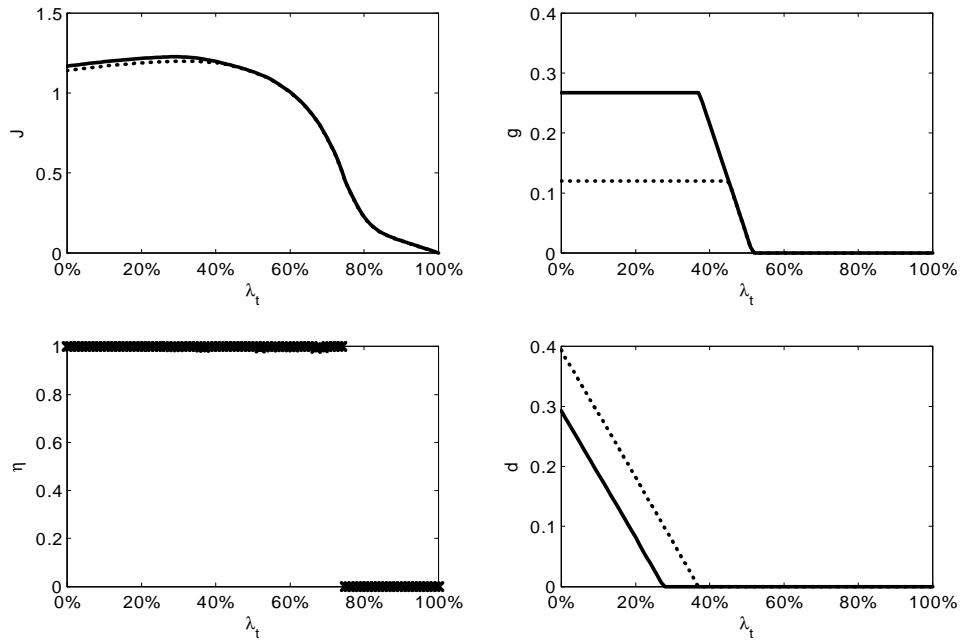
This figure shows the optimal relative value J_T^* (top left figure), the optimal policy for the investment level g_T (top right figure), the optimal policy for the hedging ratio η_T (bottom left figure), the optimal dividend payout d_T (bottom right figure) as a function of leverage. The dotted line represents the case with no investment opportunity in the decision period $(i_T = 0)$, while the continuous line represents the case when the investment opportunity is present $(i_T = 14.7\%)$. The optimal policies and relative value are computed with the numerical implementation described in Appendix H of the technical report. The cost of capital is minimized at $\lambda^* = 31.6\%$. The magnitude of the investment opportunity is equal to 12.6%, which arrives during a period with a probability of 21.2%. The expected ROIC is $\mathrm{E}[z] = 8.4\%$ with standard deviation $\sigma = 5.8\%$ (Table 1).

Figure 3: Expected Leverage Ratio for t+1 – Multi-Period Model
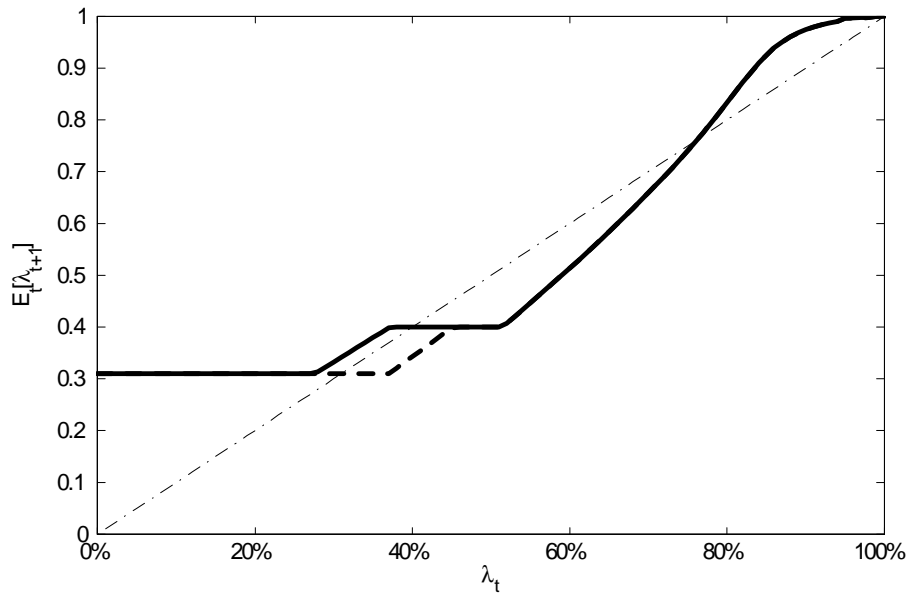
This figure shows the expected leverage at time $t+1$ as a function of the leverage ratio at time t. The dotted line represents the case with no investment opportunity in the decision period $(i_T = 0)$, while the continuous line represents the case when the investment opportunity is present $(i_T = 12.6\%)$. The expected leverage is computed with equation (41) and the associated optimal values for η_T, d_T, and g_T. The cost of capital is minimized at $\lambda_t = 31.6\%$. The magnitude of the investment opportunity is equal to 12.6%, which arrives during a period with a probability of 21.2%. The expected value and standard deviation of the source of uncertainty z is 8% and 10% respectively. The straight line is the curve $f(\lambda) = \lambda$.

Figure 4: Firm relative values in a high volatility framework
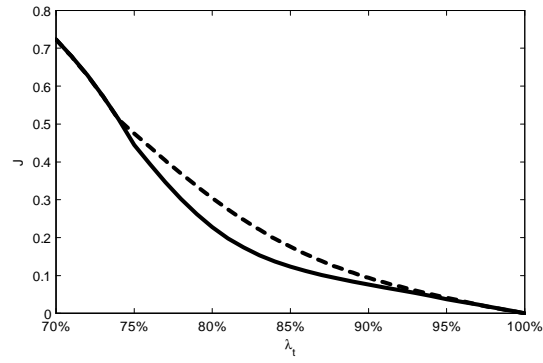
The continuous line represents the relative value of the firm for the base case where all parameters are as in Tables 2, 3 and 4. The dashed line corresponds to the firm relative value when the underlying source of uncertainty (σ) is doubled to 11.6%, all other parameters being for the base case.

Technical Report

Not part of the paper
Available upon request

D. Amaya
G. Gauthier
T.O. Leautier

B Derivation of the Bellman equations

This section deduces the Bellman equations (11a) and (11b). For $t = 0, ..., T-1$, let $J_t^* (\lambda_t, i_t)$ be the optimal value for the $T - t$ stage problem that starts at date t in state (λ_t, i_t), as defined in Equation (10). First, consider the case for $t = T$. Using the continuation value obtained in (8),

$$
\begin{aligned}
J_T^* (\lambda_T, i_T) &= \max_{\substack{g_T, \eta_T, d_T \\ s.t. \ (1), \ (2), \ (3), \ (5)}} \mathrm{E}_T \left[\sum_{s=T+1}^{\infty} \left(\prod_{k=T+1}^{s} D(\lambda_{k-1}) \right) (x_s - g_{s-1} + \delta) \left(\prod_{u=T}^{s-2} (1 + g_u - \delta) \right) \right] \\
&= \max_{\substack{g_T, \eta_T, d_T \\ s.t. \ (1), \ (2), \ (3), \ (5)}} D(\lambda_T) \mathrm{E}_T \left[x_{T+1} - g_T + \delta + (1 + g_T - \delta) \frac{\mathrm{E}[z] - g + \delta}{w(\lambda_{T+1}) - g + \delta} \right],
\end{aligned}
$$

which leads to Equation $(11a)$. Now, from Equation (10), the problem at t can be written as

$$
\begin{aligned}
J_t^* (\lambda_t, i_t) &= \max_{\substack{\{g_s, \eta_s, d_s\}_{s \geq t} \\ s.t. \ (1), \ (2), \ (3), \ (5)}} \mathrm{E}_t \left[\sum_{s=t+1}^{\infty} \left(\prod_{k=t+1}^{s} D(\lambda_{k-1}) \right) (x_s - g_{s-1} + \delta) \left(\prod_{u=t}^{s-2} (1 + g_u - \delta) \right) \right] \\
&= \max_{\substack{\{g_s, \eta_s, d_s\}_{s \geq t} \\ s.t. \ (1), \ (2), \ (3), \ (5)}} D(\lambda_t) \mathrm{E}_t \left[x_{t+1} - g_t + \delta \right. \\
&\quad \left. + (1 + g_t - \delta) \sum_{s=t+2}^{\infty} \left(\prod_{k=t+2}^{s} D(\lambda_{k-1}) \right) (x_s - g_{s-1} + \delta) \left(\prod_{u=t+1}^{s-2} (1 + g_u - \delta) \right) \right].
\end{aligned}
$$

Let $(g_u^*, \eta_u^*, d_u^*)_{u=t+1}^{T}$ be an optimal policy for the subproblem $J_{t+1}^* (\lambda_{t+1}, i_{t+1})$ that starts at date $t + 1$ and has final decision at stage T. Then

$$
J_t^* (\lambda_t, i_t) = \max_{\substack{g_t, \eta_t, d_t \\ s.t. \ (2),(1), \ and \ (3)}} D(\lambda_t) \mathrm{E}_t \left[(x_{t+1} - g_t + \delta) + (1 + g_t - \delta) J_{t+1}^* (\lambda_{t+1}, i_{t+1}) \right]
$$

which is Equation $(11b)$.

C Proofs of the first order conditions (cf. Section A.1)

This Appendix provides detailed proofs of some results stated in Section A.1. Let

$$
\Psi (\lambda_t, g_t, d_t, \eta_t) = \mathrm{E}_t \left[J_{t+1}^* (\lambda_{t+1} (\lambda_t, g_t, d_t, \eta_t), i_{t+1}) \right].
$$

Because $0 \leq \Lambda_{t+1} \leq 1$ if and only if $A_t (\lambda_t, d_t, \eta_t) \leq \varepsilon_{t+1} \leq B_t (\lambda_t, g_t, d_t, \eta_t)$, we define

$$
\widetilde{\varepsilon}_{t+1} = A_t \mathbf{1}_{\varepsilon_{t+1} \leq A_t} + \varepsilon_{t+1} \mathbf{1}_{A_t < \varepsilon_{t+1} < B_t} + B_t \mathbf{1}_{\varepsilon_{t+1} \geq A_t}
$$

and note that

$$\lambda_{t+1} = y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \widetilde{\varepsilon}_{t+1}.$$

The following lemma computes the conditional expectation of the next period firm's relative value, taking care of the truncation effect.

Lemma 4 *If ε_{t+1} and i_{t+1} are independent,*

$$
\begin{aligned}
\Psi\left(\lambda_t, g_t, d_t, \eta_t\right) \;=\;& \mathrm{E}_t\left[J_{t+1}^*\left(1, i_{t+1}\right) \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right)\right] \\
&+ \mathrm{E}_t\left[\int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \\
&+ \mathrm{E}_t\left[J_{t+1}^*\left(0, i_{t+1}\right)\left(1 - \Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right)\right].
\end{aligned}
$$

Proof of Lemma 4. The expected relative firm value is

$$
\begin{aligned}
\Psi\left(\lambda_t, g_t, d_t, \eta_t\right) \;=\;& \mathrm{E}_t\left[J_{t+1}^*\left(\lambda_{t+1}, i_{t+1}\right)\right] \\
\;=\;& \mathrm{E}_t\left[J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \widetilde{\varepsilon}_{t+1}, i_{t+1}\right)\right] \\
\;=\;& \mathrm{E}_t\left[\mathrm{E}_t\left[\left. J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \frac{\widetilde{\varepsilon}_{t+1}}{\sigma_Z}, i_{t+1}\right) \right| i_{t+1}\right]\right].
\end{aligned}
$$

Assuming that ε_{t+1} and i_{t+1} are independent, the inner conditional expectation becomes

$$
\begin{aligned}
&\int_{-\infty}^{A_t(\lambda_t, d_t, \eta_t)} J_{t+1}^*\left(1, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon \\
&+ \int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon \\
&+ \int_{B_t(\lambda_t, g_t, d_t, \eta_t)}^{\infty} J_{t+1}^*\left(0, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon \\
\;=\;& J_{t+1}^*\left(1, i_{t+1}\right) \Phi\left(A_t\right) + \int_{A_t}^{B_t} J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon \\
&+ J_{t+1}^*\left(0, i_{t+1}\right)\left(1 - \Phi'\left(B_t\right)\right)
\end{aligned}
$$

where Φ is the cumulative distribution function of a standard normal random variable and Φ' is its density function. This completes the proof. ∎

The following lemmas compute the partial first derivatives of Λ_{t+1} with respect to each control variables.

Lemma 5

$$\frac{\partial \Lambda_{t+1}}{\partial \lambda_t} (\lambda_t, g_t, d_t, \eta_t) = \frac{\mu'(\lambda_t)}{1 + g_t - \delta},$$

$$\frac{\partial \Lambda_{t+1}}{\partial g_t} (\lambda_t, g_t, d_t, \eta_t) = \frac{1 - \Lambda_{t+1}(\lambda_t, g_t, d_t, \eta_t)}{1 + g_t - \delta},$$

$$\frac{\partial \Lambda_{t+1}}{\partial d_t} (\lambda_t, g_t, d_t, \eta_t) = \frac{1}{1 + g_t - \delta},$$

$$\frac{\partial \Lambda_{t+1}}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t) = \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon_{t+1}$$

$$= \begin{cases} y_{t+1}(\lambda_t, g_t, d_t) - \Lambda_{t+1}(\lambda_t, g_t, d_t, \eta_t) & if \ \eta_t = 1 \\ \frac{y_{t+1}(\lambda_t, g_t, d_t) - \Lambda_{t+1}(\lambda_t, g_t, d_t, \eta_t)}{1 - \eta_t} & if \ \eta_t < 1 \end{cases}$$

Proof. Straight application starting from the definition of Λ_{t+1}.

Lemma 6 *If $\eta_t \neq 1$ then*

$$\frac{\partial \Psi}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t) = \mathrm{E}_t \left[\frac{y_{t+1} - \lambda_{t+1}}{1 - \eta_t} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (\lambda_{t+1}, i_{t+1}) \right]$$

$$- \frac{y_{t+1} - 1}{1 - \eta_t} \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (1, i_{t+1}) \right] \Phi(A_t) - \frac{y_{t+1}}{1 - \eta_t} \mathrm{E}_t \left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} (0, i_{t+1}) \right] (1 - \Phi(B_t)).$$

If $0 < y_{t+1}(\lambda_t, g_t, d_t) < 1$, then $\frac{\partial \Psi}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t = 1) = 0$.

Proof of Lemma 6. Starting using Lemma 4,

$$\frac{\partial \Psi}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t) = \frac{\partial}{\partial \eta_t} \mathrm{E}_t \left[J_{t+1}^* (1, i_{t+1}) \Phi(A_t(\lambda_t, d_t, \eta_t)) \right]$$

$$+ \frac{\partial}{\partial \eta_t} \mathrm{E}_t \left[\int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} J_{t+1}^* \left(y_{t+1}(\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi'(\varepsilon) d\varepsilon \right]$$

$$+ \frac{\partial}{\partial \eta_t} \mathrm{E}_t \left[J_{t+1}^* (0, i_{t+1}) (1 - \Phi(B_t(\lambda_t, g_t, d_t, \eta_t))) \right]$$

$$= \mathrm{E}_t \left[J_{t+1}^* (1, i_{t+1}) \frac{\partial}{\partial \eta_t} \Phi(A_t(\lambda_t, d_t, \eta_t)) \right]$$

$$+ \mathrm{E}_t \left[\frac{\partial}{\partial \eta_t} \int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} J_{t+1}^* \left(y_{t+1}(\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi'(\varepsilon) d\varepsilon \right]$$

$$+ \mathrm{E}_t \left[J_{t+1}^* (0, i_{t+1}) \frac{\partial}{\partial \eta_t} (1 - \Phi(B_t(\lambda_t, g_t, d_t, \eta_t))) \right]$$

where the derivative can be introduced inside the integral using the bounded convergence theorem. Using

Leibniz integral rule[10],

$$\frac{\partial \Psi}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t)$$

$$= \mathrm{E}_t \left[J^*_{t+1} (1, i_{t+1}) \Phi' (A_t (\lambda_t, d_t, \eta_t)) \frac{\partial A_t}{\partial \eta_t} (\lambda_t, d_t, \eta_t) \right]$$

$$+ \mathrm{E}_t \left[\frac{\partial B_t}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t) J^*_{t+1} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} B_t (\lambda_t, g_t, d_t, \eta_t), i_{t+1} \right) \Phi' (B_t (\lambda_t, g_t, d_t, \eta_t)) \right]$$

$$- \mathrm{E}_t \left[\frac{\partial A_t}{\partial \eta_t} (\lambda_t, d_t, \eta_t) J^*_{t+1} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} A_t (\lambda_t, d_t, \eta_t), i_{t+1} \right) \Phi' (A_t (\lambda_t, d_t, \eta_t)) \right]$$

$$+ \mathrm{E}_t \left[\int_{A_t (\lambda_t, d_t, \eta_t)}^{B_t (\lambda_t, g_t, d_t, \eta_t)} \frac{\partial J^*_{t+1}}{\partial \eta_t} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi' (\varepsilon) d\varepsilon \right]$$

$$- \mathrm{E}_t \left[J^*_{t+1} (0, i_{t+1}) \Phi' (B_t (\lambda_t, g_t, d_t, \eta_t)) \frac{\partial B_t}{\partial \eta_t} (\lambda_t, g_t, d_t, \eta_t) \right]$$

$$= \mathrm{E}_t \left[\int_{A_t (\lambda_t, d_t, \eta_t)}^{B_t (\lambda_t, g_t, d_t, \eta_t)} \frac{\partial J^*_{t+1}}{\partial \eta_t} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi' (\varepsilon) d\varepsilon \right]$$

$$= \mathrm{E}_t \left[\int_{A_t (\lambda_t, d_t, \eta_t)}^{B_t (\lambda_t, g_t, d_t, \eta_t)} \frac{\partial \lambda_{t+1}}{\partial \eta_t} \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi' (\varepsilon) d\varepsilon \right] \quad \text{(chain rule)}$$

$$= \mathrm{E}_t \left[\int_{A_t (\lambda_t, d_t, \eta_t)}^{B_t (\lambda_t, g_t, d_t, \eta_t)} \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi' (\varepsilon) d\varepsilon \right] \quad \text{(Lemma 5)}$$

$$= \mathrm{E}_t \left[\int_{A_t (\lambda_t, d_t, \eta_t)}^{B_t (\lambda_t, g_t, d_t, \eta_t)} \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} \left(y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1} \right) \Phi' (\varepsilon) d\varepsilon \right]$$

$$+ \mathrm{E}_t \left[\int_{-\infty}^{A_t (\lambda_t, d_t, \eta_t)} \frac{\sigma_Z}{1 + g_t - \delta} A_t (\lambda_t, d_t, \eta_t) \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} (1, i_{t+1}) \Phi' (\varepsilon) d\varepsilon \right]$$

$$+ \mathrm{E}_t \left[\int_{B_t (\lambda_t, g_t, d_t, \eta_t)}^{\infty} \frac{\sigma_Z}{1 + g_t - \delta} B_t (\lambda_t, g_t, d_t, \eta_t) \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} (0, i_{t+1}) \Phi' (\varepsilon) d\varepsilon \right]$$

$$- \mathrm{E}_t \left[\frac{\sigma_Z}{1 + g_t - \delta} A_t (\lambda_t, d_t, \eta_t) \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} (1, i_{t+1}) \Phi (A_t (\lambda_t, d_t, \eta_t)) \right]$$

$$- \mathrm{E}_t \left[\frac{\sigma_Z}{1 + g_t - \delta} B_t (\lambda_t, g_t, d_t, \eta_t) \frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}} (0, i_{t+1}) (1 - \Phi (B_t (\lambda_t, g_t, d_t, \eta_t))) \right].$$

Since $\lambda_{t+1} = y_{t+1} (\lambda_t, g_t, d_t) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \widetilde{\varepsilon}_{t+1}$, then if $\eta_t \neq 1$,

$$\frac{\sigma_Z}{1 + g_t - \delta} \widetilde{\varepsilon}_{t+1} = \frac{y_{t+1} (\lambda_t, g_t, d_t) - \lambda_{t+1}}{1 - \eta_t}.$$

---

[10] $\frac{d}{d\alpha} \int_{a(\alpha)}^{b(\alpha)} f (x, \alpha) dx = \frac{db}{d\alpha} (\alpha) f (b (\alpha), \alpha) - \frac{da}{d\alpha} (\alpha) f (a (\alpha), \alpha) + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha} (x, \alpha) dx$

Therefore, if $\eta_t \neq 1$, then

$$
\begin{aligned}
\frac{\partial \Psi}{\partial \eta_t}\left(\lambda_t, g_t, d_t, \eta_t\right) = \ & \mathrm{E}_t\left[\frac{y_{t+1}\left(\lambda_t, g_t, d_t\right)-\lambda_{t+1}\left(\lambda_t, g_t, d_t, \eta_t\right)}{1-\eta_t} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}\left(\lambda_t, g_t, d_t, \eta_t\right), i_{t+1}\right)\right] \\
& -\frac{y_{t+1}\left(\lambda_t, g_t, d_t\right)-1}{1-\eta_t} \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right)\right] \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right) \\
& -\frac{y_{t+1}\left(\lambda_t, g_t, d_t\right)}{1-\eta_t} \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right]\left(1-\Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right).
\end{aligned}
$$

If $\eta_t = 1$ and $0 < y_{t+1}\left(\lambda_t, g_t, d_t\right) < 1$, then

$$
\begin{aligned}
\frac{\partial \Psi}{\partial \eta_t}\left(\lambda_t, g_t, d_t, \eta_t = 1\right) & = \mathrm{E}_t\left[\int_{-\infty}^{\infty} \frac{\sigma_Z}{1+g_t-\delta} \varepsilon \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right), i_{t+1}\right) \Phi'(\varepsilon) d\varepsilon\right] \\
& = \frac{\sigma_Z}{1+g_t-\delta} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right), i_{t+1}\right) \mathrm{E}_t\left[\int_{-\infty}^{\infty} \varepsilon \Phi'(\varepsilon) d\varepsilon\right]=0.
\end{aligned}
$$

Therefore, $\eta_t = 1$ is one root of $\frac{\partial \Psi}{\partial \eta_t}\left(\lambda_t, g_t, d_t, \eta_t\right)$. $\blacksquare$

Lemma 7

$$
\begin{aligned}
\frac{\partial \Psi}{\partial d_t}\left(\lambda_t, g_t, d_t, \eta_t\right) = & \frac{1}{1+g_t-\delta} \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}\left(\lambda_t, g_t, d_t, \eta_t\right), i_{t+1}\right)\right] \\
& -\frac{1}{1+g_t-\delta}\left(\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right)\right] \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right)+\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right]\left(1-\Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right)\right).
\end{aligned}
$$

Proof of Lemma 7. Using Lemma 4,

$$
\begin{aligned}
\frac{\partial}{\partial d_t} \Psi\left(\lambda_t, g_t, d_t, \eta_t\right) = & \ \frac{\partial}{\partial d_t} \mathrm{E}_t\left[J_{t+1}^*\left(1, i_{t+1}\right) \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right)\right] \\
& +\frac{\partial}{\partial d_t} \mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right)-\sigma_Z \frac{1-\eta_t}{1+g_t-\delta} \varepsilon, i_{t+1}\right) \Phi'(\varepsilon) d\varepsilon\right] \\
& +\frac{\partial}{\partial d_t} \mathrm{E}_t\left[J_{t+1}^*\left(0, i_{t+1}\right)\left(1-\Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right)\right] \\
= & \ \mathrm{E}_t\left[J_{t+1}^*\left(1, i_{t+1}\right) \frac{\partial}{\partial d_t} \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right)\right] \\
& +\mathrm{E}_t\left[\frac{\partial}{\partial d_t} \int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right)-\sigma_Z \frac{1-\eta_t}{1+g_t-\delta} \varepsilon, i_{t+1}\right) \Phi'(\varepsilon) d\varepsilon\right] \\
& +\mathrm{E}_t\left[J_{t+1}^*\left(0, i_{t+1}\right) \frac{\partial}{\partial d_t}\left(1-\Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right)\right]
\end{aligned}
$$

where the derivative can be introduced inside the integral using the bounded convergence theorem. Using

Leibniz integral rule[11],

$$\frac{\partial}{\partial d_t} \Psi\left(\lambda_t, g_t, d_t, \eta_t\right)$$

$$= \mathrm{E}_t\left[J_{t+1}^*\left(1, i_{t+1}\right) \Phi'\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right) \frac{\partial A_t}{\partial d_t}\left(\lambda_t, d_t, \eta_t\right)\right]$$

$$+\mathrm{E}_t\left[\frac{\partial B_t}{\partial d_t}\left(\lambda_t, g_t, d_t, \eta_t\right) J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} B_t, i_{t+1}\right) \Phi'\left(B_t\right)\right]$$

$$-\mathrm{E}_t\left[\frac{\partial A_t}{\partial d_t}\left(\lambda_t, d_t, \eta_t\right) J_{t+1}^*\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} A_t, i_{t+1}\right) \Phi'\left(A_t\right)\right]$$

$$+\mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} \frac{\partial J_{t+1}^*}{\partial d_t}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$-\mathrm{E}_t\left[J_{t+1}^*\left(0, i_{t+1}\right) \Phi'\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right) \frac{\partial B_t}{\partial d_t}\left(\lambda_t, g_t, d_t, \eta_t\right)\right]$$

$$= \mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} \frac{\partial J_{t+1}^*}{\partial d_t}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} \frac{\partial \lambda_{t+1}}{\partial d_t}\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \quad \text{(chain rule)}$$

$$= \mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} \frac{1}{1+g_t-\delta}\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \quad \text{(Lemma 5)}$$

$$= \frac{1}{1+g_t-\delta}\mathrm{E}_t\left[\int_{A_t\left(\lambda_t, d_t, \eta_t\right)}^{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \frac{1}{1+g_t-\delta}\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}\left(\lambda_t, g_t, d_t, \eta_t\right), i_{t+1}\right)\right]$$

$$-\frac{1}{1+g_t-\delta}\mathrm{E}_t\left[\int_{-\infty}^{A_t\left(\lambda_t, d_t, \eta_t\right)} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} A_t\left(\lambda_t, d_t, \eta_t\right), i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$-\frac{1}{1+g_t-\delta}\mathrm{E}_t\left[\int_{B_t\left(\lambda_t, g_t, d_t, \eta_t\right)}^{\infty} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} B_t\left(\lambda_t, g_t, d_t, \eta_t\right), i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \frac{1}{1+g_t-\delta}\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}\left(\lambda_t, g_t, d_t, \eta_t\right), i_{t+1}\right)\right]$$

$$-\frac{1}{1+g_t-\delta}\left(\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right)\right] \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right) + \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right]\left(1 - \Phi\left(B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right)\right). \blacksquare$$

Lemma 8

$$\frac{\partial \Psi}{\partial g_t}\left(\lambda_t, g_t, d_t, \eta_t\right) = \frac{1}{1+g_t-\delta}\left(\mathrm{E}_t\left[\left(1-\lambda_{t+1}\right)\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}, i_{t+1}\right)\right] - \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right] \Phi'\left(-B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right).$$

---

[11] $\frac{d}{d\alpha}\int_{a(\alpha)}^{b(\alpha)} f\left(x, \alpha\right) dx = \frac{db}{d\alpha}\left(\alpha\right) f\left(b\left(\alpha\right), \alpha\right) - \frac{da}{d\alpha}\left(\alpha\right) f\left(a\left(\alpha\right), \alpha\right) + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\left(x, \alpha\right) dx$

Proof of Lemma 8. Using similar arguments as for Lemma 7's proof,[12]

$$\frac{\partial}{\partial g_t} \Psi\left(\lambda_t, g_t, d_t, \eta_t\right)$$

$$= \mathrm{E}_t\left[\int_{A_t(\lambda_t,d_t,\eta_t)}^{B_t(\lambda_t,g_t,d_t,\eta_t)} \frac{\partial J_{t+1}^*}{\partial g_t}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \mathrm{E}_t\left[\int_{A_t(\lambda_t,d_t,\eta_t)}^{B_t(\lambda_t,g_t,d_t,\eta_t)} \frac{\partial \lambda_{t+1}}{\partial g_t}\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \quad \text{(chain rule)}.$$

Using Lemma 5 and Definition (**??**),

$$\frac{\partial}{\partial g_t} \Psi\left(\lambda_t, g_t, d_t, \eta_t\right)$$

$$= \mathrm{E}_t\left[\int_{A_t(\lambda_t,d_t,\eta_t)}^{B_t(\lambda_t,g_t,d_t,\eta_t)} \frac{1 - y_{t+1}\left(\lambda_t, g_t, d_t\right) + \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon}{1+g_t-\delta} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1} - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta}\varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$+ \mathrm{E}_t\left[\int_{-\infty}^{A_t(\lambda_t,d_t,\eta_t)} \frac{1 - y_{t+1}\left(\lambda_t, g_t, d_t\right) + \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} A_t}{1+g_t-\delta} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1} - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} A_t, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$+ \mathrm{E}_t\left[\int_{B_t(\lambda_t,g_t,d_t,\eta_t)}^{\infty} \frac{1 - y_{t+1}\left(\lambda_t, g_t, d_t\right) + \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} B_t}{1+g_t-\delta} \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(y_{t+1} - \sigma_Z \frac{1-\eta_t}{1+g_t-\delta} B_t, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$- \mathrm{E}_t\left[\frac{1-1}{1+g_t-\delta}\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right) \Phi\left(A_t\left(\lambda_t, d_t, \eta_t\right)\right)\right] - \mathrm{E}_t\left[\frac{1-0}{1+g_t-\delta}\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right] \Phi\left(-B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)$$

$$= \frac{1}{1+g_t-\delta}\left(\mathrm{E}_t\left[\left(1-\lambda_{t+1}\right)\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}, i_{t+1}\right)\right] - \mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right] \Phi'\left(-B_t\left(\lambda_t, g_t, d_t, \eta_t\right)\right)\right). \blacksquare$$

D   Roots of Equation (25)

There is another root to Equation (25) provided that

$$\mathrm{E}_t\left[\left(\lambda_{t+1} - y_{t+1}\right)\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}, i_{t+1}\right)\right] = -y_{t+1}\mathrm{E}_t\left[\frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\right]\Phi\left(-B_t\right). \tag{29}$$

The right hand side is negative and nil if the truncation effect is negligible.
If λ_t is large enough so that the truncation effect is negligible, then $\Phi\left(-B_t\right) = 0$, and the Condition (29) becomes

$$\mathrm{Cov}_t\left[\lambda_{t+1} - y_{t+1}, \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}}\left(\lambda_{t+1}, i_{t+1}\right)\right] = 0.$$

Since these two random variables are strongly dependent, one can show that this covariance cannot be nil unless $\eta_t = 1$.

---

[12] $\frac{d}{d\alpha}\int_{a(\alpha)}^{b(\alpha)} f\left(x, \alpha\right) dx = \frac{db}{d\alpha}\left(\alpha\right) f\left(b\left(\alpha\right), \alpha\right) - \frac{da}{d\alpha}\left(\alpha\right) f\left(a\left(\alpha\right), \alpha\right) + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\left(x, \alpha\right) dx$

E Second derivative with respect to η

Lemma 9

$$\frac{\partial^2 \Psi}{\partial \eta_t^2} \left(\lambda_t, g_t, d_t, \eta_t = 1\right) = \left(\frac{\sigma_Z}{1 + g_t - \delta}\right)^2 \mathrm{E}_t \left[\frac{\partial^2 J_{t+1}^*}{\partial \lambda_{t+1}^2} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right), i_{t+1}\right)\right]$$

Proof of the lemma. Using the dominated convergence theorem to introduce the derivative operator inside the expectation,

$$\frac{\partial^2 \Psi}{\partial \eta_t^2} \left(\lambda_t, g_t, d_t, \eta_t\right)$$

$$= \frac{\partial}{\partial \eta_t} \mathrm{E}_t \left[\int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \mathrm{E}_t \left[\frac{\partial}{\partial \eta_t} \int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \mathrm{E}_t \left[\frac{\partial B_t}{\partial \eta_t} \left(\lambda_t, \eta_t\right) \frac{\sigma_Z}{1 + g_t - \delta} B_t \left(\lambda_t, \eta_t\right) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} B_t \left(\lambda_t, \eta_t\right), i_{t+1}\right) \Phi'\left(B_t \left(\lambda_t, \eta_t\right)\right)\right]$$

$$- \mathrm{E}_t \left[\frac{\partial A_t}{\partial \eta_t} \left(\lambda_t, \eta_t\right) \frac{\sigma_Z}{1 + g_t - \delta} A_t \left(\lambda_t, \eta_t\right) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} A_t \left(\lambda_t, \eta_t\right), i_{t+1}\right) \Phi'\left(A_t \left(\lambda_t, \eta_t\right)\right)\right]$$

$$+ \mathrm{E}_t \left[\int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} \frac{\partial}{\partial \eta_t} \frac{\sigma_Z}{1 + g_t - \delta} \varepsilon \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \quad \text{(Leibniz integral rule)}$$

$$= \mathrm{E}_t \left[\frac{\partial B_t}{\partial \eta_t} \left(\lambda_t, \eta_t\right) \frac{\sigma_Z}{1 + g_t - \delta} B_t \left(\lambda_t, \eta_t\right) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(0, i_{t+1}\right) \Phi'\left(B_t \left(\lambda_t, \eta_t\right)\right)\right]$$

$$- \mathrm{E}_t \left[\frac{\partial A_t}{\partial \eta_t} \left(\lambda_t, \eta_t\right) \frac{\sigma_Z}{1 + g_t - \delta} A_t \left(\lambda_t, \eta_t\right) \frac{\partial J_{t+1}^*}{\partial \lambda_{t+1}} \left(1, i_{t+1}\right) \Phi'\left(A_t \left(\lambda_t, \eta_t\right)\right)\right]$$

$$+ \mathrm{E}_t \left[\int_{A_t(\lambda_t, d_t, \eta_t)}^{B_t(\lambda_t, g_t, d_t, \eta_t)} \left(\frac{\sigma_Z}{1 + g_t - \delta} \varepsilon\right)^2 \frac{\partial^2 J_{t+1}^*}{\partial \lambda_{t+1}^2} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right) - \sigma_Z \frac{1 - \eta_t}{1 + g_t - \delta} \varepsilon, i_{t+1}\right) \Phi'\left(\varepsilon\right) d\varepsilon\right] \quad \text{(Chain rule)}$$

If $\eta_t = 1$, then

$$\frac{\partial^2 \Psi}{\partial \eta_t^2} \left(\lambda_t, g_t, d_t, \eta_t = 1\right)$$

$$= \left(\frac{\sigma_Z}{1 + g_t - \delta}\right)^2 \mathrm{E}_t \left[\frac{\partial^2 J_{t+1}^*}{\partial \lambda_{t+1}^2} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right), i_{t+1}\right) \int_{-\infty}^{\infty} \varepsilon^2 \Phi'\left(\varepsilon\right) d\varepsilon\right]$$

$$= \left(\frac{\sigma_Z}{1 + g_t - \delta}\right)^2 \mathrm{E}_t \left[\frac{\partial^2 J_{t+1}^*}{\partial \lambda_{t+1}^2} \left(y_{t+1}\left(\lambda_t, g_t, d_t\right), i_{t+1}\right)\right]$$

F Limit cases as $\lambda \to 1$

Lemma 10 *If $\lim_{\lambda \to 1} w\left(\lambda\right) = \infty$ and $\lim_{\lambda \to 1} \frac{w'(\lambda)}{(w(\lambda))^2} = 0$, then for any t,*

$$\lim_{\lambda \to 1} J_t\left(\lambda\right) = 0 \ \text{ and } \ \lim_{\lambda \to 1} \frac{\partial J_t}{\partial \lambda_t}\left(\lambda\right) = 0.$$

Proof of Lemma ??. The final results are obtained noticing that

$$\lim_{\lambda_{T+1}\to 1} J_{T+1}\left(\lambda_{T+1}\right) = \lim_{\lambda_{T+1}\to 1} \frac{\mathrm{E}[z] - g + \delta}{w(\lambda_{T+1}) - g + \delta} = 0$$

$$\lim_{\lambda_{T+1}\to 1} \frac{\partial J_{T+1}}{\partial \lambda_{T+1}}\left(\lambda_{T+1}\right) = \lim_{\lambda_{T+1}\to 1} -w'(\lambda_{T+1})\frac{\mathrm{E}[z] - g + \delta}{(w(\lambda_{T+1}) - g + \delta)^2} = 0.$$

Using induction, since $\lim_{\lambda\to 1}\lambda^*_{t+1}\left(\lambda\right) = 1$ and $\lim_{\lambda\to 1} D\left(\lambda\right) = \lim_{\lambda\to 1}\frac{1}{1+w(\lambda)} = 0$,

$$\lim_{\lambda\to 1} J^*_t\left(\lambda\right) = \lim_{\lambda\to 1}\Theta_t\left(g^*_t, \eta^*_t, d^*_t; \lambda\right)$$
$$= \lim_{\lambda\to 1} D\left(\lambda\right)\left(\mathrm{E}[z] - g^*_t\left(\lambda\right) + \delta + (1 + g^*_t\left(\lambda\right) - \delta)\mathrm{E}_t\left[J^*_{t+1}\left(\lambda^*_{t+1}\left(\lambda\right), i_{t+1}\right)\right]\right) = 0.$$

For $\lambda_t > \lambda^{(3)}_t$, $g^*_t\left(\lambda\right) = 0$ and $d^*_t\left(\lambda\right) = 0$. Therefore,

$$\frac{\partial J^*_t}{\partial \lambda_t}\left(\lambda\right) = \frac{D'\left(\lambda\right)}{D\left(\lambda\right)} J^*_t\left(\lambda\right) + D\left(\lambda\right)(1 - \delta)\mathrm{E}_t\left[\frac{\partial J^*_{t+1}}{\partial \lambda_t}\left(\lambda^*_{t+1}\left(\lambda\right), i_{t+1}\right)\right].$$

But both $\frac{D'(\lambda)}{D(\lambda)} = -\frac{w'(\lambda)}{1+w(\lambda)}$ and $D\left(\lambda\right)$ converges to 0 as $\lambda \to 1$. Consequently, $\frac{\partial J^*_t}{\partial \lambda_t}\left(\lambda\right) \to 0$ as $\lambda \to 1$

$$\frac{\partial J^*_t}{\partial \lambda_t}\left(\lambda_t, i_t\right) = D'\left(\lambda_t\right)\left(\mathrm{E}[z] + \delta + (1 - \delta)\mathrm{E}_t\left[J^*_{t+1}\left(\lambda^*_{t+1}\left(\lambda_t\right), i_{t+1}\right)\right]\right) + D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(\lambda^*_{t+1}\left(\lambda_t, i_t\right), i_{t+1}\right)\right]$$

$$-D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right)\Phi\left(A^*\left(\lambda_t\right)\right)\right] - D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\Phi'\left(-B^*\left(\lambda_t, i_t\right)\right)\right].$$

Consequently

$$\lim_{\lambda_t\to 1}\frac{\partial J^*_t}{\partial \lambda_t}\left(\lambda_t, i_t\right)$$
$$= \lim_{\lambda_t\to 1} D'\left(\lambda_t\right)\left(\mathrm{E}[z] + \delta + (1 - \delta)\mathrm{E}_t\left[J^*_{t+1}\left(\lambda^*_{t+1}\left(\lambda_t\right), i_{t+1}\right)\right]\right) + \lim_{\lambda_t\to 1} D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(\lambda^*_{t+1}\left(\lambda_t, i_t\right), i_{t+1}\right)\right]$$
$$- \lim_{\lambda_t\to 1} D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(1, i_{t+1}\right)\Phi\left(A^*\left(\lambda_t\right)\right)\right] - \lim_{\lambda_t\to 1} D\left(\lambda_t\right)\mathrm{E}_t\left[\mu'\left(\lambda_t\right)\frac{\partial J^*_{t+1}}{\partial \lambda_{t+1}}\left(0, i_{t+1}\right)\Phi'\left(-B^*\left(\lambda_t, i_t\right)\right)\right]$$
$$= 0.$$

G   Induction hypotheses at time T

G.1   At time $T+1$

Since

$$J_{T+1}(\lambda) = \frac{\mathrm{E}[z] - g + \delta}{w(\lambda) - g + \delta},$$

then J_{T+1} is maximized at $\overline{\lambda}_{T+1} = \lambda^*$. Because $\mathrm{E}[z] > w(\lambda^*)$, $J_{T+1}(\lambda^*) > 1$, establishing Induction Hypothesis 2.

By definition, $\widehat{\lambda}_{T+1}$ satisfies $\phi_{T+1}\left(\widehat{\lambda}_{T+1}\right) = 0$ where

$$\phi_{T+1}(\lambda) = J_{T+1}(\lambda) - 1 + (1-\lambda) J'_{T+1}(\lambda) = \frac{\mathrm{E}[z] - g + \delta}{w(\lambda) - g + \delta} - 1 - (1-\lambda) \frac{\mathrm{E}[z] - g + \delta}{(w(\lambda) - g + \delta)^2} w'(\lambda).$$

Because ϕ_{T+1} is a continuous and decreasing function on $[\lambda^*, 1]$ with $\phi_{T+1}(\lambda^*) = J_{T+1}(\lambda) - 1 > 0$ and $\lim_{\lambda \to 1} \phi_{T+1}(\lambda) = -1$, then $\lambda^* \le \widehat{\lambda}_{T+1} < 1$. Moreover, since J_{T+1} is decreasing on $[\lambda^*, 1]$, then $J'_{T+1}(\lambda) < 0$ for all $\lambda \in [\lambda^*, 1]$. In particular, $J'_{T+1}\left(\widehat{\lambda}_{T+1}\right) < 0$. Finally, since $\phi_{T+1}\left(\widehat{\lambda}_{T+1}\right) = 0$, then

$$0 = \phi_{T+1}\left(\widehat{\lambda}_{T+1}\right) = J_{T+1}\left(\widehat{\lambda}_{T+1}\right) - 1 + \left(1 - \widehat{\lambda}_{T+1}\right) J'_{T+1}\left(\widehat{\lambda}_{T+1}\right) \le J_{T+1}\left(\widehat{\lambda}_{T+1}\right) - 1,$$

implying that $J_{T+1}\left(\widehat{\lambda}_{T+1}\right) \ge 1$.

G.2   At time T

Because $\varphi_{T+1}(y) = J_{T+1}(y)$, Lemma 1 implies that full hedging is optimal whenever $y_{T+1}(\lambda_T) \in [a_{T+1}, b_{T+1}]$.

Lemma 11 *For any* $\lambda_T \in \left[0, \lambda_T^{(4)}\right]$, $y_{T+1}^*(\lambda_T) \in [a_{T+1}, b_{T+1}]$.

Proof. First, it is shown that $\lambda_{T+1}^*(\lambda_T) = \lambda^*$ for all $\lambda_T \le \lambda_T^{(1)}$. Indeed, Equation (12) is

$$\Theta_T(g_T, \eta_T, d_T; \lambda_T) = D(\lambda_T) \left(\mathrm{E}[z] - 1 + (1 + g_T - \delta)\mathrm{E}_T \left[\frac{\mathrm{E}[z] - g + \delta}{w(\lambda_{T+1}) - g + \delta} - 1 \right] \right).$$

and is maximized if $\lambda_{T+1} = \lambda^*$ and g_T is a large as possible. Therefore, for small values of λ_T, $g_T^*(\lambda_T) = i_T + \delta$ and the optimal dividend rate is such that $y_{T+1}(\lambda_T, g_T^*, d_T^*) = \lambda^*$, that is, $d_T^*(\lambda_T) = \mathrm{E}[z] + \lambda^* - (1 - \lambda^*) i_T - \mu(\lambda_T)$. This situation is feasible if and only if $d_T^*(\lambda_T)$ is positive, implying that $\lambda_T \in \left[0, \lambda_T^{(1)}(i_T)\right]$. For $\lambda_T > \lambda_T^{(1)}(i_T)$, the firm is already to endebted to pay dividends and $y_{T+1}^*(\lambda_T)$ is an increasing function of λ_T. Therefore, full hedging is the optimal decision until $y_{T+1}^*(\lambda_T)$ reaches b_{T+1}, that is, for all $\lambda_T < \lambda_T^{(4)}$. ∎

Because $\eta_T^*(\lambda_T) = 1$ for $\lambda_T \in \left[0, \lambda_T^{(4)}\right]$, the proof of Section A.3 holds, thus Proposition 1 is verified at time T. Consequently, replacing the optimal controls into Equation (12) leads to the relative firm value at

time $T: J_T^*\left(\lambda, i_T\right) =$

$$
\begin{cases}
D\left(\lambda\right)\left(\mathrm{E}[z] - i_T + (1+i_T)J_{T+1}\left(\lambda^*\right)\right) & \text{if } \lambda \in \left[0, \lambda_T^{(1)}\left(i_T\right)\right] \\[2mm]
D\left(\lambda\right)\left(\mathrm{E}[z] - i_T + (1+i_T)J_{T+1}\left(\frac{\mu(\lambda)+i_T-\mathrm{E}[z]}{1+i_T}\right)\right) & \text{if } \lambda \in \left(\lambda_T^{(1)}\left(i_T\right), \lambda_T^{(2)}\left(i_T\right)\right] \\[2mm]
D\left(\lambda\right)\left(\mathrm{E}[z] + \left(J_{T+1}\left(\widehat{\lambda}_{T+1}\right) - 1\right)\frac{\mathrm{E}[z]+\widehat{\lambda}_{T+1}-\mu(\lambda)}{1-\widehat{\lambda}_{T+1}} + J_{T+1}\left(\widehat{\lambda}_{T+1}\right)\right) & \text{if } \lambda \in \left(\lambda_T^{(2)}\left(i_T\right), \lambda_T^{(3)}\right] \\[2mm]
D\left(\lambda\right)\left(\mathrm{E}[z] + \delta + (1-\delta)J_{T+1}\left(\frac{\mu(\lambda)-\delta-\mathrm{E}[z]}{1-\delta}\right)\right) & \text{if } \lambda \in \left(\lambda_T^{(3)}, \lambda_T^{(4)}\right]
\end{cases}
$$

$$(30)$$

G.3 Verification of Induction Hypothesis 1 at time T

The concavity is verified by looking at the second derivative of J_T^* which must be negative. Starting from Equation 30, the second derivative is computed for each of the four intervals.

For $\lambda_T \in \left[0, \lambda_T^{(1)}\left(i_T\right)\right]$,

$$
\frac{\partial^2 J_t^*}{\partial \lambda_t^2}\left(\lambda_T, i_T\right) = \frac{D''\left(\lambda_T\right)}{D\left(\lambda_T\right)} J_T^*\left(\lambda_T, i_T\right) < 0
$$

because $D''\left(\lambda_T\right) < 0$.

For $\lambda_T \in \left(\lambda_T^{(1)}\left(i_T\right), \lambda_T^{(2)}\left(i_T\right)\right]$,

$$
\begin{aligned}
\frac{\partial^2 J_T^*}{\partial \lambda_t^2}\left(\lambda_T, i_T\right) &= \frac{D''\left(\lambda_T\right)}{D\left(\lambda_T\right)} J_T^*\left(\lambda_T, i_T\right) + \left(D\left(\lambda_T\right)\mu\left(\lambda_T\right)'' - D''\left(\lambda_T\right)\mu\left(\lambda_T\right)\right) J_{T+1}'\left(\frac{\mu\left(\lambda_T\right)+i_T-\mathrm{E}\left[z\right]}{1+i_T}\right) \\
&\quad + D\left(\lambda_T\right)\frac{\left(\mu'\left(\lambda_T\right)\right)^2}{1+i_T} J_{T+1}''\left(\frac{\mu\left(\lambda_T\right)+i_T-\mathrm{E}\left[z\right]}{1+i_T}\right).
\end{aligned}
$$

The first term is negative since D and J_T^* are positive and D'' is negative. The second term is negative because $D\mu'' - D''\mu > 0$ by assumption and $\lambda_T > \lambda_T^{(1)}\left(i_T\right)$ implies that $\frac{\mu(\lambda_T)+i_T-\mathrm{E}[z]}{1+i_T} > \frac{\mu\left(\lambda_T^{(1)}(i_T)\right)+i_T-\mathrm{E}[z]}{1+i_T} = \lambda^*$. Therefore, J_{T+1} being decreasing at the right of λ^*, $J_{T+1}' < 0$. Finally, the third term is negative if J_{T+1} is concave on

$$
\left]\lambda_{T+1}^*\left(\lambda_T^{(1)}\left(i_T\right)\right), \lambda_{T+1}^*\left(\lambda_T^{(2)}\left(i_T\right)\right)\right] = \left]\lambda^*, \widehat{\lambda}_{T+1}\right].
$$

For $\lambda_T \in \left(\lambda_T^{(2)}\left(i_T\right), \lambda_T^{(3)}\right]$,

$$
\frac{\partial^2 J_T^*}{\partial \lambda_t^2}\left(\lambda_T, i_T\right) = \frac{D''\left(\lambda_T\right)}{D\left(\lambda_T\right)} J_T^*\left(\lambda_T, i_T\right) - \frac{J_{T+1}\left(\widehat{\lambda}_{T+1}\right) - 1}{1 - \widehat{\lambda}_{T+1}}\left(D\left(\lambda_T\right)\mu\left(\lambda_T\right)'' - D''\left(\lambda_T\right)\mu\left(\lambda_T\right)\right).
$$

54

Again, the first term in negative for the same reasons as before. Because $J_{T+1}\left(\widehat{\lambda}_{T+1}\right) > 1$ (see Appendix G.1) and $D\mu'' - D''\mu > 0$ by assumption, the second term is also negative.

Finally, for $\lambda_T \in \left(\lambda_T^{(3)}, 1\right]$,

$$
\begin{aligned}
\frac{\partial^2 J_T^*}{\partial \lambda_t^2}\left(\lambda_T, i_T\right) =\ & \frac{D''\left(\lambda_T\right)}{D\left(\lambda_T\right)} J_T^*\left(\lambda_T, i_T\right) + \left(D\left(\lambda_T\right) \mu\left(\lambda_T\right)'' - D''\left(\lambda_T\right) \mu\left(\lambda_T\right)\right) J_{T+1}'\left(\frac{\mu\left(\lambda_T\right) - \delta - \mathrm{E}\left[z\right]}{1 - \delta}\right) \\
& + D\left(\lambda_T\right) \frac{\left(\mu'\left(\lambda_T\right)\right)^2}{1 - \delta} J_{T+1}''\left(\frac{\mu\left(\lambda_T\right) - \delta - \mathrm{E}\left[z\right]}{1 - \delta}\right).
\end{aligned}
$$

The two first terms are negative. Since J_{T+1}'' remains negative until $\frac{\mu(\lambda_T)-\delta-\mathrm{E}[z]}{1-\delta}$ reaches c_T, the third term is negative provided $\frac{\mu(\lambda_T)-\delta-\mathrm{E}[z]}{1-\delta} \leq c_T$. Therefore, there is a constant $b_T > \lambda_T^{(3)}$ for which J_T^* is concave on $[0, b_T]$. ∎

G.4 Verification of the Induction Hypothesis 2 at time T

We will show that on $[0, \lambda_t^{(3)}]$, $D^{-1}\left(\lambda_t\right) J_t^*\left(\lambda_t, i_t\right) \geq 1 + \mathrm{E}[z]$. As a consequence, since $\lambda^* \in [0, \lambda_t^{(3)}]$, then $J_t^*\left(\lambda^*, i_t\right) \geq \frac{1+\mathrm{E}[z]}{1+w(\lambda^*)} > 1$. The proof of Section A.6 is also valid for time T since in Appendix G.1, we have shown that $J_{T+1}\left(\overline{\lambda}_{T+1}\right) > 1$ and $J_{T+1}(\widehat{\lambda}_{T+1}) > 1$.

H   Numerical Implementation for Multi-period Model

To implement the dynamic programming model $(11b)$, one has to rely on a suboptimal approximation. Consider a finite partition $[\lambda_l, \lambda_{l+1})$ of the λ dimension such that $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_N = 1$. Given the discrete nature of the variable i, the state space takes values on the grid:

$$\mathcal{G} = \left\{ (\lambda_l, i_m) \mid l = 0, \ldots, N \text{ and } m = 1, \ldots, |M| \right\},$$

where $|\cdot|$ denotes the number of elements in the set.

Since the state space has been approximated by the grid $\mathcal{G}$, $\tilde{J}_{t+1}$ denotes the approximation of J_{t+1} that is available at each point of $\mathcal{G}$. This means that the suboptimal dynamic programming problem $(11b)$ to solve for $t = 1, \ldots, T-1$ is the following:

$$\tilde{J}_t \left(\lambda^l, i^m \right) = \max_{\substack{g_t, \eta_t \\ 0 \leq g_t \leq i_t}} \left\{ \frac{1}{1 + w(\lambda_t)} \left(\mathrm{E}[z] - g_t + \delta + (1 + g_t - \delta) \, \mathrm{E}_{lmt} \left[\tilde{J}_{t+1} \left(\lambda_{t+1}, i_{t+1} \right) \right] \right) \right\},$$
$$(31)$$

where $\mathrm{E}_{lmt}[\cdot]$ stands for the conditional expectation at time t given that the state variables where at $\left(\lambda^l, i^m \right)$.

Given that the state space is continuous on the λ dimension, and in order to compute the recursion step in the Bellman equation, the cost-to-go function $\tilde{J}_{t+1}(\lambda_{t+1}, i_{t+1})$ has to be approximated with some function so that the problem can be implemented numerically for $t = 1, ..., T - 1$.

The function $\tilde{J}_{t+1}$ is approximated with a piecewise quadratic interpolation on the λ dimension, which will be denoted by $\hat{J}_{t+1}$.[13] Other types of interpolation schemes can be used, such as higher order polynomials, spline functions, or Laguerre approximations; however, in the present case, these approximations will not contribute much to the precision of the results but will render the implementation more complicated.

To solve (31), one needs to approximate the term $\mathrm{E}_{lmt}[\cdot]$. Suppose that at time t the current state is $(\lambda_l, i_m) \in \mathcal{G}$. Now, consider the following regions that partition the space state:

$$\mathcal{A}_j^k = \left\{ (\lambda, i) \mid \lambda_j \leq \lambda \leq \lambda_{j+1} \text{ and } i = i_k \right\}, \ j = 0, \ldots, N - 1, k = 1, \ldots, |M|.$$

The value of $\tilde{J}_{t+1}(\lambda_{t+1}, i_{t+1})$ is approximated within each region by:

$$\hat{J}_{j,k,t+1}(\lambda_{t+1}, i_k) = \psi_{j,k,t} + \theta_{j,k,t}(\lambda_{t+1} - \lambda_j) + \gamma_{j,k,t}(\lambda_{t+1} - \lambda_j)(\lambda_{t+1} - \lambda_{j+1}), \quad (32)$$

where $\psi_{j,k,t}, \theta_{j,k,t}$, and $\gamma_{j,k,t}$ are defined in $(38), (39)$, and (40) respectively. Then, one can write the expected value of $\tilde{J}_{t+1}$ in terms of the regions $\mathcal{A}_j^k$ as:

$$\mathrm{E}_{lmt} \left[\tilde{J}_{t+1}(\lambda_{t+1}, i_{t+1}) \right] \simeq \sum_{j=0}^{N-1} \sum_{k=0}^{|M|} \mathrm{E}_{lmt} \left[\hat{J}_{j,k,t+1}(\lambda_{t+1}, i_{t+1}) \, I \left\{ \mathcal{A}_j^k \right\} \right], \quad (33)$$

---

[13]This kind of approximation has been use in the context of option pricing with a Brownian process in [2] and for GARCH process in [1].

where $I\{A_j^k\}$ is the indicator function for the event $(\lambda_{t+1}, i_{t+1}) \in A_j^k$.

Working with $\mathrm{E}_{lmt}\left[\hat{J}_{j,k,t+1}\left(\lambda_{t+1}, i_{t+1}\right) I\left\{A_j^k\right\}\right]$ for $j \in \{0, \ldots, N-2\}$, $k \in \{0, \ldots, |M|\}$ and using (32), one has:

$$
\begin{aligned}
\mathrm{E}_{lmt}\left[\hat{J}_{j,k,t+1}\left(\lambda_{t+1}, i_{t+1}\right) I\left\{A_j^k\right\}\right] &= \left(\psi_{j,k,t} - \theta_{j,k,t}\lambda_j + \gamma_{j,k,t}\lambda_j\lambda_{j+1}\right)\mathrm{E}_{lmt}\left[I\left\{A_j^k\right\}\right] + \\
&\quad \left(\theta_{j,k,t} - \gamma_{j,k,t}\left(\lambda_j + \lambda_{j+1}\right)\right)\mathrm{E}_{lmt}\left[\lambda_{t+1}I\left\{A_j^k\right\}\right] \\
&\quad + \gamma_{j,k,t}\mathrm{E}_{lmt}\left[\left(\lambda_{t+1}\right)^2 I\left\{A_j^k\right\}\right].
\end{aligned}
\tag{34}
$$

Combining $(31), (33)$, and(34) gives:

$$
\begin{aligned}
\tilde{J}_t\left(\lambda^l, i^m\right) &\simeq \max_{\substack{g_t, \eta_t \\ st:0 \le g_t \le i_t}}\left\{\frac{1}{1+w\left(\lambda_t\right)}\left(\mathrm{E}\left[z_t\right] - g_t + (1+g_t)\sum_{j=0}^{N-1}\sum_{k=0}^{|M|}\left(\tilde{\psi}_{j,k,t}\mathrm{E}_{lmt}\left[I\left\{A_j^k\right\}\right]\right.\right.\right. \\
&\quad \left.\left.\left. + \tilde{\theta}_{j,k,t}\mathrm{E}_{lmt}\left[\lambda_{t+1}I\left\{A_j^k\right\}\right] + \gamma_{j,k,t}\mathrm{E}_{lmt}\left[\left(\lambda_{t+1}\right)^2 I\left\{A_j^k\right\}\right]\right)\right)\right\},
\end{aligned}
\tag{35}
$$

where $\tilde{\psi}_{j,k,t}$ and $\tilde{\theta}_{j,k,t}$ are defined in as

$$
\begin{aligned}
\tilde{\psi}_{j,k,t} &= \psi_{j,k,t} - \theta_{j,k,t}\lambda_j + \gamma_{j,k,t}\lambda_j\lambda_{j+1}, & (36) \\
\tilde{\theta}_{j,k,t} &= \theta_{j,k,t} - \gamma_{j,k,t}\left(\lambda_j + \lambda_{j+1}\right), & (37) \\
\psi_{j,k,t} &= \tilde{J}_{t+1}\left(\lambda_j, i_k\right), & (38) \\
\theta_{j,k,t} &= \frac{\tilde{J}_{t+1}\left(\lambda_{j+1}, i_k\right) - \tilde{J}_{t+1}\left(\lambda_j, i_k\right)}{\lambda_{j+1} - \lambda_j}, & (39) \\
\gamma_{j,k,t} &= \begin{cases} \frac{\left(\tilde{J}_{t+1}(\lambda_j, i_k)(\lambda_{j+2}-\lambda_{j+1}) - \tilde{J}_{t+1}(\lambda_{j+1}, i_k)(\lambda_{j+2}-\lambda_j) + \tilde{J}_{t+1}(\lambda_{j+2}, i_k)(\lambda_{j+1}-\lambda_j)\right)}{(\lambda_{j+2}-\lambda_{j+1})(\lambda_{j+2}-\lambda_j)(\lambda_{j+1}-\lambda_j)} & j < N-1 \\ 0 & j = N-1 \end{cases} & (40)
\end{aligned}
$$

respectively.

Given that the underlying source of volatility of the profitability is assumed Gaussian, closed-form expressions for $\mathrm{E}_{lmt}\left[I\left\{A_j^k\right\}\right]$, $\mathrm{E}_{lmt}\left[\lambda_{t+1}I\left\{A_j^k\right\}\right]$, and $\mathrm{E}_{lmt}\left[\left(\lambda_{t+1}\right)^2 I\left\{A_j^k\right\}\right]$ can be computed.

References

[1] Ben-Ameur, H., M. Breton and J.M. Martinez, 2009, Dynamic Programming Approach for Valuing Options in the GARCH Model, Management Science 55, 252-266.

[2] Ben-Ameur H. , M. Breton, and P. L'Ecuyer, 2002, A Dynamic Programming Procedure for Pricing American-Style Asian Options, Management Science 48, 625-643.

H.1 Expressions for expected values

This appendix presents closed-form expressions for the expected values $E_{lmt}\left[I\left\{\mathcal{A}_j^k\right\}\right]$, $E_{lmt}\left[\lambda_{t+1}I\left\{\mathcal{A}_j^k\right\}\right]$, and $E_{lmt}\left[(\lambda_{t+1})^2 I\left\{\mathcal{A}_j^k\right\}\right]$.

First, equation (5) implies that λ_{t+1} is a truncated normal variable over the interval $[0,1]$ with expected value and variance conditional on t given by

$$E\left[\lambda_{t+1}|\lambda_t, g_t, \eta_t\right] = \mu_{\Lambda_{t+1}} - \sigma_{\Lambda_{t+1}}\frac{\phi\left(\bar{\lambda}_{t+1}^1\right) - \phi\left(\bar{\lambda}_{t+1}^0\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)}, \tag{41}$$

$$\mathbf{Var}\left[\lambda_{t+1}|\lambda_t, g_t, \eta_t\right] = \sigma_{\Lambda_{t+1}}^2\left(1 - \frac{\bar{\lambda}_{t+1}^1\phi\left(\bar{\lambda}_{t+1}^1\right) - \bar{\lambda}_{t+1}^0\phi\left(\bar{\lambda}_{t+1}^0\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} - \left(\frac{\phi\left(\bar{\lambda}_{t+1}^1\right) - \phi\left(\bar{\lambda}_{t+1}^0\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)}\right)^2\right) \tag{42}$$

where

$$\mu_{\Lambda_{t+1}} = \frac{g_t - \delta - \mu_z + \lambda_t\left(1 + r\left(\lambda_t\right)\right) + d_t}{1 + g_t - \delta},$$

$$\sigma_{\Lambda_{t+1}}^2 = \left(\frac{1 - \eta_t}{1 + g_t - \delta}\right)^2 \sigma_z^2,$$

$\bar{\lambda}_{t+1}^1 = \frac{1 - \mu_{\Lambda_{t+1}}}{\sigma_{\Lambda_{t+1}}}$, $\bar{\lambda}_{t+1}^0 = \frac{-\mu_{\Lambda_{t+1}}}{\sigma_{\Lambda_{t+1}}}$, and ϕ and Φ are the density and cumulative distribution of the normal variable.

Given the independence between z and i, the probability of being in a certain region is equivalent to the multiplication of being in the region defined on the dimension of λ, multiplied by the probability of being in the state $i_{t+1} = i^k$, which is P_k. Consequently, one has:

$$E_{lmt}\left[I\left\{\mathcal{A}_j^k\right\}\right] = P\left(\lambda_j \leq \lambda_{t+1} \leq \lambda_{j+1}\right)P_k = \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)}P_k,$$

with $\tilde{\lambda}^j = \frac{\lambda_j - \mu_{\lambda_{t+1}}}{\sigma_{\lambda_{t+1}}}$, $\bar{\lambda}_{t+1}^1 = \frac{1 - \mu_{\lambda_{t+1}}}{\sigma_{\lambda_{t+1}}}$, and $\bar{\lambda}_{t+1}^0 = \frac{-\mu_{\lambda_{t+1}}}{\sigma_{\lambda_{t+1}}}$.

To compute the terms $E_{lmt}\left[\lambda_{t+1}I\left\{\mathcal{A}_j^k\right\}\right]$ and $E_{lmt}\left[(\lambda_{t+1})^2 I\left\{\mathcal{A}_j^k\right\}\right]$, the moment generating function of the truncated normal is used:

$$M\left(t\right) = E\left[e^{Xt}\big| X \in [\alpha_1, \alpha_2]\right] = e^{\mu t + \frac{\sigma^2 t^2}{2}}\frac{\Phi\left(\frac{\alpha_2 - \mu}{\sigma} - \sigma t\right) - \Phi\left(\frac{\alpha_1 - \mu}{\sigma} - \sigma t\right)}{\Phi\left(\frac{\alpha_2 - \mu}{\sigma}\right) - \Phi\left(\frac{\alpha_1 - \mu}{\sigma}\right)},$$

where μ and σ are the expected value and standard deviation of the normal distribution.

Expressions for $\mathrm{E}_{lmt}\left[\lambda_{t+1}I\left\{\mathcal{A}_j^k\right\}\right]$ with $j \in \{0,\ldots,N-1\}$, can be computed as follows:

$$
\begin{aligned}
\mathrm{E}_{lmt}\left[\lambda_{t+1}I\left\{\mathcal{A}_j^k\right\}\right] &= P_k \int_{\lambda_j}^{\lambda_{j+1}} \lambda_{t+1} \frac{\frac{1}{\sigma_{\lambda_{t+1}}}\varphi\left(\frac{\lambda_{t+1}-\mu_{\lambda_{t+1}}}{\sigma_{\lambda_{t+1}}}\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} d\lambda_{t+1} \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} \int_{\lambda_j}^{\lambda_{j+1}} \lambda_{t+1} \frac{\frac{1}{\sigma_{\lambda_{t+1}}}\varphi\left(\frac{\lambda_{t+1}-\mu_{\lambda_{t+1}}}{\sigma_{\lambda_{t+1}}}\right)}{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)} d\lambda_{t+1} \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} \mathrm{E}\left[\lambda_{t+1}|\,\lambda_{t+1} \in [\lambda_j, \lambda_{j+1}]\right] \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} \left.M'(t)\right|_{t=0} \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi\left(\bar{\lambda}_{t+1}^1\right) - \Phi\left(\bar{\lambda}_{t+1}^0\right)} \left(\mu_{\lambda_{t+1}} - \sigma_{\lambda_{t+1}} \frac{\varphi\left(\tilde{\lambda}_{j+1}\right) - \varphi\left(\tilde{\lambda}_j\right)}{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}\right).
\end{aligned}
$$

In a similar way, for $j \in \{0,\ldots,N-2\}$ one has:

$$
\begin{aligned}
&\mathrm{E}_{lmt}\left[(\lambda_{t+1})^2 I\left\{\mathcal{A}_j^k\right\}\right] \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi_{t+1}\left(\tilde{\lambda}\right) - \Phi_{t+1}\left(\tilde{\lambda}_{t+1}\right)} \left.M''(t)\right|_{t=0} \\
&= P_k \frac{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)}{\Phi_{t+1}\left(\tilde{\lambda}\right) - \Phi_{t+1}\left(\tilde{\lambda}_{t+1}\right)} \left(\sigma_{\lambda_{t+1}}^2 + \mu_{\lambda_{t+1}}^2 - \sigma_{\lambda_{t+1}}^2 \frac{\tilde{\lambda}_{j+1}\varphi\left(\tilde{\lambda}_{j+1}\right) - \tilde{\lambda}_j\varphi\left(\tilde{\lambda}_j\right)}{\Phi\left(\tilde{\lambda}_{j+1}\right) - \Phi\left(\tilde{\lambda}_j\right)} - 2\mu_{\lambda_{t+1}}\sigma_{\lambda_{t+1}} \frac{\varphi\left(\tilde{\lambda}_{j+1}\right) -}{\Phi\left(\tilde{\lambda}_{j+1}\right) -}\right.
\end{aligned}
$$

Figure 5: WACC curve
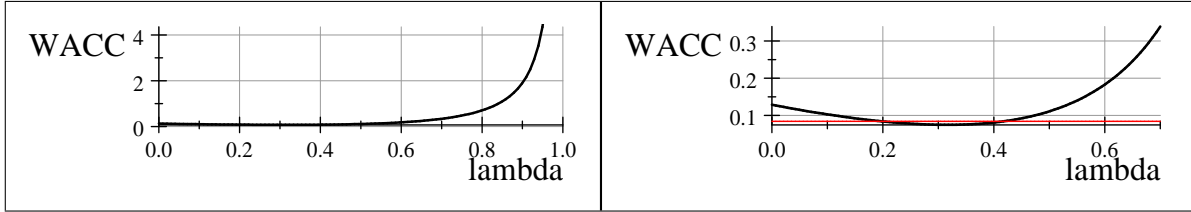
The **WACC curve** is

$$w(\lambda) = \frac{\beta_0 + \beta_1\lambda + \beta_2\lambda^2}{1 - \lambda}$$

with $\beta_0 = 0.1287$, $\beta_1 = -0.4159$ and $\beta_2 = 0.5389$.
With this specification,

$$w'(\lambda) = \frac{\beta_0 + \beta_1 + 2\beta_2\lambda - \beta_2\lambda^2}{(1 - \lambda)^2} \text{ and } w''(\lambda) = 2\frac{\beta_0 + \beta_1 + \beta_2}{(1 - \lambda)^3}.$$

Moreover, $\lambda^* = 0.31658$ since it is the solution to $w'(\lambda^*) = \frac{\beta_0 + \beta_1 + 2\beta_2\lambda - \beta_2\lambda^2}{(1-\lambda)^2} = 0$. The red line correspond to $\mathrm{E}[z]$.
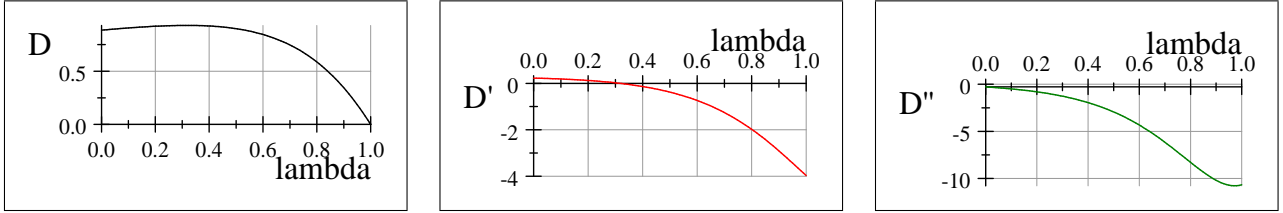
I   Numerical example

I.1   Notation and assumptions

I.2   The firm's relative value at time $T + 1$

Parameter estimates. The following parameters have been estimated:

$$\begin{aligned}
\mathrm{E}[z] &= 0.084 \\
\sigma &= 0.058 \\
\delta &= 0.120 \\
i &= I = 0.126 + 0.021 \\
g &= \delta + 0.021 \\
q &= g - \delta
\end{aligned}$$
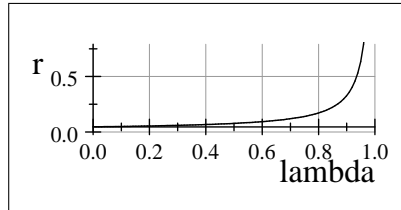
Figure 6: Discount factor

The **discount factor** is

$$D\left(\lambda\right) = \frac{1}{1 + w\left(\lambda\right)} = \frac{1 - \lambda}{\left(1 + \beta_0\right) - \left(1 - \beta_1\right)\lambda + \beta_2\lambda^2}$$

The concavity is verified by the mean of its second derivative which should be negative:

$$D'\left(\lambda\right) = \frac{-\left(\beta_0 + \beta_1\right) - 2\beta_2\lambda + \beta_2\lambda^2}{\left(\left(1 + \beta_0\right) - \left(1 - \beta_1\right)\lambda + \beta_2\lambda^2\right)^2}$$

$$D''\left(\lambda\right) = \frac{-2\left(\left(\beta_0 + \beta_1 + \beta_2\right) + \beta_0\left(\beta_2 - \beta_1\right) - \beta_1^2\right) + 6\beta_2\left(\beta_0 + \beta_1\right)\lambda + 6\beta_2^2\lambda^2 - 2\beta_2^2\lambda^3}{\left(\left(1 + \beta_0\right) - \left(1 - \beta_1\right)\lambda + \beta_2\lambda^2\right)^3}$$
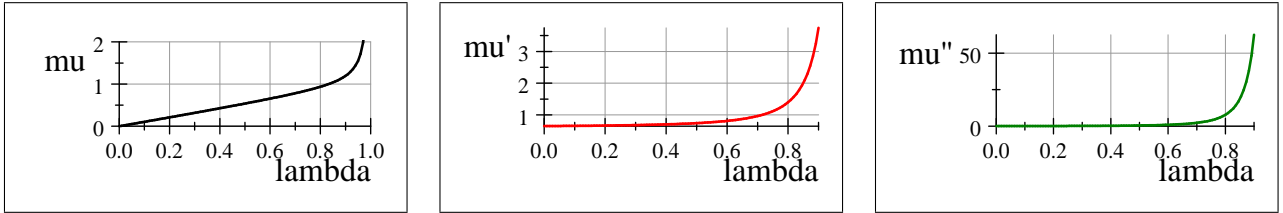
Figure 7: Interest rate

The **interest rate** r is

$$r\left(\lambda\right) = \left(1 - \tau\right)\frac{\alpha_0 + \alpha_1\lambda}{1 - \lambda}$$

with $\alpha_0 = 0.077$ and $\alpha_1 = -0.0248$ and the taxation rate $\tau = 0.40$.

Figure 8: Debt and interest paiment by unit of invested capital

The debt and interest paiment by unit of invested capital function μ is

$$\mu\left(\lambda\right) = \lambda\left(1 + r\left(\lambda\right)\right) = \lambda\left(1 - \tau\right)\left(\frac{1 + \alpha_0 - \left(1 - \alpha_1\right)\lambda}{1 - \lambda}\right).$$

Properties of $\mu\left(\cdot\right)$ are verified using its two first derivatives:

$$\mu'\left(\lambda\right) = \frac{\left(1 + \alpha_0\right) - 2\left(1 - \alpha_1\right)\lambda + \left(1 - \alpha_1\right)\lambda^2}{\left(1 - \lambda\right)^2} = \left(1 - \tau\right)\left(1 - \alpha_1 + \frac{\alpha_0 + \alpha_1}{\left(1 - \lambda\right)^2}\right),$$

$$\mu''\left(\lambda\right) = 2\left(1 - \tau\right)\frac{\left(\alpha_0 + \alpha_1\right)}{\left(1 - \lambda\right)^3}.$$

Therefore, $\mu\left(\cdot\right)$ is convex if and only if $\alpha_0 + \alpha_1 > 0$. Moreover, $\alpha_1 < 1$ and $\alpha_0 + \alpha_1 > 0$ implies that μ is increasing.
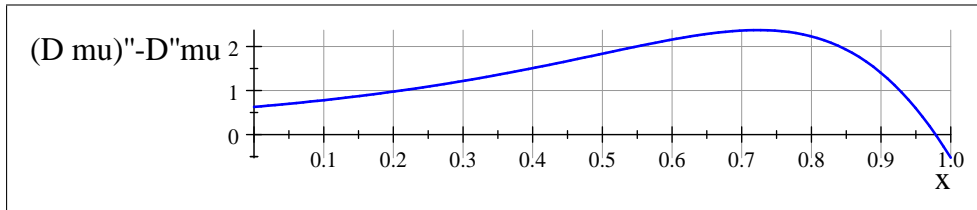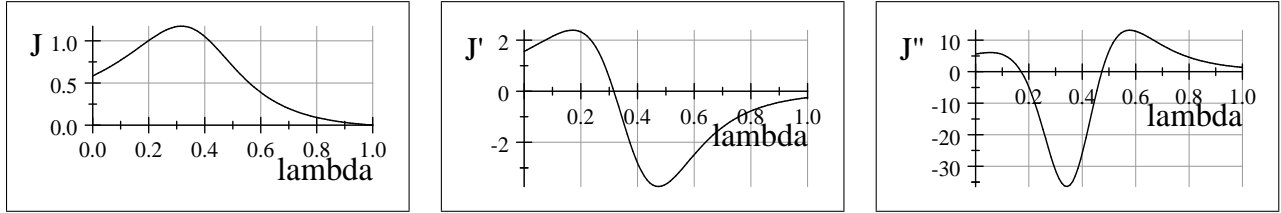
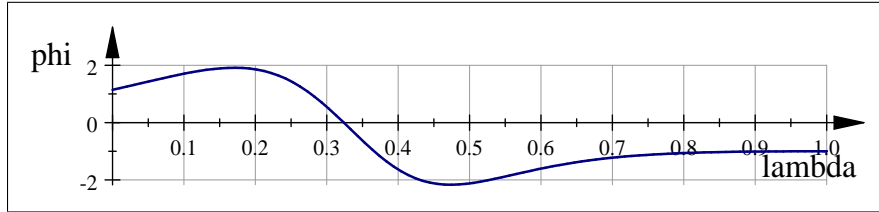Figure 9: ...

Figure 10: Continuation value at time T+1

The continuation value at time $T + 1$ is

$$J_{T+1}\left(\lambda_{T+1}\right) = \frac{\mathrm{E}\left[z\right] - g + \delta}{w\left(\lambda_{T+1}\right) - g + \delta}.$$

Note that $J'_{T+1}\left(\lambda^*\right) = 0$ implies that $\lambda^* = l = 0.316\,58$.

Figure 11: Partial derivative of the value function with respect to the investment level

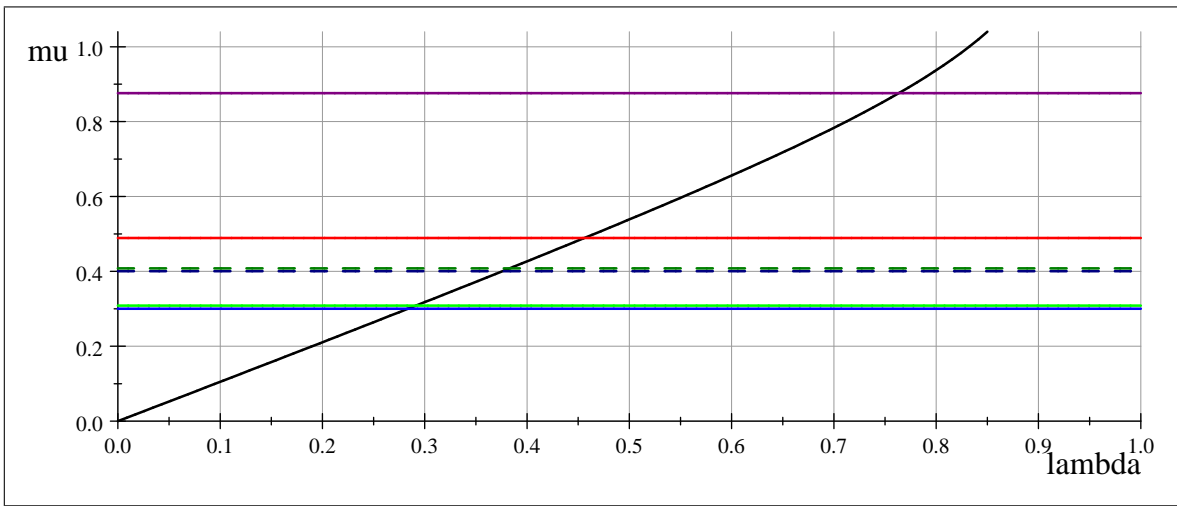$\widehat{\lambda}$ is the solution to $\phi\left(\widehat{\lambda}\right) = 0$ where

$$
\begin{aligned}
\phi\left(\lambda\right) & = J_{T+1}\left(\lambda\right) - 1 + \left(1 - \lambda\right) J'_{T+1}\left(\lambda\right) \\
& = \frac{\mathrm{E}\left[z\right] - \left(g - \delta\right)}{w\left(\lambda\right) - \left(g - \delta\right)} - 1 + \left(1 - \lambda\right)\left(-\frac{\mathrm{E}\left[z\right] - \left(g - \delta\right)}{\left(w\left(\lambda\right) - \left(g - \delta\right)\right)^2}\left(\frac{\beta_0 + \beta_1 + 2\beta_2\lambda - \beta_2\lambda^2}{\left(1 - \lambda\right)^2}\right)\right)
\end{aligned}
$$

In this application, $\widehat{\lambda} = L = 0.323\,87$

Figure 12: Leverage's thresholds at maturity

In this numerical case

$$
\begin{array}{ll}
i = I = 0.147 & \lambda^* = l = 0.31658 \\
g - \delta = 0.021 & \widehat{\lambda} = L = 0.32387 \\
\delta = 0.12 & b = 0.76337
\end{array}
$$

$\mu\left(\cdot\right)$ (black),

$\mathrm{E}\left[z\right] + \lambda^* - \left(1 - \lambda^*\right)i = 0.30012$ (light blue) $\Rightarrow \lambda_T^{(1)}\left(i\right) = 0.28351$,

$\mathrm{E}\left[z\right] + \lambda^* = 0.40058$ (dark blue - dashed) $\Rightarrow \lambda_T^{(1)}\left(0\right) = 0.3761$,

$\widehat{\lambda} - \left(1 - \widehat{\lambda}\right)i + \mathrm{E}\left[z\right] = 0.30848$ (light green) $\Rightarrow \lambda_T^{(2)}\left(i\right) = 0.29127$,

$\widehat{\lambda} + \mathrm{E}\left[z\right] = 0.40787$ (green - dashed) $\Rightarrow \lambda_T^{(2)}\left(0\right) = 0.38275$,

$\mathrm{E}\left[z\right] + \widehat{\lambda}\left(1 - \delta\right) + \delta = 0.48901$ (red) $\Rightarrow \lambda_T^{(3)} = 0.45597$,

$\mathrm{E}\left[z\right] + b\left(1 - \delta\right) + \delta = 0.87577$ (purple) $\Rightarrow b = 0.76337$.

64