Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications[*]

Bruno Biais[†]   Thomas Mariotti[‡]   Guillaume Plantin[§]   Jean-Charles Rochet[¶]

First Draft: October 2004
Revised: September 2006

Abstract

An entrepreneur with limited liability needs to finance an infinite horizon investment project. An agency problem arises because she can divert operating cash-flows before reporting them to the financiers. We first study the optimal contract in discrete time. This contract can be implemented by cash reserves, debt and equity. The latter is split between the financiers and the entrepreneur, and pays dividends when retained earnings reach a threshold. To provide appropriate incentives to the entrepreneur, the firm is downsized when it runs short of cash. We then study the continuous-time limit of the model. We prove the convergence of the discrete-time value functions and optimal contracts. Our analysis yields rich implications for the dynamics of security prices. Stock prices follow a diffusion reflected at the dividend barrier and absorbed at zero. Their volatility, as well as the leverage ratio of the firm, increase after bad performance. Stock prices and book-to-market ratios are in a non-monotonic relationship. A more severe agency problem entails lower price earning ratios and firm liquidity, and higher default risk.

[†]Université de Toulouse (GREMAQ/CNRS, IDEI, CRG) and CEPR.
[‡]Université de Toulouse (GREMAQ/CNRS, IDEI) and CEPR.
[§]London Business School.
[¶]Université de Toulouse (GREMAQ, IDEI) and CEPR.

1. INTRODUCTION

While corporate finance studies the design and issuance of securities, asset pricing investigates the dynamics of their market valuations. Our objective is to analyze these two aspects in a unified framework. In line with the corporate finance literature, we emphasize the role of agency problems for the optimal design of securities. One of the main contributions of this paper is to assess the implications of agency costs for the dynamics of key financial variables such as stock prices, leverage, book-to-market ratios, or default risk. Another contribution of this paper is methodological. While most of the literature on dynamic contracting relies on discrete-time specifications, asset pricing models are typically formulated in continuous time. Each of these approaches has its advantages and shortcomings. Discrete-time models enable one to clarify the extensive form of the game played by the different agents. Yet, continuous-time models are often simpler and analytically more tractable. Drawing from these two approaches, we first study a discrete-time agency model of the firm, and then examine its continuous-time limit obtained as the time interval between consecutive periods vanishes. Besides laying down clear conceptual foundations for continuous-time modelling, this allows us to derive a rich set of testable asset pricing implications.

Our discrete-time model is a stationary version of DeMarzo and Fishman's (2003) cash-flow diversion model. An entrepreneur has discovered an investment project, which only she can manage successfully. Because she has limited resources and limited liability, she needs to contract with financiers to fund the initial investment and the operating costs of the project, as well as her personal consumption. Both the entrepreneur and the financiers are assumed to be risk-neutral. After the initial investment has been sunk, the project generates random cash-flows according to a binomial process. Specifically, in each period, cash-flows can be negative, corresponding to operating losses, or positive, corresponding to operating profits. Cash-flows are independently distributed across periods, with constant mean and variance per unit of time. As in Bolton and Scharfstein (1990), an agency problem arises because the entrepreneur privately observes the cash-flows, and can divert them for personal consumption. An incentive compatible long-term contract incites her to report cash-flow realizations truthfully by specifying appropriate transfers and liquidation decisions contingent on each history of reports.

In line with Green (1987), Spear and Srivastava (1987) or Thomas and Worrall (1990), we use recursive techniques to characterize the optimal incentive compatible contract. In any period, a sufficient statistic of the history of past reports is the current expected discounted utility of the entrepreneur. In the optimal contract, this state variable evolves randomly between a liquidation threshold, below which the project is stochastically terminated or downsized, and a dividend threshold, above which the entrepreneur is compensated with cash. The utility of the entrepreneur increases after high cash-flow realizations, while it decreases after low cash-flow realizations. Its sensitivity to the performance of the project goes up with the severity of the agency problem, as measured by the rate at which the entrepreneur can convert concealed funds into personal consumption. These results are the stationary counterpart of DeMarzo and Fishman's (2003) characterization of the optimal financial contract in a finite horizon setting.

The second step of our analysis consists in offering a new implementation of this optimal contract with standard financial instruments. We focus on the case where financiers hold securities, that is, claims to non-negative cash-flows. This is consistent with the financiers

forming a diffuse investor base, with limited liability. Since the project can generate operating losses and the entrepreneur also has limited liability, the firm must hold cash reserves to cover these losses and continue the project. These reserves are held on a bank account, which ensures that they are observable and that their use is contractible. The securities arising in the implementation are claims on the cash-flows generated by the firm's assets, which consist of the investment project and the cash reserves. Two types of securities are issued by the firm. (i) Debt is held by the financiers only, and pays a coupon at each period until the firm is liquidated. Downsizing or liquidation occur when the firm runs short of cash and cannot service its debt. Debt is performance sensitive, as the coupon decreases with the firm's performance record. Specifically, covenants of the optimal financial contract state that as cash reserves increase and default risk decreases, the coupon rate is reduced. This property of our implementation is consistent with observed features of venture capital contracts (Kaplan and Strömberg (2003)), and with the increased reliance on step-up bonds (Lando and Mortensen (2004)) or performance pricing loans (Asquith, Beatty and Weber (2005)). (ii) Equity is held both by the entrepreneur and the financiers. In line with the empirical findings of DeAngelo, DeAngelo and Stulz (2006), dividends are distributed when retained earnings reach a performance threshold. Thus equity holders receive dividends only when the firm displays a sufficiently large ratio of liquid to total assets. To maintain incentive compatibility, the entrepreneur is compensated by holding a non-tradeable fraction of equity, which increases with the magnitude of the agency problem. Overall, the entrepreneur only receives cash when the accumulated performance of the firm reaches a contractually specified threshold (Kaplan and Strömberg (2003)).

While the implementation of the discrete-time optimal contract generates some interesting qualitative results, it falls short of delivering precise asset pricing implications. To gain more insight into the pricing of securities and the optimal financial structure of the firm, we examine a limit model obtained by letting the length of each period go to zero. Our analysis proceeds in two steps. (i) The first step is purely analytic, and consists in characterizing the continuous-time limit of the value function that describes the aggregate social surplus in the discrete-time optimal contract. Using an operator approach, we show that, as the period length goes to zero, this value function converges uniformly to the solution of a free boundary problem. While no explicit solution to this problem is available, characterizing the continuous-time limit of the value function through a simple ordinary differential equation improves significantly the analytical tractability of the model. (ii) The second step is probabilistic, and consists in characterizing the continuous-time limit of the entrepreneur's utility process in the discrete-time optimal contract. Using stability results for reflected diffusion processes (Słomiński (1993)), we show that, as the period length goes to zero, this utility process converges in law to the solution of a stochastic differential equation with a reflecting boundary condition. This limit process can be described as follows. Like its discrete-time counterpart, it evolves randomly between an absorbing liquidation boundary, at which the project is terminated, and a reflecting dividend boundary, at which the entrepreneur is compensated with cash. The instantaneous variations of this limit process can be decomposed into two terms. The first term is a diffusion, that grows on average at a rate equal to the entrepreneur's discount rate, with a constant volatility proportional to the intensity of the incentive problem and the volatility of the cash-flows. The second term is the change in a non-decreasing process which increases only when the utility of the entrepreneur hits the dividend threshold. This reflecting process corresponds to the accumulated cash compensation of the entrepreneur.

These convergence results are complementary to the contemporaneous work of DeMarzo and Sannikov (2006). Building on elegant martingale techniques introduced by Sannikov (2003), they directly analyze a continuous-time version of the cash-flow diversion model, in which cumulative cash-flows follow a Brownian motion with drift, corresponding to the limit of our discrete-time binary cash-flow process. Their characterization of the continuous-time optimal contract exactly matches our characterization of the continuous-time limit of the discrete-time optimal contract. This convergence result is not a priori obvious. Indeed, while it is customary in finance to approximate Brownian motion by binary processes, for instance to discretize asset prices processes, it cannot be taken for granted that this type of approximation remains valid in the context of an agency problem, nor that the optimal contract in the binary model will yield a good approximation of the optimal contract in the Brownian model (see Hellwig and Schmidt (2002) for a discussion of related difficulties in the context of Holmström and Milgrom's (1987) continuous-time principal-agent model).

In the last part of the paper, we use the continuous-time limit of the optimal contract and of its implementation with cash reserves, stocks and bonds to derive a rich set of novel asset pricing implications. (i) The endogenous stock price process generated by the implementation of the optimal financial contract is reminiscent of that postulated by Black and Scholes (1973). Yet, it displays several new features. First, the stock price is reflected at the dividend barrier. Next, as the stock price approaches zero, its volatility remains strictly positive. As a result of this, the price of the stock can reach zero, an event corresponding to the liquidation of the firm. Last, the volatility of the stock goes up after bad performance, in line with the leverage effect documented by Black (1976) and Nelson (1991). (ii) Performance shocks have a persistent impact on the financial structure of the firm. This is consistent with the empirical findings of Welch (2004). When the firm is financially distressed, or has low cash reserves, the stock price reacts very strongly to current operating performance. This is also true of the bond price, but to a lesser extent. Hence, leverage falls as the firm becomes more profitable and moves out of financial distress. (iii) Our theoretical analysis also sheds light on the dynamics of balance sheet items. We show that the sign of the correlation between stock returns and changes in the book-to-market ratio is not constant. For financially distressed or recently established firms with low cash reserves, stock returns and contemporaneous changes in the book-to-market ratio should be negatively correlated, while they should be positively correlated for financially healthy or mature firms. (iv) The credit yield spread provides a convenient measure of the risk of default. Our continuous-time implementation enables us to characterize this spread and to describe how it varies with the key parameters of the model, such as the intensity of the incentive problem. With a more severe agency problem, additional incentives must be given to the entrepreneur in the form of cash payments, so that the optimal contract entails lower cash reserves. This makes default more likely to occur in the near future, and consequently the credit yield spread is higher. This increase in the risk of default also reduces the valuation of the stock. An empirical implication of this is that firms facing more severe agency problems should have relatively lower price earning ratios.

The present paper is in line with several recent insightful analyses of dynamic financial contracting. Gromb (1999) studies renegotiation-proof contracts in a multi-period version of Bolton and Scharfstein's (1990) model. More recently, recursive techniques have been used by DeMarzo and Fishman (2003), Quadrini (2004) and Clementi and Hopenhayn (2006) to analyze different versions of the cash-flow diversion model. Albuquerque and Hopenhayn (2004) consider a symmetric information framework in which the entrepreneur can default

on her obligations toward the lenders. With the exception of DeMarzo and Fishman (2003), these papers do not draw the implications of optimal contracting for the firm's financial structure. Our implementation differs from DeMarzo and Fishman's (2003) in that it relies on securities and cash reserves, while they emphasize the role of credit lines.

The paper is organized as follows. Section 2 presents the model and the discrete-time optimal contract, and discusses its implementation with cash reserves, debt and equity. In Section 3, we study the convergence of the value function and of the optimal contract to their continuous-time limits. Section 4 spells out the continuous-time asset pricing implications of our model. Section 5 concludes. Proofs not given in the text are in the appendices.

2. THE DISCRETE-TIME MODEL

2.1. The contractual environment

Our basic model is an infinite horizon version of DeMarzo and Fishman's (2003) cash-flow diversion model in which cash-flows are assumed to follow a stationary binomial process. Time is discrete and indexed by nh, where $n \in \mathbb{N}$ is the number of periods elapsed, and $h > 0$ is the length of a period. We will examine in Section 3 the continuous-time limit of the model obtained as h goes to 0. There are two types of agents, an entrepreneur and financiers. The financiers are risk-neutral and have unlimited wealth. Their discount factor is $1/(1+rh)$. The entrepreneur is risk-neutral and has initial wealth $A \geq 0$. Her discount factor is $1/(1 + \rho h)$. We assume that $\rho > r$, hence the entrepreneur is more impatient than the financiers. The entrepreneur is protected by limited liability, thus contracts cannot stipulate negative payments to her.

We consider an investment project which can be managed by the entrepreneur only. Undertaking this project requires an investment outlay $I \geq 0$ at date 0. At each date, the project can be continued or liquidated. The liquidation value is set equal to 0. As long as the project is operated, it generates random cash-flows at intervals h according to a binomial process. Specifically, at any date nh prior to liquidation, the net cash-flow is given by:

$$c_{nh}^h = \mu h + \sigma \varepsilon_n \sqrt{h}, \tag{1}$$

where μ and σ are positive constants and the random variables $\{\varepsilon_n\}_{n \in \mathbb{N}}$ are identically and independently distributed, with distribution:

$$\varepsilon_n = \begin{cases} \varepsilon_+ = \sqrt{\dfrac{1-p}{p}} & \text{with probability } p, \\[4mm] \varepsilon_- = -\sqrt{\dfrac{p}{1-p}} & \text{with probability } 1-p. \end{cases} \tag{2}$$

Thus each cash-flow has a Bernoulli distribution with mean μh and variance $\sigma^2 h$. We denote by $c_+^h = \mu h + \sigma \varepsilon_+ \sqrt{h}$ and $c_-^h = \mu h + \sigma \varepsilon_- \sqrt{h}$ the possible realizations of the cash-flows. We shall assume hereafter that h is close enough to 0, so that $c_-^h < 0$. One interpretation is that the project involves operating costs $-c_-^h$ and yields a gross income $c_+^h - c_-^h$ with probability p, while with probability $1 - p$ the gross income is 0.

There are three reasons why the entrepreneur would like to contract with financiers. First, whenever $A < I$, she needs to borrow to finance the project. Second, because of

limited liability, she needs to borrow to fund operating costs. Third, because she is more impatient than the financiers, she would like to borrow to finance consumption.

Contracting between the entrepreneur and the financiers is subject to an ex-post moral hazard problem. Following Bolton and Scharfstein (1990), Gromb (1999), DeMarzo and Fishman (2003), Quadrini (2004) and Clementi and Hopenhayn (2006), we assume that the realizations of the cash-flows are privately observed by the entrepreneur. She can then conceal them from the financiers and divert them for her personal consumption. We denote by $\lambda \in (0, 1]$ the rate at which the entrepreneur can convert concealed funds into personal consumption. Thus $1 - \lambda$ is the cost of diversion. For h close enough to 0, one has:

$$p\lambda(c_+^h - c_-^h) > \mu h. \tag{3}$$

Under condition (3), the cost of diversion is low, and the temptation to divert funds is high relative to the expected cash-flow. As a result of this, the agency problem is severe. It is easy to check that, in the static version of the model, condition (3) implies that even if the project requires no investment outlay, $I = 0$, it cannot be financed with probability 1 whenever the entrepreneur has low initial wealth. Thus static contracting would lead to credit rationing.

Like the cash-flow realizations, the consumption and saving decisions of the entrepreneur are not observable by the financiers. As a result of this, financial contracts can be contingent only on payments and reports made by the entrepreneur. Given a public history of such payments and reports up to date nh, the contract first specifies the probability x_{nh}^h with which the project is continued. If the project is liquidated, the financiers and the entrepreneur get their reservation utility, which is 0 as the project has no liquidation value.[1] If the project is continued, the financiers pay $-c_-^h$ to the entrepreneur to cover operating costs, the use of these funds for operating the project being verifiable. After the gross income $c_{nh}^h - c_-^h$ is realized, the entrepreneur makes a payment to the financiers, possibly concealing some of the cash-flow, and sends a report. Taking into account this new information, the contract then specifies a non-negative transfer u_{nh}^h from the financiers to the entrepreneur. Finally, the entrepreneur chooses how much to consume from concealed funds, at rate λ, and from her current transfer. Remaining funds accumulate on a private saving account at an interest rate r' at most equal to the market rate r. The balances on this account are not observable by the financiers.

To focus on a single source of market imperfection, we assume that the entrepreneur and the financiers can fully commit to a long-term contract.[2] One can then simplify the analysis considerably by relying on a version of the revelation principle. As shown by DeMarzo and Fishman (2003, Proposition 1), there is no restriction in focusing on incentive compatible direct revelation mechanisms such that the entrepreneur pays $c_{nh}^h - c_-^h$ to the financiers and abstains from private saving, and in which other reports are irrelevant. This is because both diversion and private savings are weakly inefficient, $\lambda \le 1$ and $r' \le r$. Equivalently, the entrepreneur reports truthfully each realization of the cash-flow, and receives a non-negative transfer contingent on her report, which she consumes immediately. To induce truthful revelation of the cash-flows, financiers will rely on transfers contingent on the history of reports, as well as on the threat of liquidating the project.

---

[1] In full generality, one should also allow for transfers from the financiers to the entrepreneur in case of liquidation. However, promising positive transfers to the entrepreneur in this event would weaken incentives and be suboptimal. To ease the exposition, we therefore rule out such transfers from the outset.

[2] DeMarzo and Fishman (2003) also provide an analysis of renegotiation-proof contracts.

Remark. In this model, an incentive problem arises because the entrepreneur can divert operating cash-flows. An alternative approach would be to assume that the manager must exert unobservable effort to enhance cash-flows, as in Innes (1990), Holmström and Tirole (1997) and in a previous draft of the present paper. While the two interpretations of the agency problem are slightly different and complementary, it turns out that when the effort choice is binary, the formal analysis of the two models is identical, except in one respect. In the cash-flow diversion model, the revelation principle implies that there is no loss of generality in requiring truthful revelation of cash-flow realizations. By contrast, in the unobservable effort model, one must impose additional restrictions to ensure that it is optimal to request effort in all contingencies, which raises additional mathematical difficulties. To simplify the exposition, we have therefore chosen to focus on the cash-flow diversion scenario.

2.2. The dynamic programming formulation

Following DeMarzo and Fishman (2003), we rely on a dynamic programming formulation to solve for the optimal contract. Specifically, it can be shown as in Green (1987), Spear and Srivastava (1987) or Thomas and Worrall (1990) that, at any date, the expected discounted utility of the entrepreneur in the continuation of the optimal contract is a sufficient statistic for the information provided by the history of payments and reports up to this date. At any date prior to liquidation, and given a current promised utility w for the entrepreneur, the optimal contract thus specifies a continuation probability x and, conditional on the project not being liquidated, transfers to the entrepreneur, u_+ and u_-, as well as continuation utilities for the entrepreneur, w_+ and w_-, contingent on the report c_+^h or c_-^h she makes about the current realization of the cash-flow. The contract must satisfy the following feasibility and limited liability constraints:

$$(x, u_+, u_-, w_+, w_-) \in [0, 1] \times \mathbb{R}_+^4. \tag{4}$$

We first solve for the optimal contract assuming that the entrepreneur cannot save privately. We then argue that this optimal contract is robust to private savings.

Let $F^h(w)$ be the expected discounted utility of the financiers given a promised utility w for the entrepreneur in the optimal incentive compatible contract. At any date prior to liquidation, the financiers receive the cash-flow from the project, net of the transfer to the entrepreneur. The function F^h satisfies the following Bellman equation:

$$F^h(w) = \max \left\{ x \left[\mu h - p u_+ - (1 - p) u_- + \frac{p F^h(w_+) + (1 - p) F^h(w_-)}{1 + rh} \right] \right\} \tag{5}$$

for all $w \geq 0$, subject to (4) and to constraints (6)–(7) stated below. First, the contract must satisfy a consistency condition, stating that the utility of the entrepreneur at the beginning of any period must be equal to the expected payment made to her during the period, plus the expected present value of her continuation utility:

$$w = x \left[p u_+ + (1 - p) u_- + \frac{p w_+ + (1 - p) w_-}{1 + \rho h} \right]. \tag{6}$$

Second, we impose the condition that the entrepreneur always reports truthfully high cash-flow realizations. Since by assumption the entrepreneur cannot save privately, this is the case if the total payoff following a high cash-flow realization is larger than the sum of the total

6

payoff following a low cash-flow realization and of the gain in terms of current consumption of diverting the gross income $c_+^h - c_-^h$, that is:

$$u_+ + \frac{w_+}{1 + \rho h} \geq u_- + \frac{w_-}{1 + \rho h} + \lambda(c_+^h - c_-^h). \tag{7}$$

At this point, we could appeal to DeMarzo and Fishman's (2003) results to characterize the optimal contract, adjusting for the fact that our model is stationary while theirs is framed in a finite horizon setting. Our restriction to binary cash-flows enables us however to provide a simple and intuitive derivation of the optimal contract, which paves the way to the analysis of its implementation and of its continuous-time limit.

The first step of our analysis consists in temporarily eliminating current transfers u_+ and u_- from the formulation of the problem. To do so, it is convenient to introduce an auxiliary function defined by $V^h(w) = w + F^h(w)$ for all $w \geq 0$. The function V^h measures the social surplus generated by the project, which is independent of current transfers. One has then the following lemma.

Lemma 1. *The function V^h satisfies the following Bellman equation:*

$$V^h(w) = \max \left\{ x \left\{ \mu h + \frac{p V^h(w_+) + (1-p) V^h(w_-)}{1 + rh} - \frac{(\rho - r)h[p w_+ + (1-p) w_-]}{(1 + rh)(1 + \rho h)} \right\} \right\} \tag{8}$$

for all $w \geq 0$, subject to the constraints:

$$(x, w_+, w_-) \in [0, 1] \times \mathbb{R}_+^2, \tag{9}$$

$$w \geq x \left[\frac{w_-}{1 + \rho h} + p\lambda(c_+^h - c_-^h) \right], \tag{10}$$

$$w \geq x \left[\frac{p w_+ + (1-p) w_-}{1 + \rho h} \right]. \tag{11}$$

It follows immediately from (9)–(10) that, for $w < p\lambda(c_+^h - c_-^h)$, liquidation must occur with positive probability. The probability of liquidation goes to 1 as w goes to 0. This reflects the fact that the optimal contract relies on the threat of project termination as well as on continuation utilities to provide incentives to the entrepreneur. When the entrepreneur's utility is low, most incentives come from the threat of liquidation. Note that this threat is socially costly since, even at the time of liquidation, the project remains potentially profitable. The next lemma states that the solution V^h to (8)–(11) exists and is unique, and establishes some of its properties.

Lemma 2. *There exists a unique continuous and bounded solution V^h to (8) subject to (9)–(11). The function V^h is non-decreasing, concave, and vanishes at 0. Moreover, there exist two thresholds $w^{h,m} \geq w^{h,l} \geq p\lambda(c_+^h - c_-^h)$ such that:*

   (i) *If $w \in [0, w^{h,l})$, the project is continued with probability $w/w^{h,l}$ and liquidated with probability $1 - w/w^{h,l}$. If $w \in [w^{h,l}, \infty)$, the project is continued with probability 1.*

   (ii) *The function V^h is linear over the liquidation region $[0, w^{h,l})$, strictly increasing over $[0, w^{h,m}]$, and constant over $[w^{h,m}, \infty)$.*

In the liquidation region $[0, w^{h,l})$, the continuation probability x is a linear function of the entrepreneur's utility. If $w \in [0, w^{h,l})$ and the project is continued, the entrepreneur's utility first jumps to $w/x = w^{h,l}$, and the optimal contract starting at $w^{h,l}$ is then immediately executed. If the project is liquidated, the continuation utilities are 0 for both the entrepreneur and the financiers. As a result of this, $V^h(w) = (w/w^{h,l})V^h(w^{h,l})$ for any $w \in [0, w^{h,l})$, and V^h is linear over the liquidation region.

2.3. The optimal contract

We now derive the solution to (8)–(11), from which the optimal contract will follow. At the upper end of the state space, the interval $(w^{h,m}, \infty)$ corresponds to the region where constraints (10)–(11) are slack. As can be seen from (8), it is then optimal to set $x = 1$ and:

$$w_+ = w_- = w^{h,r}, \tag{12}$$

where $w^{h,r}$ is the smallest point at which the mapping $w \mapsto V^h(w) - (\rho - r)hw/(1 + \rho h)$ reaches its maximum. By construction, $w^{h,r} \leq w^{h,m}$. At $w^{h,m}$, (10) is just binding, so that:

$$w^{h,m} = \frac{w^{h,r}}{1 + \rho h} + p\lambda(c_+^h - c_-^h), \tag{13}$$

while (11) is slack. When $w \geq w^{h,m}$, incentives are provided by current transfers, and not through continuation utilities or liquidation threats. To determine the optimal transfers, we return to the original formulation (4)–(7) of the problem. The consistency constraint (6) pins down the expected transfer $pu_+ + (1 - p)u_-$, while the incentive compatibility constraint (7) pins down the difference $u_+ - u_-$. Defining:

$$w^{h,d} = w^{h,m} - \lambda(c_+^h - c_-^h), \tag{14}$$

this yields:

$$u_+ = w - w^{h,d}, \tag{15}$$

$$u_- = w - w^{h,m}. \tag{16}$$

Additional information on the thresholds $w^{h,l}$, $w^{h,r}$ and $w^{h,m}$ is provided by the following lemma.

Lemma 3. *For h close enough to 0, $w^{h,l} \leq w^{h,r} < w^{h,m}$.*

Note that, by construction, the utility of the entrepreneur never leaves the interval $[0, w^{h,r}]$ once it has entered it, which occurs after at most one period. A key implication of Lemma 3 is that $w^{h,r}$ is a reflecting barrier for the utility of the entrepreneur. This result hinges on the assumption that $\rho > r$. If the entrepreneur and the financiers were equally patient, that is if $\rho = r$, as in the model of Clementi and Hopenhayn (2006), there would be no cost in delaying entrepreneurial compensation, and capitalizing the rewards promised to the entrepreneur at rate r. In that case, it would never be optimal to make transfers to the entrepreneur before reaching the point at which enough profits have been accumulated to finance the incentive

costs without ever relying on the liquidation threat.[3] By contrast, when $\rho > r$, there is a trade-off between the efficiency gains from letting the entrepreneur consume early and the risk of liquidation. As a result of this, it is optimal to let the entrepreneur consume while the risk of liquidation is not fully eliminated, so that there always remains a strictly positive probability that the project will be liquidated in the future.

It remains to characterize the solution to (8)–(11) over the interval $[w^{h,l}, w^{h,m}]$, where $x = 1$ and at least one of the constraints (10)–(11) is binding. Using the concavity of V^h, one obtains that it is optimal to let (10) be binding, so that:

$$w_- = (1 + \rho h)[w - p\lambda(c_+^h - c_-^h)]. \tag{17}$$

Concerning w_+, two cases may occur. Either (11) is slack, so that an optimal choice is:

$$w_+ = w^{h,r}. \tag{18}$$

As can be seen from (11) and (13)–(14), this is the case if and only if $w > w^{h,d}$. Alternatively, if $w \leq w^{h,d}$, (11) is binding, so that from (11) and (17):

$$w_+ = (1 + \rho h)[w + (1 - p)\lambda(c_+^h - c_-^h)]. \tag{19}$$

To determine the optimal transfers, we return to the original formulation (4)–(7) of the problem. If $w > w^{h,d}$ this yields:

$$u_+ = w - w^{h,d}, \tag{20}$$

$$u_- = 0, \tag{21}$$

while if $w \leq w^{h,d}$ this yields:

$$u_+ = u_- = 0. \tag{22}$$

The following proposition offers a summary of the results derived in (12)–(22),[4] providing a stationary analogue to DeMarzo and Fishman's (2003, Propositions 2–4) characterization of the optimal contract in a finite horizon setting.

Proposition 1. *The optimal contract without private savings is characterized by two regimes:*

(i) *If $w \in [0, w^{h,l})$, the project is continued with probability $x = w/w^{h,l}$ and liquidated with probability $1 - x = 1 - w/w^{h,l}$. If the project is continued, the optimal contract starting at $w/x = w^{h,l}$ is immediately executed.*

(ii) *If $w \in [w^{h,l}, \infty)$, the project is continued with probability 1. The optimal continuation utilities are given by:*

$$w_+ = \min\{(1 + \rho h)[w + (1 - p)\lambda(c_+^h - c_-^h)], w^{h,r}\}, \tag{23}$$

$$w_- = \min\{(1 + \rho h)[w - p\lambda(c_+^h - c_-^h)], w^{h,r}\}, \tag{24}$$

---

[3]This case is not conducive to continuous-time analysis, since the present value $(1 + rh)p\lambda(c_+^h - c_-^h)/(rh)$ of these incentive costs goes to ∞ as h goes to 0. While the continuous-time limit of the discrete-time value function is still well-defined, the discrete-time optimal contract does not converge to a well-defined limit.

[4]We thank an anonymous referee for suggesting this formulation to us.

while the optimal current transfers are given by:

$$u_+ = \max\{w - w^{h,d}, 0\}, \tag{25}$$

$$u_- = \max\{w - w^{h,m}, 0\}. \tag{26}$$

As a result of this, the total utility of the entrepreneur given a cash-flow realization c is $w - p\lambda(c_+^h - c_-^h) + \lambda(c - c_-^h)$. If this exceeds $w^{h,r}/(1 + \rho h)$, the entrepreneur is paid the excess over $w^{h,r}/(1 + \rho h)$ in cash.

This result can be interpreted as follows. For incentive compatibility reasons, the total utility of the entrepreneur increases in the cash-flow she reports. As shown by (23)–(24), her utility increases following a high cash-flow realization, while it decreases following a low cash-flow realization. Another key implication of (23)–(24) is that, as λ increases and thus the agency problem intensifies, the utility of the entrepreneur becomes more sensitive to the performance of the firm. Finally, as shown by (25)–(26), the entrepreneur receives immediate cash compensation only when the accumulated performance of the firm reaches a contractually specified threshold. Below this threshold, the entrepreneur is compensated solely by the promise of future payments.

We have derived the optimal contract under the assumption that the entrepreneur cannot save privately. In line with DeMarzo and Fishman (2003, Corollary 3A), a key observation is that the marginal benefit of reporting a higher cash-flow realization is constant and equal to λ. As a result of this, since the discount rate ρ of the entrepreneur is greater than the return r' on private savings, she has no incentive to conceal a high cash-flow realization in order to report higher cash-flow realizations in the future. That is, the optimal contract derived in Proposition 1 is robust to private savings, and therefore coincides with the optimal contract when the entrepreneur can save privately.

2.4. Initializing the optimal contract

To fully characterize the dynamics of the contractual relationship, one needs only to specify the initial value of the entrepreneur's utility, w_0^h. Recall that $F^h(w) = V^h(w) - w$ measures the expected discounted utility of the financiers in the optimal contract given a promised utility w for the entrepreneur. In line with Holmström and Tirole (1997), one can then interpret $\max_{w \in \mathbb{R}_+}\{F^h(w)\}$ as the maximum income that can be pledged to financiers. Two cases can then arise. If $\max_{w \in \mathbb{R}_+}\{F^h(w)\} < I - A$, then the maximum pledgeable income is lower than the need for outside funds, so that the project cannot be financed with probability 1.[5] Note that this may occur despite the project having a positive net present value in the absence of agency costs, $\mu(1 + rh)/r > I$, reflecting the possibility of credit rationing. If, on the contrary, $\max_{w \in \mathbb{R}_+}\{F^h(w)\} \geq I - A$, then the project can be financed with probability 1, in contrast with the static case. The initial rent $F^h(w_0^h) - (I - A)$ of the financiers then depends on the allocation of bargaining power between the entrepreneur and the financiers. If the entrepreneur has all the bargaining power at date 0, she maximizes her initial utility subject to the financiers' participation constraint:

$$w_0^{h,E} = \max\{w \in \mathbb{R}_+ \,|\, F^h(w) \geq I - A\}.$$

---

[5] As shown by DeMarzo and Fishman (2003), randomization at this initial stage would enable funding, although with a probability less than 1.

If, on the contrary, the financiers have all the bargaining power at date 0, they choose the entrepreneur's initial utility so as to obtain the maximum pledgeable income:

$$w_0^{h,F} \in \arg\max_{w \in \mathbb{R}_+} \{F^h(w)\}.$$

Since F^h is concave and bounded above, this situation leads to a lower initial utility for the entrepreneur, $w_0^{h,F} \leq w_0^{h,E}$, and thus to a lower social surplus, $V^h(w_0^{h,F}) \leq V^h(w_0^{h,E})$. More generally, the higher is the financiers' bargaining power, the lower are the entrepreneur's initial utility and the social surplus. This reflects the rent-efficiency trade-off typically arising in principal-agent problems. Since the entrepreneur is cash-constrained, she cannot generate a Pareto improvement by offering a monetary transfer in exchange for an increase in her initial utility.

Once the initial value of the entrepreneur's utility has been specified, the dynamics of the optimal contract is fully determined by Proposition 1. We denote by $\{x_{nh}^h\}_{n \in \mathbb{N}}$, $\{w_{nh}^h\}_{n \in \mathbb{N}}$ and $\{u_{nh}^h\}_{n \in \mathbb{N}}$ the stochastic processes describing the continuation probabilities, the utility of the entrepreneur and the transfers to the entrepreneur in the optimal contract. Note that the utility of the entrepreneur enters the interval $[0, w^{h,r}]$ after at most one period and never leaves it thereafter. Hence there is no significant loss of generality in assuming that $w_0^h \leq w^{h,r}$. In this case, it follows from (26) that transfers are never distributed to the entrepreneur in case of a low cash-flow realization. Proposition 1 can then be rephrased as follows. As long as the project is continued, which occurs in each period with probability:

$$x_{nh}^h = \min\left\{\frac{w_{nh}^h}{w^{h,l}}, 1\right\}, \tag{27}$$

the utility of the entrepreneur evolves according to:

$$w_{(n+1)h}^h = \min\left\{(1 + \rho h)\left[\frac{w_{nh}^h}{x_{nh}^h} + \lambda(c_{nh}^h - \mu h)\right], w^{h,r}\right\}, \tag{28}$$

while the transfers to the entrepreneur are given by:

$$u_{nh}^h = \max\left\{\frac{w_{nh}^h}{x_{nh}^h} - w^{h,d}, 0\right\} 1_{\{c_{nh}^h = c_+^h\}}. \tag{29}$$

Once the project is liquidated, the entrepreneur receives no further transfers, and her utility stays constant at 0.

2.5. Implementing the optimal contract with cash reserves and securities

So far, our results are closely in line with DeMarzo and Fishman (2003). We now depart from their analysis, starting with a new implementation of the optimal contract characterized in Proposition 1. In general, several implementations are possible. Indeed, as long as incentive, limited liability and participation constraints are satisfied, the Modigliani and Miller (1958) logic applies, and slicing and dicing of cash-flows is irrelevant. To narrow down the set of possible implementations, we impose that financiers hold securities, defined as claims with limited liability, which can be held by a diffuse investor basis. This contrasts with DeMarzo and Fishman (2003), who show how to implement the optimal contract using a credit line, together with long-term debt and equity. Ownership of securities by a diffuse investor basis

11

hinders renegotiation. This provides a motivation for our assumption that parties can commit to a long-term contract. As argued below, our implementation is consistent with stylized facts and empirical evidence about the financial structure of firms. In Section 4, we also show that it generates a rich set of testable implications for the pricing of securities.

At each date, operating costs can exceed operating revenues, resulting in a negative net cash-flow. To meet these cash outlays while preserving the entrepreneur's and the financiers' limited liability, the firm must hold cash reserves. A consequence of limited liability is that the project must be liquidated when the firm runs out of cash. Cash reserves are held on a bank account, and their use is assumed to be verifiable. This differs from the operating cash-flow generated inside the firm, which can be diverted and concealed by the entrepreneur. Interest is earned at rate r on the cash deposited at the bank. In equilibrium, the change in cash reserves at each date is thus equal to the net operating cash-flow, plus the interest on cash reserves, less the payments to the entrepreneur and the financiers. Since no further investments are made after date 0, these additional cash reserves are accounted for as retained earnings in the financial statements of the firm.

For implementation purposes, it is convenient to interpret x as an irreversible downsizing factor rather than as a liquidation probability. When $x < 1$, a fraction $1 - x$ of the project is liquidated, and the firm continues to operate at a reduced scale. We assume constant returns to scale, so that current and future cash-flows are scaled down by a factor x, along with the utilities of the entrepreneur and of the financiers. According to this downsizing interpretation, w represents the size-adjusted entrepreneur's utility, while $V^h(w)$ represents the size-adjusted social surplus. The size-adjusted utility of the entrepreneur evolves as follows. If $w \geq w^{h,l}$, no downsizing takes place in the current period. The size-adjusted continuation utilities are then given by (23)–(24), and the size-adjusted transfers by (25)–(26). By contrast, if $w < w^{h,l}$, the firm is first scaled down by a factor $x = w/w^{h,l}$, and the optimal contract starting at a size-adjusted utility level $w/x = w^{h,l}$ is then immediately executed. This interpretation is consistent with all the formal results hitherto obtained. In particular, the Bellman equation (8)–(11) is the same under the downsizing and the probabilistic interpretation of liquidation. We return to the latter in Section 3, and we show in Section 4 that the two interpretations coincide in the continuous-time limit of the model.

We denote by $\{\hat{x}_{nh}^h\}_{n \in \mathbb{N}}$, $\{\hat{w}_{nh}^h\}_{n \in \mathbb{N}}$ and $\{\hat{u}_{nh}^h\}_{n \in \mathbb{N}}$ the stochastic processes describing the downsizing factors, the size-adjusted utility of the entrepreneur and the size-adjusted transfers to the entrepreneur in the optimal contract. These processes are formally defined in Appendix A. By construction, they coincide with the processes $\{x_{nh}^h\}_{n \in \mathbb{N}}$, $\{w_{nh}^h\}_{n \in \mathbb{N}}$ and $\{u_{nh}^h\}_{n \in \mathbb{N}}$ defined by (27)–(29) until the period in which liquidation can occur according to the probabilistic interpretation. The size of the firm at the beginning of period n is given by $\prod_{i=0}^{n-1} \hat{x}_{ih}^h$, with the convention $\prod_{i=0}^{-1} = 1$.

In the abstract characterization of the optimal contract, the relevant state variable is the size-adjusted utility of the entrepreneur. In the implementation of the optimal contract, payoffs and decisions are contingent on size-adjusted cash reserves, defined as the ratio of cash reserves to the current size of the firm. We design the implementation in such a way that these two state variables are informationally equivalent. Specifically, the size-adjusted cash reserves $\hat{m}_{nh}^h$ will be constructed so that $\hat{m}_{nh}^h = \hat{w}_{nh}^h/\lambda^h$ at any date nh, where $\lambda^h = (1+\rho h)\lambda/(1+rh)$.[6] In turn, the cash reserves held by the firm on its bank account at the beginning of period

---

[6]Similarly, Shim (2006) analyzes a dynamic agency model of bank regulation, in which the level of book value capital plays the role of a record keeping device, linked to the banker's continuation utility.

n are given by $m_{nh}^h = (\prod_{i=0}^{n-1} \hat{x}_{ih}^h)\hat{m}_{nh}^h$. We focus hereafter on the case where $\hat{w}_0^h \leq w^{h,r}$. If $\hat{w}_0^h > w^{h,r}$, one can pay a special dividend to the entrepreneur at date 0, and then apply our implementation of the optimal contract.

Proposition 2. *Whenever λ is close enough to 1 and h is close enough to 0, the optimal contract can be implemented as follows. At date 0, the firm issues stocks and bonds. The entrepreneur contributes her initial wealth A and is granted a non-tradeable fraction λ of the stocks. The financiers receive the remaining fraction $1 - \lambda$ of the stocks and all of the bonds. The firm uses the entrepreneur's contribution and the proceeds from the issuance to pay the investment cost I and to hoard an amount of cash $\hat{m}_0^h = \hat{w}_0^h/\lambda^h$. At the beginning of period n, the size of the firm is $\prod_{i=0}^{n-1} \hat{x}_{ih}^h$ as specified by the optimal contract. Then, at any date nh, the implementation is characterized by two regimes:*

   (i) If $\hat{m}_{nh}^h \in [0, w^{h,l}/\lambda^h)$, the firm is scaled down by a factor $\hat{x}_{nh}^h = \hat{m}_{nh}^h/(w^{h,l}/\lambda^h)$, after which the implementation starting at $\hat{m}_{nh}^h/\hat{x}_{nh}^h = w^{h,l}/\lambda^h$ is immediately executed.

   (ii) If $\hat{m}_{nh}^h \in [w^{h,l}/\lambda^h, w^{h,r}/\lambda^h]$, $\hat{x}_{nh} = 1$ and no downsizing takes place. Stocks distribute a dividend:

$$\hat{e}_{nh}^h = \left(\prod_{i=0}^n \hat{x}_{ih}^h\right)\max\left\{\frac{\lambda^h \hat{m}_{nh}^h}{\lambda} - \frac{w^{h,d}}{\lambda}, 0\right\}1_{\{c_{nh}^h = c_+^h\}}, \tag{30}$$

while bonds distribute a coupon:

$$\hat{b}_{nh}^h = \left(\prod_{i=0}^n \hat{x}_{ih}^h\right)\left[\mu h - \frac{(\rho - r)h\hat{m}_{nh}^h}{1 + rh}\right]. \tag{31}$$

In particular, stocks distribute a dividend if and only if, following a high cash-flow realization, the size-adjusted cash reserves $\hat{m}_{(n+1)h}^h$ reach $w^{h,r}/\lambda^h$, while no dividend is distributed following a low cash-flow realization.

Rules (i)–(ii) in Proposition 2 are specified at date 0 as covenants of the optimal contract. Afterwards, in the execution of the contract, neither the entrepreneur nor the financiers have discretion over these covenants. In line with our full commitment assumption, there is thus no scope for renegotiating the liquidation and payout decisions.

In this implementation, the entrepreneur is compensated with a share λ of the stocks. Since $\lambda^h \hat{m}_{nh}^h = \hat{w}_{nh}^h$ by construction, the dividend (30) paid by the stocks reflects that, when no downsizing takes place in period n, the size-adjusted transfer $\hat{u}_{nh}^h$ received by the entrepreneur in case of success is given by $\max\{\hat{w}_{nh}^h - w^{h,d}, 0\}$, which is the analogue of (25) under the downsizing interpretation. The assumption that λ is high enough and h is small enough guarantees that the coupon (31) paid by the bonds remains non-negative for each level of the size-adjusted cash reserves.

Our implementation has several attractive features, consistent with stylized facts and empirical findings about the financial structure of firms.

Balance sheet. The implementation essentially consists in creating a firm that is jointly owned by the entrepreneur and the financiers. The value of this firm reflects not only the investment project, but also the level of its cash reserves. On the liability side of the firm's balance sheet, the counterparts of these assets are debt, held by the financiers only, and equity, held by both the entrepreneur and the financiers. These securities are limited liability claims on the cash-flows generated by the assets of the firm.

Financial distress. The firm is downsized when operating losses drive size-adjusted cash reserves to a very low level. This is consistent with the stylized fact that financially distressed firms are often downsized (Denis and Shome (2005)). A natural interpretation of downsizing is that the firm must maintain a minimum liquidity ratio $w^{h,l}/\lambda^h$, below which a fraction of its assets must be liquidated.

Stocks. Stocks are claims on dividends which are distributed only when cash reserves reach a contractually specified threshold. As pointed out by Brealey and Myers (2000, Section 24.6), bond covenants typically include clauses restricting the set of circumstances in which dividends can be paid. In particular, contractual clauses often preclude dividend payments when the firm does not have a sufficiently large ratio of liquid assets to total assets. Our results are also in line with the empirical findings of DeAngelo, DeAngelo and Stulz (2006), who observe a significant positive relationship between the distribution of dividends and the level of retained earnings, which amount to cash reserves in our model. Consistent with our emphasis on cash reserves, firms that consistently paid dividends in their sample displayed a large ratio of cash balances to total assets. Our model also implies that dividends are paid only after the firm has established a sufficient performance record. This is in line with the stylized fact that dividends are paid by large and mature firms, while young firms, especially in the high-tech industry, pay no cash to shareholders for long periods of time (Bulan, Subramaniam and Tanlu (2005)).

Managerial compensation. In line with stylized facts on executive compensation (Dial and Murphy (1995), Murphy (1999)), the entrepreneur, who is also the manager of the project, is compensated with restricted stocks, which she cannot sell. If she could, this would curb her incentives to refrain from diverting cash-flows. The entrepreneur's share of equity increases with the intensity of the incentive problem as measured by λ. Dividends on these stocks are paid after the firm has accumulated high enough cash reserves. This is consistent with the empirical findings of Kaplan and Strömberg (2003) that financial contracts typically specify that managers receive cash compensation when performance milestones are reached. Kaplan and Strömberg (2004) find that the use of such performance benchmarks increases with asymmetric information about the operations of the firm.

Bonds. While stocks distribute dividends only when cash reserves reach a threshold, bonds pay a coupon in each period. If the entrepreneur and the financiers were equally patient, the size-adjusted coupon would be constant, as in a standard debt contract. By contrast, when $\rho > r$, the coupon decreases with size-adjusted cash reserves, which reflect accumulated firm performance. This is consistent with clauses observed in practice in financial contracts. Kaplan and Strömberg (2003) find that venture capitalists often hold preferred shares, which are similar to bonds in that they deliver contractually specified revenues, to be paid before any dividend. As observed by Kaplan and Strömberg (2003, Table 3), the contracts defining these claims typically include clauses stating that their revenue is reduced if performance goals are attained, or that their owner is entitled to additional compensation if the performance of the firm lies below a certain threshold. The payments on the bonds in Proposition 2 are also in line with those of step-up bonds, which have been issued in large amounts over the recent years, especially in the European telecom industry (Lando and Mortensen (2004)). Such bonds have provisions stating that the coupon payments increase as the credit rating of the issuer deteriorates, which in our model corresponds to a decline

in cash reserves. The bonds in our implementation are also similar to performance pricing loans, for which the interest is tied to some pre-specified measure of the performance of the borrower.[7] Asquith, Beatty and Weber (2005) document the prevalence of such clauses in bank loans.

3. THE CONTINUOUS-TIME LIMIT

In this section, we examine the properties of the continuous-time limit of the model obtained when the length of each period goes to 0. This will enable us in Section 4 to derive sharp predictions on the pricing of securities used to implement the optimal contract.

3.1. Convergence of the value functions

To begin with, we study the limit of the value functions characterized in Lemma 1. Our first result shows that the liquidation region $[0, w^{h,l})$ shrinks to $\{0\}$ as h goes to 0. Thus in the continuous-time limit, liquidation is no longer probabilistic, or alternatively no size adjustments need to take place before the firm is liquidated, unlike in the discrete-time framework.

Lemma 4. $\lim_{h\downarrow 0} w^{h,l} = 0$.

The analysis now proceeds in several steps. First, we provide a heuristic derivation of the continuous-time limit of the discrete-time value functions. Next, we state and discuss our main convergence result. Finally, we provide an outline of the proof of this result.

A heuristic derivation. We first provide an intuitive approximation of the value function V^h for h close to 0. To do so, we focus on the region $[w^{h,l}, w^{h,d}]$ of the state space in which the project is continued with probability 1, but no transfers are made to the entrepreneur. For any w in this region, the continuation utilities w_- and w_+ are given by (17) and (19). By (1)–(2), $c_+^h - c_-^h = \sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h}$, so that Taylor-Young approximations yield:

$$V^h(w_+) = V^h(w) + V^{h\prime}(w)[\rho w h + (1-p)\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h}]$$

$$+ \frac{1}{2} V^{h\prime\prime}(w)(1-p)^2\lambda^2\sigma^2(\varepsilon_+ - \varepsilon_-)^2 h + o(h)$$

and:

$$V^h(w_-) = V^h(w) + V^{h\prime}(w)[\rho w h - p\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h}]$$

$$+ \frac{1}{2} V^{h\prime\prime}(w)p^2\lambda^2\sigma^2(\varepsilon_+ - \varepsilon_-)^2 h + o(h).$$

Taking expectations and using the fact that $p(1-p)(\varepsilon_+ - \varepsilon_-)^2 = 1$ by (2) leads to:

$$pV^h(w_+) + (1-p)V^h(w_-) = V^h(w) + V^{h\prime}(w)\rho w h + \frac{1}{2} V^{h\prime\prime}(w)\lambda^2\sigma^2 h + o(h).$$

[7]Tchistyi (2006) extends DeMarzo and Fishman's (2003) analysis to the case where cash-flows are serially correlated. He offers an implementation of the optimal contract using a credit line in which a higher interest rate is charged when the balance is high. Such performance pricing is similar to our findings.

Substituting this into (8) and using the fact that $pw_+ + (1-p)w_- = (1+\rho h)w$ by (17) and (19), one obtains the following approximation:

$$rV^h(w) = \mu - (\rho - r)w + \rho w V^{h\prime}(w) + \frac{\lambda^2 \sigma^2}{2} V^{h\prime\prime}(w) + o(1). \qquad (32)$$

This intuition for (32) is that the total social return of the project is approximatively equal to the expected cash-flow μ, plus the social benefit $\rho w V^{h\prime}(w) - (\rho - r)w$ of providing utility w to the entrepreneur, less the social cost $-\lambda^2 \sigma^2 V^{h\prime\prime}(w)/2$ of incentives. The above derivation is heuristic in that it presumes that the function V^h is twice differentiable, which need not always be the case. Moreover, while we know that $\lim_{h\downarrow 0} w^{h,l} = 0$, we have little information so far about the behavior of $w^{h,d}$ for h close to 0, so that the range over which (32) is valid remains unclear. Finally, (32) gives us information about the discrete-time value functions, not directly about their continuous-time limit. What needs to be established is that an exact version of (32) holds in the limit as h goes to 0.

The convergence result. The main result of this subsection can be stated as follows.

Proposition 3. *As h goes to 0, the value function V^h converges uniformly to the unique solution V to the free boundary problem:*

$$rV(w) = \begin{cases} \mu - (\rho - r)w + \rho w V'(w) + \dfrac{\lambda^2 \sigma^2}{2} V''(w) & \text{if } w \in [0, w^m], \\[2mm] rV(w^m) & \text{if } w \in (w^m, \infty), \end{cases} \qquad (33)$$

with boundary conditions:

$$V(0) = 0, \qquad (34)$$

$$V'(w^m) = 0, \qquad (35)$$

$$V''(w^m) = 0. \qquad (36)$$

The ordinary differential equation in (33) is the natural continuous-time counterpart of the Taylor-Young approximation (32). Condition (34) simply reflects that the discrete-time value functions V^h all vanish at 0 and converge uniformly to V. Condition (35) is a smooth pasting condition which expresses the fact that w^m can be interpreted as a dividend boundary in the continuous-time limit of the model, as discussed in Subsection 3.2 below. Condition (36) is a super contact condition which reflects the optimality of the dividend boundary w^m.

Our next result is a useful corollary of Proposition 3.

Corollary 1. $\lim_{h\downarrow 0} w^{h,m} = w^m.$

By (13)–(14), an immediate implication of this is that $\lim_{h\downarrow 0} w^{h,r} = \lim_{h\downarrow 0} w^{h,d} = w^m$. In particular, the region $(w^{h,d}, w^{h,m}]$ in which the entrepreneur receives positive transfers in case of success shrinks to $\{w^m\}$ as h goes to 0.

16

Outline of the proof. The proof of Proposition 3 proceeds in several steps. First, we show that the free boundary problem (33)–(36) has a unique solution.

Lemma 5. *There exists a unique solution V to (33)–(36).*

Next, we establish two important properties of V.

Lemma 6. *V is strictly increasing and strictly concave over $[0, w^m]$.*

We now come to the key part of the proof. For any period length $h > 0$, and for any bounded continuous function v defined over $\mathbb{R}_+$, let $T^{h,c}v$ be the incentive constrained maximum social surplus given a continuation value function v, assuming that stochastic liquidation is not feasible. The operator $T^{h,c}$ is more tractable than the Bellman operator T^h implicit in (8)–(11) because the constraints that enter its definition are linear, see equations (A.1)–(A.4) in the Appendix. Specifically, $T^{h,c}$ differs from T^h only in that x is constrained to be equal to 1. From (10), this implies that $T^{h,c}v$ is defined over $[p\lambda(c_+^h - c_-^h), \infty)$ only. Moreover, by construction, $T^{h,c}V^h = T^hV^h = V^h$ over $[w^{h,l}, \infty)$. We examine what happens when one applies the operator $T^{h,c}$ to the solution V to (33)–(36) for h close enough to 0. For this purpose, it will be convenient to denote by $\|\cdot\|_h$ the supremum norm over $[p\lambda(c_+^h - c_-^h), \infty)$. One has then the following result.

Lemma 7. $\lim_{h \downarrow 0} \|T^{h,c}V - V\|_h/h = 0.$

The proof relies on the differential operator (33) characterization of V, together with the smooth pasting and super contact conditions (35)–(36). The intuitive meaning of this result is that, when it is applied to V, the operator $T^{h,c}$ converges rapidly to the identity as h goes to 0. Alternatively, note that the operator $T^{0,c}$ obtained by letting $h = 0$ in (A.1)–(A.4) coincides with the identity on the subspace of concave functions defined over $\mathbb{R}_+$. Then, since V is itself concave by Lemma 6, what Lemma 7 asserts is that the derivative at $h = 0$ of the operator $T^{h,c}$, evaluated at the solution V to (33)–(36), is equal to 0:

$$\lim_{h \downarrow 0} \frac{\|(T^{h,c} - T^{0,c})V\|_h}{h} = 0.$$

Since $T^{h,c}V^h = V^h$ over $[w^{h,l}, \infty)$ by construction, and $\lim_{h \downarrow 0} w^{h,l} = 0$ by Lemma 4, this suggests that V^h converges to V as h goes to 0. This can be formally established as follows. Note that $V^h - V = T^{h,c}V^h - T^{h,c}V + T^{h,c}V - V$ over $[w^{h,l}, \infty)$. Since $w^{h,l} \geq p\lambda(c_+^h - c_-^h)$, one has, by definition of the norm $\|\cdot\|_h$:

$$\sup_{w \in [w^{h,l}, \infty)} \{|V^h(w) - V(w)|\} \leq \|T^{h,c}V^h - T^{h,c}V\|_h + \|T^{h,c}V - V\|_h$$

$$\leq \frac{\|V^h - V\|_0}{1 + rh} + \|T^{h,c}V - V\|_h,$$

(37)

where $\|\cdot\|_0$ is the standard supremum norm over $\mathbb{R}_+$, and the second inequality follows from a straightforward adaptation of Blackwell's theorem. We now use (37) to obtain an upper bound for $\|V^h - V\|_0$. A difficulty is that the left-hand side of (37) may be strictly smaller than $\|V^h - V\|_0$, as the function $|V^h - V|$ may reach its maximum over $\mathbb{R}_+$ in $[0, w^{h,l}]$. Two

cases must therefore be distinguished. Either $\|V^h - V\|_0 = \sup_{w \in [w^{h,l}, \infty)} \{|V^h(w) - V(w)|\}$ along a subsequence of h's that goes to 0. It then follows from (37) that:

$$\|V^h - V\|_0 \leq \left(\frac{1+rh}{r}\right)\frac{\|T^{h,c}V - V\|_h}{h}.$$

Lemma 7 then implies that $\|V^h - V\|_0$ goes to 0, and thus V^h converges to V uniformly along this subsequence. Alternatively, $\|V^h - V\|_0 = \max_{w \in (0, w^{h,l})} \{|V^h(w) - V(w)|\}$ along a subsequence of h's that goes to 0. It then follows that:

$$\|V^h - V\|_0 = |V(\tilde{w}^h) - \alpha^h \tilde{w}^h|,$$

where α^h is the constant slope of V^h over $(0, w^{h,l})$, and $\tilde{w}^h \in (0, w^{h,l})$ is given by the first-order condition $V'(\tilde{w}^h) = \alpha^h$. Since V' is bounded, the slopes α^h remain bounded. Lemma 4 then implies that $\|V^h - V\|_0$ goes to 0, and thus V^h converges to V uniformly along this subsequence. This completes the proof of Proposition 3.

Remark. It should be noted that the proof of Proposition 3 only makes use of the Bellman operator characterization of the discrete-time value functions, along with the differential operator characterization of the continuous-time value function. In particular, it relies on the discrete-time value functions only to the extent that they are defined as fixed points of a family of Bellman operators. As can be seen from the proof of Lemma 7, however, the application of the operator $T^{h,c}$ to the continuous-time value function V leads to thresholds similar to those used in the derivation of the discrete-time optimal contract.

3.2. Convergence of the optimal contracts

The results obtained so far are purely analytic. The objective of this subsection is to study their probabilistic counterpart, that is, the limit of the optimal contracts characterized in Proposition 1. We first examine the continuous-time limit of the cumulated cash-flow process. By (1), for any fixed period length $h > 0$, the total revenue generated by the project up to any date nh prior to liquidation is given by:

$$r_{nh}^h = \mu nh + \sigma \sum_{i=1}^{n} \varepsilon_i \sqrt{h}. \tag{38}$$

A continuous-time process $R^h = \{R_t^h\}_{t \geq 0}$ can be obtained from the discrete-time revenue process $\{r_{nh}^h\}_{n \in \mathbb{N}}$ given by (38) by setting $R_t^h = r_{[\![t/h]\!]h}^h$ for all $t \geq 0$, where $[\![t/h]\!]$ is the integer part of t/h, that is, the unique integer n such that $t \in [nh, (n+1)h)$. Accordingly, each sample path of R^h is a step function, taking the value r_{nh}^h when $t \in [nh, (n+1)h)$. It follows then from Jacod and Shiryaev (1987, Chapter IX, Theorem 4.21) that, as h goes to 0, this process converges in law to an arithmetic Brownian motion $R = \{R_t\}_{t \geq 0}$ with drift μ and volatility σ,

$$R_t = \mu t + \sigma Z_t \tag{39}$$

for all $t \geq 0$, where $Z = \{Z_t\}_{t \geq 0}$ is a standard Brownian motion. Similarly, for any fixed period length $h > 0$, a continuous-time process $W^h = \{W_t^h\}_{t \geq 0}$ can be obtained from the discrete-time utility process $\{w_{nh}^h\}_{n \in \mathbb{N}}$ by setting $W_t^h = w_{[\![t/h]\!]h}^h$ for all $t \geq 0$. Our objective is to study the limit in law of this process as h goes to 0.

18

The analysis proceeds in several steps. First, we provide a heuristic derivation of the continuous-time limit of the discrete-time utility process. Next, we establish an intuitive probabilistic representation of the continuous-time value function. Finally, we state and discuss our main convergence result.

A heuristic derivation. We use the results of Subsections 2.3 and 3.1 to identify the limit of the utility process W^h as h goes to 0. Note first that since the value function V^h converges uniformly to V, the initial value W_0^h of the process W^h converges to a well-defined limit W_0, holding the bargaining power of the financiers and that of the entrepreneur fixed as h goes to 0. Since there is no loss of generality in assuming that $W_0^h \leq w^{h,r}$ for all $h > 0$ as the utility of the entrepreneur enters the interval $[0, w^{h,r}]$ after at most one period and never leaves it afterwards, and since $\lim_{h \downarrow 0} w^{h,r} = w^m$ by Corollary 1, one can assume that $W_0 \in [0, w^m]$.[8] Next, by Proposition 1, the utility of the entrepreneur is reflected at $w^{h,r}$, and it grows at an expected rate ρh and with constant volatility $\lambda \sigma \sqrt{h}$ over the interval $(w^{h,l}, w^{h,r})$. Moreover, by Lemma 4, the liquidation region $[0, w^{h,l})$ shrinks to $\{0\}$ as h goes to 0. Finally, by Corollary 1, the region $(w^{h,d}, w^{h,m}]$ in which the entrepreneur receives positive transfers in case of success shrinks to $\{w^m\}$ as h goes to 0. This suggests that, in the continuous-time limit, the utility of the entrepreneur follows a diffusion process that is reflected at w^m and killed at 0, and that grows at an expected rate ρ and with constant volatility $\lambda \sigma$ over the interval $(0, w^m)$.

To establish this result, we rely on tools from the theory of reflected diffusion processes. We first need an adequate formalization of the candidate limit process. This can be achieved as follows. Given the Brownian motion $Z = \{Z_t\}_{t \geq 0}$, consider the so called Skorokhod problem of finding a continuous adapted process $W = \{W_t\}_{t \geq 0}$ with initial value W_0 and a non-decreasing adapted process $L = \{L_t\}_{t \geq 0}$ with initial value 0 that jointly satisfy:

$$dW_t = \rho W_t dt + \lambda \sigma dZ_t - dL_t, \tag{40}$$

$$W_t \leq w^m, \tag{41}$$

$$L_t = \int_0^t 1_{\{W_s = w^m\}} dL_s \tag{42}$$

for all $t \in [0, \tau]$, where $\tau = \inf\{t \geq 0 \,|\, W_t = 0\}$, and $W_t = 0$ for all $t \geq \tau$. Condition (42) requires that L increases only when W hits the boundary w^m, while (40)–(41) express the fact that this causes W to be reflected back at w^m. From Tanaka (1979, Theorem 4.1), there exists a pathwise unique solution (W, L) to (40)–(42).

A probabilistic representation. The above heuristic derivation suggests that W is the continuous-time limit of the entrepreneur's utility process, and therefore that L represents the entrepreneur's continuous-time cumulative payment process. In line with this conjecture, the following representation result holds.

Proposition 4. *Let $F(w) = V(w) - w$ denote the financiers' utility given a promised utility*

---

[8]Alternatively, one could allow for $W_0 > w^m$, and let the continuous-time limit of the discrete-time optimal contract start with a transfer to the entrepreneur bringing back her utility to w^m.

w for the entrepreneur in the continuous-time limit of the model. Then, for any $w \in [0, w^m]$,

$$w = \mathbb{E}^{(w,0)}\left[\int_0^\tau e^{-\rho t} dL_t\right], \tag{43}$$

$$F(w) = \mathbb{E}^{(w,0)}\left[\int_0^\tau e^{-rt}(\mu dt - dL_t)\right], \tag{44}$$

where $\mathbb{E}^{(w,0)}$ is the expectation operator induced by the process (W, L) starting at $(w, 0)$.

This result is a simple consequence of Itô's formula, together with the properties (33)–(35) of V and the dynamics (40)–(42) of W. The interpretation of (43) is that the utility of the entrepreneur is equal to the present value of the payments she receives, discounted at rate ρ. Similarly, (44) says that the utility of the financiers is equal to the present value of the cash-flows from the project, less the payments to the entrepreneur, discounted at rate r.

The structure of payments is then intuitively as follows. As long as W stays in $[0, w^m)$, the entrepreneur receives no payments. By contrast, whenever W hits w^m, the entrepreneur receives an immediate payment dL, and W bounces back in $[0, w^m)$. Therefore w^m can be interpreted as a dividend boundary in the continuous-time limit of the model. The smooth pasting condition (35), which plays a key role in the proof of (44), accordingly expresses the fact that the marginal cost of compensating the entrepreneur at w^m is just equal to the marginal cost of an immediate payment, that is $F'(w^m) = -1$. Finally, the project is liquidated at the first time τ at which W hits 0, and the continuation utilities are then 0 for both the entrepreneur and the financiers.

The convergence result. The main result of this subsection can be stated as follows.

Proposition 5. *As h goes to 0, the process W^h converges in law to the process W.*

This result strengthens the conclusions of Proposition 3 by establishing that the discrete-time optimal contract, as summarized by the utility process of the entrepreneur, converges to a well-defined continuous-time limit as h goes to 0. This justifies rigorously our interpretation of the representation (43)–(44).

The proof of Proposition 5 proceeds in three steps. Leaving aside the issues raised by the potential liquidation of the project, we first study the stochastic processes $\tilde{W}^h$ and $\tilde{W}$ generated by (23)–(24) and (40)–(42) over the whole intervals $(-\infty, w^{h,r}]$ and $(-\infty, w^m]$. A key observation is that the discounted version $\tilde{W}^{h,\rho}$ of $\tilde{W}^h$ solves a discretized and discounted version of the Skorokhod problem (40)–(42). One can then rely on stability properties of this problem (Słomiński (1993, Proposition 4)) to establish that, as h goes to 0, $\tilde{W}^{h,\rho}$ converges in law to the discounted version $\tilde{W}^\rho$ of $\tilde{W}$. In the second step of the proof, we argue that since $\lim_{h\downarrow 0} w^{h,l} = 0$, this convergence result is preserved if one kills $\tilde{W}^h$ and $\tilde{W}$, respectively at the first times at which $\tilde{W}^h < w^{h,l}$ and $\tilde{W} \leq 0$. Note that the process $\tilde{W}$ killed at 0 coincides with W. The third step of the proof deals with the liquidation region $[0, w^{h,l})$, and examines the relationship between the discounted version $W^{h,\rho}$ of W^h and the process $\tilde{W}^{h,\rho}$ killed at the first time at which $\tilde{W}^h < w^{h,l}$. By construction, the processes $W^{h,\rho}$ and $\tilde{W}^{h,\rho}$ coincide up to the point where they reach the liquidation region. Using again the fact that $\lim_{h\downarrow 0} w^{h,l} = 0$, we show that the distance between the paths of $W^{h,\rho}$ and those of the process $\tilde{W}^{h,\rho}$ killed at the first time at which $\tilde{W}^h < w^{h,l}$ converges to 0 in probability as h goes to 0. Combining this with the first step of the proof implies the result.

3.3. Comparison with DeMarzo and Sannikov (2006)

In contemporaneous work, DeMarzo and Sannikov (2006) study a continuous-time version of the cash-flow diversion model, in which cumulative cash-flows follow a Brownian motion with drift as in (39). In line with Sannikov (2003), they use martingale techniques to derive the optimal contract. Their continuous-time characterization of the financiers' value function and of the entrepreneur's utility process matches our characterization of the continuous-time limit of the discrete-time optimal contracts. Propositions 3 and 5 therefore show that their continuous-time analysis emerges as the limit of a discrete-time cash-flow diversion model in which cash-flows follow the binomial process defined by (1)–(2).

4. DYNAMIC PRICING OF SECURITIES

In this section, we investigate the implications of our continuous-time analysis for the pricing of the securities used to implement the optimal contract.

4.1. Security prices

We first describe the continuous-time limit of the implementation of the discrete-time optimal contract proposed in Subsection 2.5. For a short period length $h > 0$, this implementation generates a stochastic process $\{\hat{m}_{nh}^h\}_{n\in\mathbb{N}}$ for size-adjusted cash reserves that is designed in such a way that $\hat{m}_{nh}^h = \hat{w}_{nh}^h/\lambda^h$ at each date nh. The cash reserves held by the firm on its bank account at the beginning of period n are given by $m_{nh}^h = (\prod_{i=0}^{n-1} \hat{x}_{ih}^h)\hat{m}_{nh}^h$. A continuous-time process $M^h = \{M_t^h\}_{t\geq 0}$ can be obtained from the discrete-time cash reserves process $\{m_{nh}^h\}_{n\in\mathbb{N}}$ by setting $M_t^h = m_{\llbracket t/h \rrbracket h}^h$ for all $t \geq 0$. The processes W^h and M^h are not obviously related to each other, since the former derives from the probabilistic interpretation of liquidation, while the latter derives from the downsizing interpretation of liquidation. However, the following corollary of Proposition 5 shows that M^h and W^h/λ^h have the same limit $M = W/\lambda$ as h goes to 0, implying that the two interpretations of liquidation coincide in the continuous-time limit of the model. Intuitively, this is because the liquidation region $[0, w^{h,l})$ shrinks to $\{0\}$ as h goes to 0.

Corollary 2. *As h goes to 0, the process M^h converges in law to the process M.*

Since $M = W/\lambda$, it follows from (40) that:

$$dM_t = \rho M_t dt + \sigma dZ_t - \frac{1}{\lambda} dL_t \tag{45}$$

for all $t \in [0, \tau]$, and $M_t = 0$ for all $t \geq \tau$. Note that $\tau = \inf\{t \geq 0 \,|\, M_t \leq 0\}$ by construction, so that the firm is liquidated as soon as it runs out of cash. This is an endogenous feature of the implementation of the optimal contract, rather than an exogenous and possibly suboptimal restriction, as in the dividend distribution models of Jeanblanc-Picqué and Shiryaev (1995) or Radner and Shepp (1996). From (41), the cash reserves are bounded above by w^m/λ at all times. One can alternatively represent the dynamics of the cash reserves by:

$$dM_t = (rM_t + \mu)dt + \sigma dZ_t - dL_t - dP_t \tag{46}$$

for all $t \in [0, \tau]$, where dL is the payment to the entrepreneur and dP is the payment to the

financiers. It follows then from (45)–(46) that:

$$dP_t = [\mu - (\rho - r)M_t]dt + \frac{1-\lambda}{\lambda}\,dL_t \qquad (47)$$

for all $t \in [0,\tau]$. Equations (45) and (47) provide the continuous-time analogue of our discrete-time implementation. There are two types of securities, namely stocks and bonds. As in the discrete-time implementation, the entrepreneur is granted a non-tradeable fraction λ of the stocks. Stocks pay out only when cash reserves reach a threshold, while bonds pay out a continuous stream of cash-flows. Specifically, in analogy with (30), stocks distribute a total dividend dL/λ when the cash reserves M reach the threshold w^m/λ. A fraction λ of the dividend goes to the entrepreneur, who therefore receives a payment dL, while the remaining fraction $1-\lambda$ of the dividend goes to the financiers. Next, in analogy with (31), bonds continuously distribute to the financiers a coupon $\mu - (\rho-r)M$, that varies with the level M of cash reserves. Since M is bounded above by w^m/λ, one must have $\mu - (\rho-r)w^m/\lambda \geq 0$ in order to preserve limited liability. As can be seen from (33)–(36), this is the case when λ is close enough to 1.

We now derive continuous-time pricing formulas for these securities. A key implication of our implementation is that, at any date, stock and bond prices are deterministic functions of the current level of cash reserves. Specifically, let S_t be the market value of stocks at date t. Since financiers are risk-neutral, for each $t \in [0,\tau]$ this is the present value of future dividends, discounted at rate r:

$$S_t = \mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}\frac{1}{\lambda}\,dL_s\right], \qquad (48)$$

where $\mathbb{E}_t$ is the expectation operator conditional on information up to date t. Using Itô's formula as in the proof of Proposition 4, together with the dynamics (45) of M, it is easy to check that $S_t = \mathcal{S}(M_t)$ for all $t \geq 0$, where $\mathcal{S}$ is the solution over $[0, w^m/\lambda]$ to the boundary value problem:

$$r\mathcal{S}(m) = \rho m \mathcal{S}'(m) + \frac{\sigma^2}{2}\,\mathcal{S}''(m), \qquad (49)$$

$$\mathcal{S}(0) = 0, \qquad (50)$$

$$\mathcal{S}'\left(\frac{w^m}{\lambda}\right) = 1. \qquad (51)$$

Note that condition (51) simply expresses the fact that w^m/λ is the threshold value for cash reserves at which stocks distribute a dividend. When this boundary is reached, each additional revenue generated by the firm is earned as a dividend by the shareholders.

Remark. By construction, the instantaneous expected return on stocks is equal to the market rate r, which is less than the entrepreneur's discount rate ρ. Hence, if her share of the stocks were tradeable, the entrepreneur would like to sell it to the financiers. This is precluded in the optimal contract to preserve incentive compatibility.

Similarly, let D_t be the market value of bonds at date t. For each $t \in [0,\tau]$ this is the

present value of future coupons, discounted at rate r:

$$D_t = \mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}[\mu - (\rho - r)M_s]\,ds\right]. \tag{52}$$

Proceeding as for stock prices, it is easy to check that $D_t = \mathcal{D}(M_t)$ for all $t \geq 0$, where $\mathcal{D}$ is the solution over $[0, w^m/\lambda]$ to the boundary value problem:

$$r\mathcal{D}(m) = \mu - (\rho - r)m + \rho m \mathcal{D}'(m) + \frac{\sigma^2}{2}\mathcal{D}''(m), \tag{53}$$

$$\mathcal{D}(0) = 0, \tag{54}$$

$$\mathcal{D}'\left(\frac{w^m}{\lambda}\right) = 0. \tag{55}$$

We denote by $S = \{S_t\}_{t\geq 0}$ and $D = \{D_t\}_{t\geq 0}$ the stochastic processes of stock and bond prices. The processes W, M, S and D are related as follows.

Proposition 6. *At any date $t \geq 0$, the following holds:*

$$F(W_t) + M_t = (1 - \lambda)S_t + D_t. \tag{56}$$

The right-hand side of (56) is the market value of the securities held by the financiers, that is, the present value of the cash-flows these securities will distribute. The left-hand side of (56) corresponds to the assets generating these cash-flows. The first term, $F(W_t)$, is the present value of the fraction of the operating cash-flows allocated to the financiers. The second term, M_t, is the level of cash reserves, which generate interest income and correspond to the proceeds that would accrue to the bondholders if the firm were exogenously liquidated at time t. At date 0, (56) states that the value of the securities allocated to the financiers, $(1 - \lambda)S_0 + D_0$, is equal to the proceeds from the issuance, $I - A + M_0$, plus the initial rent of the financiers, $F(W_0) - (I - A)$.

In the remainder of this section, we use the characterizations (49)–(51) and (53)–(55) of stock and bond prices to derive empirical implications and comparative statics results.

4.2. Empirical implications

Stock price. Our first result is a direct implication of (49)–(51).

Proposition 7. *The stock price is a strictly increasing and strictly concave function of the level of cash reserves.*

The fact that S increases in M is the result of two effects. First, an increase in cash reserves makes dividend distribution more likely. Second, it reduces the risk of default. The combination of these two effects leads to an increase in the value of the stock. The fact that S is a concave function of M is less immediate. Intuitively, the stock price reacts less to changes in the level of cash reserves when the firm has established a sufficient performance record. This is because the larger the cash reserves accumulated by the firm, the less likely default will occur in the near future, see below. By contrast, after a sequence of low cash-flow

23

realizations, cash reserves are low, and default risk is high. As a result of this, the stock price reacts strongly to firm performance and ensuing changes in the level of cash reserves. Note that since $\mathcal{S}'(w^m/\lambda) = 1$ and $\mathcal{S}$ is a strictly concave function that vanishes at 0, $\mathcal{S}(M) > M$ as long as $M > 0$. Thus the market value of stocks exceeds the level of cash reserves at any date prior to liquidation.

These results have several important implications for the dynamics of stock prices. Using the characterization (49)–(51) of the function $\mathcal{S}$ together with Itô's formula, we obtain that:

$$dS_t = rS_t dt + S_t \sigma^S(S_t) dZ_t - \frac{1}{\lambda} dL_t \qquad (57)$$

for all $t \in [0, \tau]$, where for each $s \in [0, \mathcal{S}(w^m/\lambda)]$, $\sigma^S(s) = \sigma\mathcal{S}'(\mathcal{S}^{-1}(s))/s$ represents the volatility of the return on stocks. Together with (57), Proposition 7 implies that the dynamics of the stock price S differs in three important ways from that postulated by Black and Scholes (1973). First, the stock price is reflected back each time dividends are distributed, which occurs when S hits $\mathcal{S}(w^m/\lambda)$. Next, since the volatility $S\sigma^S(S) = \sigma\mathcal{S}'(\mathcal{S}^{-1}(S))$ of S is bounded away from 0, the stock price can attain 0 with positive probability, which happens when the firm is liquidated. Last, since the function $\mathcal{S}$ is strictly increasing and strictly concave, $S\sigma^S(S)$ and a fortiori $\sigma^S(S)$ are decreasing functions of S, so that changes in the volatility of the stock price and in the volatility of the stock return are negatively correlated with stock price movements. This is in line with the leverage effect documented by Black (1976) and Nelson (1991): volatility tends to rise in response to bad news, and to fall in response to good news. Another implication is that stock prices are more volatile than cash-flows, $S\sigma^S(S) \geq \sigma$, with equality in the limit only at the dividend boundary $\mathcal{S}(w^m/\lambda)$.

Leverage ratio. Our implementation also enables us to study how the financial structure of the firm is affected by the level of cash reserves, and hence varies with the performance of the firm and the price of the stock. At any date $t \in [0, \tau]$, the leverage ratio expressed in market values is D_t/S_t. Then the following holds.

Proposition 8. *The leverage ratio expressed in market values is a strictly decreasing function of the level of cash reserves, and hence of the stock price.*

This result reflects the fact that, when the firm is consistently successful and cash reserves increase, the value of stocks increases more than that of bonds. Intuitively, this is because stocks pay out dividends only after significant accumulated performance, while bonds are relatively safer, and distribute a coupon even when performance is weaker. Hence the value of stocks is more responsive to the performance of the firm than the value of bonds. Proposition 8 implies that performance shocks and stock price movements induce persistent changes in the leverage of the firm. Such persistent changes have been documented by recent empirical analyzes. In discussing these findings, Welch (2004) questions why firms do not issue or repurchase debt or equity to counterbalance the impact of stock price movements on their capital structure. This phenomenon is indeed puzzling if one relies upon a static model, such as for instance the trade-off theory, according to which there exists an optimal leverage ratio to which the firm should endeavor to revert. By contrast, in our implementation of the optimal contract, the financial structure of the firm adjusts optimally through the change in the market value of securities, without requiring further issuing activities.

Book-to-market ratio. The empirical finance literature has emphasized the correlation between the book-to-market ratio and key economic variables such as returns. Our model can be used to generate implications for the joint dynamics of the book and market values of balance sheet items. Consider first the balance sheet of the firm at date 0. On the liability side, one finds debt and equity, accounted at their market values D_0 and S_0. On the asset side, one finds the initial cash reserves M_0, the book value I of property, plant and equipment, and an intangible asset G_0, the goodwill, reflecting the net value of the project.[9] Thus, we have the accounting identity:

$$D_0 + S_0 = I + M_0 + G_0. \tag{58}$$

From (56) and (58), it is easy to check that the goodwill is equal to the present value of the entrepreneur's compensation, discounted at rate r, plus the financiers' rent, less the financial contribution of the entrepreneur to the project:

$$G_0 = \lambda S_0 + F(W_0) - (I - A) - A.$$

Now consider the balance sheet at time $t > 0$. The only item that has changed on the asset side of the balance sheet is the level M_t of cash reserves. On the liability side, debt is still accounted for at its historical value. Consequently, the book value of equity is:

$$I + M_t + G_0 - D_0 = S_0 + M_t - M_0. \tag{59}$$

The first term on the right-hand side of (59) is the par value of the stock, equal to its initial book value. The second term reflects accumulated retained earnings. The ratio of the book value of stocks to their market value, that is, the book-to-market ratio, is thus given by:

$$B_t = \frac{S_0 + M_t - M_0}{S_t}. \tag{60}$$

From (60) and the characterization (49)–(51) of the stock price, the book-to-market ratio is a deterministic function of the cash reserves, which can be characterized as follows.

Proposition 9. *If $M_0 < w^m/\lambda$, the book-to-market ratio is a U-shaped function of the level of cash reserves, and hence of the stock price.*

The book-to-market ratio goes to infinity when cash reserves go to 0, since the market value of stocks then goes to 0, while their book value is bounded below by $S_0 - M_0$, which is strictly positive since the market value of stocks exceeds the level of cash reserves at any date prior to liquidation. As a result of this, the book-to-market ratio is decreasing at 0. By contrast, the increase in the book-to-market ratio occurring when cash reserves rise towards w^m/λ reflects the concavity of $\mathcal{S}$. Proposition 9 generates the following testable implications. For financially distressed or recently established firms with low cash reserves, the correlation between stock returns and contemporaneous changes in the book-to-market ratio should be

---

[9]In accounting terms, this can be justified as follows. Initially, the entrepreneur owns a venture consisting of a tangible asset corresponding to her initial wealth A, and an intangible asset corresponding to the investment project that only she can manage. At date 0, this entity is partially acquired by the financiers, who contribute $I - A$ and M_0. At this point, the investment I is sunk. The implementation of the optimal contract determines the financial structure of the new entity thus created. At date 0, its tangible assets are M_0 and I. The assets also include the goodwill, which is an intangible asset reflecting the economic value of the investment project. For more details, see Brealey and Myers (2000, Section 33.4).

negative. By contrast, for healthy or mature firms with high cash reserves, stock returns and contemporaneous changes in the book-to-market ratio should be positively correlated. As far as we know, this is a new implication, which has not been tested yet.[10]

Default risk. As a measure of the risk of default at time $t \in [0, \tau)$, we take the credit yield spread Δ_t on a consol bond that pays one unit of account at each date until the time of default. For any $t \in [0, \tau)$, Δ_t is defined by the following formula:

$$\int_t^\infty e^{-(r+\Delta_t)(s-t)} \, ds = \mathbb{E}_t \left[\int_t^\tau e^{-r(s-t)} \, ds \right]. \tag{61}$$

Rearranging (61), we obtain that:

$$\Delta_t = \frac{rT_t}{1 - T_t}, \tag{62}$$

where $T_t = \mathbb{E}_t[e^{-r(\tau-t)}]$ for all $t \in [0, \tau]$. Using Itô's formula as in the proof of Proposition 4, together with the dynamics (45) of M, it is easy to check that $T_t = \mathcal{T}(M_t)$ for all $t \geq 0$, where $\mathcal{T}$ is the solution over $[0, w^m/\lambda]$ to the boundary value problem:

$$r\mathcal{T}(m) = \rho m \mathcal{T}'(m) + \frac{\sigma^2}{2} \mathcal{T}''(m), \tag{63}$$

$$\mathcal{T}(0) = 1, \tag{64}$$

$$\mathcal{T}'\left(\frac{w^m}{\lambda}\right) = 0. \tag{65}$$

One has then the following result.

Proposition 10. *The credit yield spread is a strictly positive, strictly decreasing and strictly convex function of the level of cash reserves.*

The fact that an increase in the cash reserves reduces the credit yield spread simply reflects that default occurs when the firm runs out of cash. The intuition for the convexity of the credit yield spread with respect to the level of cash reserves can then be grasped by observing that, when M reaches the dividend threshold w^m/λ, cash inflows are distributed as dividends and thus do not reduce the credit yield spread, as can be seen from (62) and (65). By contrast, changes in the level of cash reserves have a large impact on the credit yield spread when the firm is close to default.

4.3. Comparative statics

Initial investment and liquidity. We first discuss some comparative statics properties of the initial balance sheet. For simplicity, the entrepreneur is assumed to have all the bargaining power at date 0, that is, financial markets are perfectly competitive.

Proposition 11. *The investment capacity of the firm and the initial level of the cash reserves are strictly increasing with respect to μ, and strictly decreasing with respect to λ and σ.*

---

[10]Indeed, the empirical asset pricing literature has primarily focused on the relationship between the value of the book-to-market ratio at a given point in time and subsequent returns.

A high value of μ means that the investment project is potentially very profitable, while high values of λ and σ imply that the agency problem is severe and that the operating risk is high. Proposition 11 is in line with static corporate finance models, which typically predict that the investment capacity of the firm increases with respect to the profitability of the project, and decreases with respect to the magnitude of agency costs and the operating risk (Innes (1990), Holmström and Tirole (1997)). A further implication of Proposition 11 is that firms facing severe agency problems have lower cash reserves. In order to test this empirical prediction, one needs to find a proxy for λ. To this end, one may for instance rely on the measure of internal risk developed by Kaplan and Strömberg (2004). In addition, the ability to divert cash-flows is likely to be greater in more innovative firms, or firms with high immaterial expenses, such as R&D or marketing, where operating cash-flows are more difficult to benchmark and evaluate.

Default risk. We next study how the credit yield spread is affected by changes in the structural parameters of the model.

Proposition 12. *For any strictly positive level of cash reserves, the credit yield spread is strictly decreasing with respect to μ, and strictly increasing with respect to λ.*

The impact of the parameters λ and μ on the credit yield spread runs through their impact on the dividend threshold w^m/λ. As λ increases, the entrepreneur's temptation to divert cash-flows is greater, and the incentive problem becomes more severe. To make reporting high cash-flows realizations more attractive for the entrepreneur, it is optimal to lower the threshold level w^m/λ of cash reserves at which she receives dividends from her share of the stocks. This has a cost, however, since increasing the frequency of dividend distributions depletes the cash reserves of the firm. This in turn increases the risk that the firm will run out of cash and therefore the risk of default. As a result, the credit yield spread increases. By contrast, an increase in μ raises the total cash-flows that can be promised to all parties, including the entrepreneur, which relaxes the agency problem. One can then afford to raise the dividend threshold, in order to reduce the risk of default and the resulting social cost of liquidation.

Stock prices. Further comparative statics results can be drawn from the characterization (49)–(51) of the stock price.

Proposition 13. *For any strictly positive level of cash reserves, the stock price is strictly increasing with respect to μ, and strictly decreasing with respect to λ.*

This result implies that, controlling for the level of cash reserves, the market value of stocks decreases with the magnitude of the agency problem. The intuition is that, when the incentive problem becomes more severe, the risk of default increases, which reduces the present value of future dividends, that is, the stock price. In addition, initial cash reserves are lower, which further depresses stock prices. Our analysis thus predicts that, all other things being equal, firms facing more severe incentive problems should have lower stock price valuations. An empirical implication of this is that the magnitude of agency costs should be negatively correlated with price earning ratios. This result also sheds light on the link between insider ownership and stock price performance, which has been a longstanding question of interest in corporate finance. Our theoretical analysis underlines that both these variables

are endogenous and jointly reflect the intensity of the agency problem as captured by our structural parameter λ, in line with the empirical studies of Himmelberg, Hubbard and Palia (1999), Demsetz and Villalonga (2001), and Coles, Lemmon and Meschke (2006).

5. CONCLUSION

This paper develops a dynamic model of the firm in which capital structure matters because of internal agency costs. A simple incentive problem suffices to rationalize a number of appealing features of the firm's balance sheet. At any given date, payout and liquidation decisions depend solely on the current amount of liquidities accumulated by the firm. A key feature of the optimal long-term contract is that dividends are paid out when cash reserves hit a given threshold, while the firm is closed when illiquid. As a consequence, shocks on the firm's solvency and liquidity ratios are highly persistent.

We characterize the continuous-time limit of the optimal financial contract. This enables us to derive several testable implications for the market valuation of optimal securities. Stock prices are more volatile than earnings, and unlike the latter, have a stochastic volatility that increases after stock price drops. Leverage is negatively related to the past performance of the firm, which also determines the sign of the correlation between stock returns and changes in the book-to-market ratio. Finally, more severe agency problems result in lower price earning ratios and firm liquidity, and in higher credit yield spreads.

We hope that this analysis can serve as a first step to bridge the gap between asset pricing and corporate finance. An important extension, which we leave for further research, would be to depart from the assumption that financiers are risk-neutral, and to study the relationship between required expected returns, market risk, and frictions within the corporation.

APPENDIX A

Proof of Lemma 1. It is immediate from (5)–(6) that the function V^h satisfies the Bellman equation (8) for all $w \geq 0$, subject to constraints (4) and (6)–(7). We now eliminate u_+ and u_- from the constraints. Given a triple $(x, w_+, w_-) \in [0,1] \times \mathbb{R}_+^2$, one can find a pair (u_+, u_-) such that (4) and (6)–(7) hold if and only if the set of $(u_+, u_-) \in \mathbb{R}_+^2$ such that:

$$u_+ - u_- \geq \lambda(c_+^h - c_-^h) - \frac{w_+ - w_-}{1 + \rho h},$$

$$\frac{w}{x} - \frac{pw_+ + (1-p)w_-}{1 + \rho h} \geq pu_+ + (1-p)u_-$$

is non-empty. The second of these constraints immediately implies (11). If the right-hand side of the first of these constraints is negative, then $u_+ = u_- = 0$ is obviously a feasible choice. In the opposite case in which this term is non-negative, the feasible set is not empty if and only if $(u_+, u_-) = (\lambda(c_+^h - c_-^h) - (w_+ - w_-)/(1 + \rho h), 0)$ is feasible, which is equivalent to:

$$\frac{w}{x} - \frac{pw_+ + (1-p)w_-}{1 + \rho h} \geq p \left[\lambda(c_+^h - c_-^h) - \frac{w_+ - w_-}{1 + \rho h} \right],$$

which, after simplifications, yields (10). The result follows. ∥

Proof of Lemma 2. Let T^h be the Bellman operator associated to (8)–(11), and consider a bounded continuous function $v \in C_b(\mathbb{R}_+)$. Note that the mapping $w \mapsto v(w) - (\rho - r)hw/(1 + \rho h)$ goes to $-\infty$

at ∞ since $v \in C_b(\mathbb{R}_+)$ and $\rho > r$. Accordingly, let M_v^h be the maximum value of this function, and $w_v^{h,r}$ the smallest point at which it reaches its maximum. Clearly $T^h v \geq 0$, so $T^h v$ is bounded below. Similarly, $T^h v \leq \mu h + M_v^h/(1 + rh)$, so $T^h v$ is bounded above. For any value of w, there is no loss of generality in restricting w_+ and w_- to be in $[0, w_v^{h,r}]$. Hence Berge's maximum theorem applies, and thus T^h maps $C_b(\mathbb{R}_+)$ into itself. Using Blackwell's theorem, it is immediate to check that T^h is a contraction, and thus it has a unique fixed point $V^h \in C_b(\mathbb{R}_+)$ by the contraction mapping theorem.

We now prove that V^h is non-decreasing and concave. Since the set of continuous, bounded, non-decreasing and concave functions that vanish at 0 is a closed subset of $C_b(\mathbb{R}_+)$, it is sufficient to prove that T^h maps this set into itself. Specifically, let v be such a function, and let $w' \geq w \geq 0$. Assume that (x, w_+, w_-) is an optimal choice in the program that defines $T^h v(w)$. Since $w' \geq w$, it follows from (9)–(11) that (x, w_+, w_-) is a feasible choice in the program that defines $T^h v(w')$, which yields the same value as $T^h v(w)$. Hence $T^h v(w') \geq T^h v(w)$, and $T^h v$ is non-decreasing. That $T^h v$ vanishes at 0 follows directly from (8) together with the fact that $w = 0$ implies $x = 0$ because of (10). We now prove that $T^h v$ is concave. In line with Clementi and Hopenhayn (2006), we decompose (8)–(11) into two subproblems. First, we consider the problem of maximizing the expected social surplus conditional on not liquidating the project:

$$T^{h,c} v(w) = \max \left\{ \mu h + \frac{p v(w_+) + (1-p) v(w_-)}{1 + rh} - \frac{(\rho - r) h [p w_+ + (1-p) w_-]}{(1+rh)(1+\rho h)} \right\} \qquad \text{(A.1)}$$

for all $w \geq p\lambda(c_+^h - c_-^h)$, subject to:

$$(w_+, w_-) \in \mathbb{R}_+^2, \qquad \text{(A.2)}$$

$$w \geq \frac{w_-}{1 + \rho h} + p\lambda(c_+^h - c_-^h), \qquad \text{(A.3)}$$

$$w \geq \frac{p w_+ + (1-p) w_-}{1 + \rho h}. \qquad \text{(A.4)}$$

The value function $T^h v$, taking into account the possibility of liquidation, is then given by:

$$T^h v(w) = \max \{ x T^{h,c} v(w^{h,c}) \} \qquad \text{(A.5)}$$

for all $w \geq 0$, subject to:

$$(x, w^{h,c}) \in [0,1] \times [p\lambda(c_+^h - c_-^h), \infty) \qquad \text{(A.6)}$$

and:

$$w = x w^{h,c}. \qquad \text{(A.7)}$$

By the same argument used to show that $T^h v$ is non-decreasing over $\mathbb{R}_+$, it follows that $T^{h,c} v$ is non-decreasing over $[p\lambda(c_+^h - c_-^h), \infty)$. Moreover, it is concave over this interval. To check this, let $w \geq w' \geq p\lambda(c_+^h - c_-^h)$, $\theta \in [0,1]$ and $w_\theta = \theta w + (1-\theta)w'$. Assume that (w_+, w_-) is optimal in the program that defines $T^{h,c} v(w)$, that (w'_+, w'_-) is optimal in the program that defines $T^{h,c} v(w')$, and define $(w_{\theta+}, w_{\theta-}) = \theta(w_+, w_-) + (1-\theta)(w'_+, w'_-)$. Since the constraints (A.2)–(A.4) are linear, it follows that $(w_{\theta+}, w_{\theta-})$ is a feasible choice in the program that defines $T^{h,c} v(w_\theta)$. Hence, since v is concave, one has $T^{h,c} v(w_\theta) \geq \theta T^{h,c} v(w) + (1-\theta) T^{h,c} v(w')$, and thus $T^{h,c} v$ is concave. By definition, $w_v^{h,r}$ is the smallest point at which the mapping $w \mapsto v(w) - (\rho - r)hw/(1+\rho h)$ reaches its maximum. Let $w_v^{h,m} = w_v^{h,r}/(1 + \rho h) + p\lambda(c_+^h - c_-^h)$. Then, for any $w \geq w_v^{h,m}$, $(w_v^{h,r}, w_v^{h,r})$ is a feasible choice in the program that defines $T^{h,c} v(w)$, and it yields the maximum utility in (A.1). As a result of this,

$T^{h,c}v$ is constant over $[w_v^{h,m}, \infty)$. Next, (A.5)–(A.7) can be rewritten as:

$$T^h v(w) = \max\left\{\frac{T^{h,c}v(w^{h,c})w}{w^{h,c}}\right\}$$

for all $w \geq 0$, subject to:

$$w^{h,c} \geq \max\{p\lambda(c_+^h - c_-^h), w\}.$$

Since $T^{h,c}v$ is continuous over $[p\lambda(c_+^h - c_-^h), \infty)$ and constant over $[w_v^{h,m}, \infty)$, the mapping $w^{h,c} \mapsto T^{h,c}v(w^{h,c})/w^{h,c}$ reaches its maximum in $[p\lambda(c_+^h - c_-^h), w_v^{h,m}]$. Moreover, since $T^{h,c}v$ is concave, the set of points at which this mapping reaches its maximum is an interval $[\underline{w}_v^{h,c}, \overline{w}_v^{h,c}]$, possibly reduced to a point. It follows that $T^h v$ is given by:

$$T^h v(w) = \begin{cases} \dfrac{T^{h,c}v(\overline{w}_v^{h,c})w}{\overline{w}_v^{h,c}} & \text{if } w \leq \overline{w}_v^{h,c}, \\[3mm] T^{h,c}v(w) & \text{if } w > \overline{w}_v^{h,c}, \end{cases}$$

which implies that $T^h v$ is concave. We let $w^{h,l} = \overline{w}_{V^h}^{h,c}$, $w^{h,r} = w_{V^h}^{h,r}$ and $w^{h,m} = w_{V^h}^{h,m}$. Observe that V^h satisfies items (i)–(ii) of Lemma 2 by construction. To conclude the proof, we need only to check that V^h is strictly increasing over $[0, w^{h,m}]$. Suppose by way of contradiction that $V^h(w) = V^h(w')$ for some $w < w' < w^{h,m}$. Then $V^h(w) = V^h(w') = V^h(w^{h,m})$ by concavity of V^h. Therefore, the optimal choice in (8)–(11) given w must be $(1, w^{h,r}, w^{h,r})$. However, because $w < w^{h,m}$ and $w^{h,m} = w^{h,r}/(1 + \rho h) + p\lambda(c_+^h - c_-^h)$, this choice violates (10), a contradiction. Hence the result. ‖

Proof of Lemma 3. To prove that $w^{h,r} < w^{h,m}$ for h close enough to 0, one may be tempted to argue that because $V^{h\prime}(w^{h,r}) = (\rho - r)h/(1 + \rho h) > 0 = V^{h\prime}(w^{h,m})$, one has $w^{h,r} < w^{h,m}$ by concavity of V^h. This presumes, however, that V^h is differentiable at $w^{h,r}$ and $w^{h,m}$, which need not be the case. Instead, we rely on the following variational argument. Since $w^{h,r} \in \arg\max_{w\in\mathbb{R}_+}\{V^h(w) - (\rho - r)hw/(1 + \rho h)\}$, $w^{h,m} \in \arg\max_{w\in\mathbb{R}_+}\{V^h(w)\}$ and $\rho > r$, one has $w^{h,r} \leq w^{h,m}$. Suppose that $w^{h,r} = w^{h,m}$. It follows then from (13) that the contract $(x, u_+, u_-, w_+, w_-) = (1, \lambda(c_+^h - c_-^h), 0, w^{h,m} - (1+\rho)\varepsilon, w^{h,m} - (1+\rho)\varepsilon)$ satisfies (9)–(11) for all $\varepsilon > 0$ close enough to 0, and delivers a utility $w^{h,m} - \varepsilon$ to the entrepreneur. By (8), one must then have:

$$V^h(w^{h,m} - \varepsilon) \geq \mu h + \frac{V^h(w^{h,m} - (1 + \rho h)\varepsilon)}{1 + rh} - \frac{(\rho - r)h[w^{h,m} - (1 + \rho h)\varepsilon]}{(1 + rh)(1 + \rho h)}. \tag{A.8}$$

Moreover, since $w^{h,r} = w^{h,m}$, one has:

$$V^h(w^{h,m}) = \mu h + \frac{V^h(w^{h,m})}{1 + rh} - \frac{(\rho - r)hw^{h,m}}{(1 + rh)(1 + \rho h)}, \tag{A.9}$$

so rearranging (A.8) yields:

$$\frac{1 + \rho h}{1 + rh}\left[\frac{V^h(w^{h,m}) - V^h(w^{h,m} - (1 + \rho h)\varepsilon)}{(1 + \rho h)\varepsilon}\right] - \frac{V^h(w^{h,m}) - V^h(w^{h,m} - \varepsilon)}{\varepsilon} \geq \frac{(\rho - r)h}{1 + rh}.$$

Taking limits as ε goes to 0, and using the fact that $\rho > r$, one gets $V_-^{h\prime}(w^{h,m}) \geq 1$. Since V^h is concave and $V^h(0) = 0$, it follows that $V^h(w^{h,m}) \geq w^{h,m}$. However, since we have assumed that $w^{h,r} = w^{h,m}$, one can use (13) to obtain that $w^{h,m} = (1 + \rho h)p\lambda(c_+^h - c_-^h)/(\rho h)$, and from this and (3), it follows that $\mu h < p\lambda(c_+^h - c_-^h) = \rho h w^{h,m}/(1 + \rho h)$ for h close enough to 0. Using (A.9), it is then straightforward to check that this implies that $V^h(w^{h,m}) < w^{h,m}$, a contradiction. Hence $w^{h,r} < w^{h,m}$, as claimed.

We next prove that for h close enough to 0, $w^{h,l} \leq w^{h,r}$. Suppose by way of contradiction that $w^{h,r} < w^{h,l}$. Then, since V^h is linear over $[0, w^{h,l}]$ and $w^{h,r}$ is by definition the smallest point at which the mapping $w \mapsto V^h(w) - (\rho - r)hw/(1 + \rho h)$ reaches its maximum, one must have $w^{h,r} = 0$. By (13), this implies that $w^{h,m} = p\lambda(c_+^h - c_-^h)$ and that:

$$V^h(w) = \begin{cases} \dfrac{\mu h w}{p\lambda(c_+^h - c_-^h)} & \text{if } w < p\lambda(c_+^h - c_-^h), \\[3mm] \mu h & \text{if } w \geq p\lambda(c_+^h - c_-^h). \end{cases}$$

However, by (1)–(2), $c_+^h - c_-^h = \sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h}$. One then has:

$$V^{h\prime}(0) = \frac{\mu\sqrt{h}}{p\lambda\sigma(\varepsilon_+ - \varepsilon_-)} > \frac{(\rho - r)h}{1 + \rho h}$$

for h close enough to 0, and the mapping $w \mapsto V^h(w) - (\rho - r)hw/(1 + \rho h)$ cannot reach its maximum at 0, a contradiction. The result follows.   ‖

Proof of Proposition 2. We first define the processes $\{\hat{x}_{nh}^h\}_{n \in \mathbb{N}}$, $\{\hat{w}_{nh}^h\}_{n \in \mathbb{N}}$ and $\{\hat{u}_{nh}^h\}_{n \in \mathbb{N}}$. One sets $\hat{w}_0^h = w_0^h$, and as long as $\hat{w}_{nh}^h > 0$, one has:

$$\hat{x}_{nh}^h = \min\left\{\frac{\hat{w}_{nh}^h}{w^{h,l}}, 1\right\}, \tag{A.10}$$

$$\hat{w}_{(n+1)h}^h = \min\left\{(1 + \rho h)\left[\frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} + \lambda(c_{nh} - \mu h)\right], w^{h,r}\right\}, \tag{A.11}$$

$$\hat{u}_{nh}^h = \max\left\{\frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} - w^{h,d}, 0\right\}1_{\{c_{nh}^h = c_+^h\}}. \tag{A.12}$$

Let $\nu = \sup\{n \in \mathbb{N} \,|\, \hat{w}_{nh}^h > 0\}$, which may be finite or infinite. Then, in any period $n > \nu$, one has $\hat{x}_{nh}^h = \hat{w}_{nh}^h = \hat{u}_{nh}^h = 0$. We generalize the definitions (30)–(31) of the dividend and the coupon to take into account the possibility of downsizing in the current period. In any period $n \leq \nu$, this yields:

$$\hat{e}_{nh}^h = \left(\prod_{i=0}^n \hat{x}_{ih}^h\right)\max\left\{\frac{\lambda^h \hat{m}_{nh}^h}{\lambda \hat{x}_{nh}^h} - \frac{w^{h,d}}{\lambda}, 0\right\}1_{\{c_{nh}^h = c_+^h\}}, \tag{A.13}$$

$$\hat{b}_{nh}^h = \left(\prod_{i=0}^n \hat{x}_{ih}^h\right)\left[\mu h - \frac{(\rho - r)h\hat{m}_{nh}^h}{(1 + rh)\hat{x}_{nh}^h}\right]. \tag{A.14}$$

The proof now proceeds in three steps.

Step 1. We first argue that for λ close enough to 1 and h close enough to 0, $\hat{b}_{nh}^h$ stays non-negative for each value of $\hat{m}_{nh}^h/\hat{x}_{nh}^h$ in $[0, w^{h,r}/\lambda^h]$. By (A.14), this is the case if and only if:

$$\mu - \frac{(\rho - r)w^{h,r}}{(1 + \rho h)\lambda} \geq 0, \tag{A.15}$$

where we have used the fact that $\lambda^h = (1 + \rho h)\lambda/(1 + rh)$. One has:

$$V^h(w^{h,m}) = \mu h + \frac{V^h(w^{h,r})}{1 + rh} - \frac{(\rho - r)hw^{h,r}}{(1 + rh)(1 + \rho h)},$$

so that, since $w^{h,r} < w^{h,m}$ for h close enough to 0 by Lemma 3 and V^h is strictly increasing over $[0, w^{h,m}]$,

$$\mu - \frac{(\rho - r)w^{h,r}}{1 + \rho h} > r[V^h(w^{h,m}) - \mu h],$$

and, since $V^h(w^{h,m}) > 0$,

$$\frac{(\rho - r)w^{h,r}}{1 + \rho h} < (1 + rh)\mu.$$

Since $\lambda \in (0, 1]$, it follows from these two inequalities that:

$$\mu - \frac{(\rho - r)w^{h,r}}{(1 + \rho h)\lambda} > r[V^h(w^{h,m}) - \mu h] + \left(1 - \frac{1}{\lambda}\right)\frac{(\rho - r)w^{h,r}}{1 + \rho h}$$

$$\geq r[V^h(w^{h,m}) - \mu h] + \left(1 - \frac{1}{\lambda}\right)(1 + rh)\mu \qquad (A.16)$$

$$\geq r[V_1^h(w_1^{h,m}) - \mu h] + \left(1 - \frac{1}{\lambda}\right)(1 + rh)\mu,$$

where $V_1^h(w_1^{h,m})$ is the minimum over λ of $V^h(w^{h,m})$, which is clearly reached for $\lambda = 1$ from (8)–(11). One can check independently that when h goes to 0, $V_1^h(w_1^{h,m})$ converges to a strictly positive value, see Proposition 3. Taking limits as (λ, h) goes to $(1, 0)$ in (A.16) then yields that (A.15) holds for any pair (λ, h) close enough to $(1, 0)$, as claimed.

Step 2. Next, we show that the entrepreneur's utility process in the implementation is the same as in the optimal contract. The proof proceeds by induction. At date 0, $\hat{w}_0^h = \lambda^h \hat{m}_0^h$ by assumption. Now suppose that our implementation works until date nh, so that $\hat{w}_{nh}^h = \lambda^h \hat{m}_{nh}^h$ and the size of the firm at the beginning of period nh is $\prod_{i=0}^{n-1} \hat{x}_{ih}^h$. At date nh, the size-adjusted dividend is $\tilde{e}_{nh}^h = \hat{e}_{nh}^h / \prod_{i=0}^n \hat{x}_{ih}^h$, while the size-adjusted coupon is $\tilde{b}_{nh}^h = \hat{b}_{nh}^h / \prod_{i=0}^n \hat{x}_{ih}^h$. Two cases must be distinguished. Suppose first that $\hat{w}_{nh}^h \in [w^{h,l}, w^{h,r}]$, so that $\hat{x}_{nh}^h = 1$ by (A.10). Since $\hat{w}_{nh}^h = \lambda^h \hat{m}_{nh}^h$, (A.12)–(A.13) imply that $\tilde{e}_{nh}^h = \max\{\hat{w}_{nh}^h - w^{h,d}, 0\}1_{\{c_{nh}^h = c_+^h\}}/\lambda = \hat{u}_{nh}^h/\lambda$, which reflects the fact that the entrepreneur is compensated with a share λ of equity. Size-adjusted cash reserves evolve according to:

$$\hat{m}_{(n+1)h}^h = (1 + rh)(\hat{m}_{nh}^h + c_{nh}^h - \tilde{b}_{nh}^h - \tilde{e}_{nh}^h)$$

$$= (1 + rh)\left[\frac{(1 + \rho h)\hat{m}_{nh}^h}{1 + rh} + c_{nh}^h - \mu h - \tilde{e}_{nh}^h\right] \qquad (A.17)$$

$$= (1 + \rho h)\left[\hat{m}_{nh}^h + \frac{1 + rh}{1 + \rho h}(c_{nh}^h - \mu h - \tilde{e}_{nh}^h)\right],$$

where the second equality follows from (A.14). The optimal contract at date nh is implemented if and only if $\hat{w}_{nh}^h = \lambda^h \hat{m}_{nh}^h$ implies that $\hat{w}_{(n+1)h}^h = \lambda^h \hat{m}_{(n+1)h}^h$, where $\hat{w}_{(n+1)h}^h$ is as specified by (A.11). Multiplying (A.17) by λ^h, and using the fact that $\hat{w}_{nh}^h = \lambda^h \hat{m}_{nh}^h$ together with (A.13), one gets:

$$\lambda^h \hat{m}_{(n+1)h}^h = (1 + \rho h)[\hat{w}_{nh}^h + \lambda(c_{nh}^h - \mu h - \tilde{e}_{nh}^h)]$$

$$(A.18)$$

$$= (1 + \rho h)[\hat{w}_{nh}^h + \lambda(c_{nh}^h - \mu h) - \max\{\hat{w}_{nh}^h - w_h^d, 0\}1_{\{c_{nh}^h = c_{h+}\}}]$$

Using that $(1 + \rho h)[w - p\lambda(c_+^h - c_-^h)] \leq w^{h,r}$ for all $w \leq w^{h,m}$ together with the fact that $w^{h,d}, w^{h,r}$ and $w^{h,m}$ are related by (13)–(14), it is straightforward to check that the right-hand side of (A.18)

is equal to the right-hand side of (A.11) with $\hat{x}_{nh}^h = 1$. Hence $\hat{w}_{(n+1)h}^h = \lambda^h \hat{m}_{(n+1)h}^h$, as required. Suppose now that $\hat{w}_{nh}^h \in [0, w^{h,l})$. Then, according to the implementation, the firm is scaled down by a factor $\hat{m}_{nh}^h / (w^{h,l}/\lambda^h) = \hat{w}_{nh}^h / w^{h,l} = \hat{x}_{nh}^h$, as prescribed by the optimal contract, and the size-adjusted cash reserves jump to $w^{h,l}/\lambda^h$. One can then apply the implementation starting with a level $\hat{m}_{nh}^h / \hat{x}_{nh}^h = w^{h,l}/\lambda^h$ of size-adjusted cash reserves, and the claim follows.

Step 3. Finally, we check that our implementation delivers the required utility to the financiers. This is the case if and only if, in any period $n \le \nu$, the value of claims held by the financiers is equal to the sum of their utility $F^h(\hat{w}_{nh}^h)$ in the optimal contract and of the size-adjusted cash reserves $\hat{m}_{nh}^h$. For notational simplicity, we establish this result at date 0. It is then straightforward to adapt the proof to an arbitrary period $n \le \nu$. For each $n \in \mathbb{N}$, let $\mathbb{E}_{nh}$ be the expectation operator conditional on information up to date nh. One needs to show that:

$$\hat{F}^h(\hat{w}_0^h) + \hat{m}_0^h = \mathbb{E}_0 \left[\sum_{n=0}^{\nu} \frac{1}{(1+rh)^n} \left[(1-\lambda)\hat{e}_{nh}^h + \hat{b}_{nh}^h \right] \right]. \tag{A.19}$$

To establish (A.19), note that since the process $\{\hat{w}_{nh}^h\}_{n\in\mathbb{N}}$ is uniformly bounded, the series:

$$\hat{w}_0^h = \sum_{n=0}^{\nu} \frac{1}{(1+rh)^n} \left(\prod_{i=0}^{n} \hat{x}_{ih} \right) \left(\frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} - \frac{\hat{w}_{(n+1)h}^h}{1+rh} \right) \tag{A.20}$$

converges absolutely even when $\nu = \infty$. Using (A.11) together with the fact that $w^{h,d}$, $w^{h,r}$ and $w^{h,m}$ are related by (13)–(14), it is straightforward to check that, for each $n \le \nu$,

$$\mathbb{E}_{nh} \left[\frac{\lambda^h}{\lambda} \max \left\{ \frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} - w^{h,d}, 0 \right\} 1_{\{c_{nh}^h = c_+^h\}} - \frac{(\rho - r)h\hat{w}_{nh}^h}{(1+rh)\hat{x}_{nh}^h} \right] = \mathbb{E}_{nh} \left[\frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} - \frac{\hat{w}_{(n+1)h}^h}{1+rh} \right]. \tag{A.21}$$

Taking expectations in (A.20) and using (A.21) thus yields:

$$\hat{w}_0^h = \mathbb{E}_0 \left[\sum_{n=0}^{\nu} \frac{1}{(1+rh)^n} \left(\prod_{i=0}^{n} \hat{x}_{ih}^h \right) \left[\frac{\lambda^h}{\lambda} \max \left\{ \frac{\hat{w}_{nh}^h}{\hat{x}_{nh}^h} - w^{h,d}, 0 \right\} 1_{\{c_{nh}^h = c_+^h\}} - \frac{(\rho - r)h\hat{w}_{nh}^h}{(1+rh)\hat{x}_{nh}^h} \right] \right]. \tag{A.22}$$

For each $n \le \nu$, $\hat{w}_{nh}^h = \lambda^h \hat{m}_{nh}^h$. It then follows from (A.13)–(A.14) and (A.22) that:

$$\hat{m}_0^h = \mathbb{E}_0 \left[\sum_{n=0}^{\nu} \frac{1}{(1+rh)^n} \left(\prod_{i=0}^{n} \hat{x}_{ih}^h \right) (\tilde{e}_{nh}^h + \tilde{b}_{nh}^h - \mu h) \right], \tag{A.23}$$

where $\tilde{e}_{nh}^h$ and $\tilde{b}_{nh}^h$ are defined as in Step 2. To conclude the proof, observe that by construction, $F^h(\hat{w}_0^h)$ is equal to the present value, discounted at rate r, of the cash-flows from the project, less the payments to the entrepreneur:

$$F^h(\hat{w}_0^h) = \mathbb{E}_0 \left[\sum_{n=0}^{\nu} \frac{1}{(1+rh)^n} \left(\prod_{i=0}^{n} \hat{x}_{ih}^h \right) (\mu h - \lambda \tilde{e}_{nh}^h) \right]. \tag{A.24}$$

Summing (A.23) and (A.24) yields (A.19). The result follows. ∥

APPENDIX B

Proof of Lemma 4. We prove that $w^{h,l} < (1+\rho h)\lambda(c_+^h - c_-^h)$ for all h close enough to 0, which implies that $\lim_{h\downarrow 0} w^{h,l} = 0$. Suppose that $w^{h,l} \ge (1+\rho h)\lambda(c_+^h - c_-^h)$ along a subsequence of h's that goes to

0. Then, by Lemma 3, one has $w^{h,r} \geq (1 + \rho h)\lambda(c_{h+} - c_{h-})$ along this subsequence. It follows that for any such h, the choice $(w_+, w_-) = ((1 + \rho h)\lambda(c_+^h - c_-^h), 0)$ is feasible and optimal in (A.1)–(A.4) for $w = p\lambda(c_+^h - c_-^h)$. Denoting by α^h the slope of V^h over $(0, w^{h,l})$ and using the fact that, by assumption, $w^{h,l} \geq (1 + \rho h)\lambda(c_+^h - c_-^h)$, one thus has:

$$T^{h,c}V^h(p\lambda(c_+^h - c_-^h)) = \mu h + \frac{p\alpha^h(1 + \rho h)\lambda(c_+^h - c_-^h)}{1 + rh} - \frac{p(\rho - r)h\lambda(c_+^h - c_-^h)}{1 + rh},$$

or, equivalently, by (1)–(2):

$$\frac{T^{h,c}V^h(p\lambda(c_+^h - c_-^h))}{p\lambda(c_+^h - c_-^h)} = \frac{\mu\sqrt{h}}{p\lambda\sigma(\varepsilon_+ - \varepsilon_-)} + \frac{\alpha^h(1 + \rho h)}{1 + rh} - \frac{(\rho - r)h}{1 + rh}.$$

Since $\rho > r$, it follows that for h close enough to 0,

$$\frac{T^{h,c}V^h(p\lambda(c_+^h - c_-^h))}{p\lambda(c_+^h - c_-^h)} > \alpha^h,$$

which contradicts the fact that $V^h(p\lambda(c_+^h - c_-^h))) = \alpha^h p\lambda(c_+^h - c_-^h)$ since $p\lambda(c_+^h - c_-^h)$ belongs to the interval $[0, w^{h,l}]$, see (10). The result follows. $\parallel$

Proof of Corollary 1. Suppose first that $\liminf_{h\downarrow 0} w^{h,m} < w^m$. Then, since V^h converges uniformly to V, $V^h(w^{h,m})$ converges to $V(\liminf_{h\downarrow 0} w^{h,m}) < V(w^m)$ along a subsequence of h's that goes to 0. However, since V^h attains its maximum at $w^{h,m}$ and V attains its maximum at w^m, uniform convergence implies that $\lim_{h\downarrow 0} V^h(w^{h,m}) = V(w^m)$, a contradiction. Therefore $\liminf_{h\downarrow 0} w^{h,m} \geq w^m$. We now prove that $\limsup_{h\downarrow 0} w^{h,r} \leq w^m$, which implies the result since $\limsup_{h\downarrow 0} w^{h,m} = \limsup_{h\downarrow 0} w^{h,r}$ by (13). It follows from Lemma 3 that, for h close enough to 0,

$$V^h(w^{h,m}) = \mu h + \frac{V^h(w^{h,r})}{1 + rh} - \frac{(\rho - r)hw^{h,r}}{(1 + rh)(1 + \rho h)} < \mu h + \frac{V^h(w^{h,m})}{1 + rh} - \frac{(\rho - r)hw^{h,r}}{(1 + rh)(1 + \rho h)},$$

so that in particular:

$$\frac{(\rho - r)w^{h,r}}{(1 + rh)(1 + \rho h)} < \mu - \frac{rV^h(w^{h,m})}{1 + rh}.$$

Taking limits as h goes to 0 and using the fact that $\mu - rV(w^m) = (\rho - r)w^m$ by (33) and (35)–(36) yields $\limsup_{h\downarrow 0} w^{h,r} \leq w^m$ as required. This completes the proof. $\parallel$

Proof of Lemma 5. To begin with, fix some $w^m > 0$, and consider the differential equation:

$$rV(w) = \mu - (\rho - r)w + \rho w V'(w) + \frac{\lambda^2 \sigma^2}{2} V''(w) \qquad (B.1)$$

for $w \in [0, w^m]$. Any solution to (B.1) can be written as the sum of a particular solution, namely $w + \mu/r$, and a general solution to the second-order linear homogenous differential equation:

$$rH(w) = \rho w H'(w) + \frac{\lambda^2 \sigma^2}{2} H''(w). \qquad (B.2)$$

Consider the solutions H_0 and H_1 of (B.2) over $\mathbb{R}_+$ characterized by the initial conditions $H_0(0) = 1$, $H_0'(0) = 0$, $H_1(0) = 0$ and $H_1'(0) = 1$, and let $W_{H_0,H_1} = H_0 H_1' - H_1 H_0'$ be the Wronskian of H_0 and H_1. Then $W_{H_0,H_1}(0) = 1 \neq 0$ which shows that H_0 and H_1 are linearly independent over $\mathbb{R}_+$. As a result of this, (H_0, H_1) is a basis of the 2-dimensional space of solutions to (B.2). One can thus write:

$$V(w) = w + \frac{\mu}{r} + \alpha_0 H_0(w) + \alpha_1 H_1(w)$$

34

for all $w \in [0, w^m]$. To pin down the coefficients α_0 and α_1, we impose the boundary conditions $V(0) = 0$ and $V'(w^m) = 0$. This yields:

$$V(w) = w + \frac{\mu}{r}\left[1 - H_0(w)\right] + \frac{H_1(w)}{H_1'(w^m)}\left[\frac{\mu}{r}\, H_0'(w^m) - 1\right]. \tag{B.3}$$

We now establish that there exists a unique $w^m > 0$ and a unique function V satisfying (B.3) such that $V''(w^m) = 0$. First, we show that $H_1' > 0$ over $\mathbb{R}_+$. Suppose by way of contradiction that $\tilde{w} = \inf\{w \geq 0 \mid H_1'(w) \leq 0\} < \infty$. Then $H_1'(\tilde{w}) = 0$ and $H_1''(\tilde{w}) \leq 0$. By (B.2), it follows that $H_1(\tilde{w}) \leq 0$. However, since $H_1'(0) = 1$, one has $\tilde{w} > 0$, and since H_1 is strictly increasing over $[0, \tilde{w}]$, it follows that $H_1(\tilde{w}) > H_1(0) = 0$, a contradiction. Hence $H_1' > 0$ over $\mathbb{R}_+$, as claimed. As a result of this, $H_1 > 0$ over $\mathbb{R}_{++}$. Next, by (B.1) and the boundary condition $V'(w^m) = 0$, one has:

$$\frac{\lambda^2 \sigma^2}{2}\, V''(w^m) = rV(w^m) - \mu + (\rho - r)w^m.$$

Using (B.3) and the definition of W_{H_0, H_1}, we obtain that:

$$\frac{\lambda^2 \sigma^2}{2}\, V''(w^m) = \rho w^m - \frac{\mu W_{H_0, H_1}(w^m) + rH_1(w^m)}{H_1'(w^m)}. \tag{B.4}$$

Then by (B.2),

$$W_{H_0, H_1}'(w) = H_0(w)H_1''(w) - H_1(w)H_0''(w)$$

$$= \frac{2}{\lambda^2 \sigma^2}\left\{H_0(w)[rH_1(w) - \rho w H_1'(w)] - H_1(w)[rH_0(w) - \rho w H_0'(w)]\right\} \tag{B.5}$$

$$= -\frac{2\rho w}{\lambda^2 \sigma^2}\, W_{H_0, H_1}(w)$$

for all $w \geq 0$. Since $W_{H_0, H_1}(0) = 1$, one obtains Abel's identity by integration of (B.5):

$$W_{H_0, H_1}(w) = \exp\left(\frac{-\rho w^2}{\lambda^2 \sigma^2}\right)$$

for all $w \geq 0$. Define a function φ by:

$$\varphi(w) = [\rho w H_1'(w) - rH_1(w)]\exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right)$$

for all $w \geq 0$. From (B.4) and the expression for W_{H_0, H_1}, we need to show that there exists a unique $w^m > 0$ such that $\varphi(w^m) = \mu$. Since $\varphi(0) = 0$, it is sufficient to prove that φ is strictly increasing and strictly convex over $\mathbb{R}_+$. By (B.2), one has:

$$\varphi'(w) = \left\{(\rho - r)H_1'(w) + \rho w H_1''(w) + \frac{2\rho w}{\lambda^2 \sigma^2}\left[\rho w H_1'(w) - rH_1(w)\right]\right\}\exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right)$$

$$= (\rho - r)H_1'(w)\exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right),$$

which is strictly positive for all $w \geq 0$ as $H_1' > 0$ over $\mathbb{R}_+$. Hence φ is strictly increasing over $\mathbb{R}_+$. By (B.2) again, one has:

$$\varphi''(w) = (\rho - r)\left[H_1''(w) + \frac{2\rho w}{\lambda^2 \sigma^2}\, H_1'(w)\right]\exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right) = \frac{2r}{\lambda^2 \sigma^2}\,(\rho - r)H_1(w)\exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right),$$

which is strictly positive for all $w > 0$ as $H_1 > 0$ over $\mathbb{R}_{++}$. Hence φ is strictly convex over $\mathbb{R}_+$, which concludes the proof. $\parallel$

Proof of Lemma 6. Using the fact that the solution V to (33)–(36) is smooth over $[0, w^m)$, one can differentiate the differential equation in (33) to obtain:

$$rV'(w) = -\rho + r + \rho V'(w) + \rho w V''(w) + \frac{\lambda^2 \sigma^2}{2} V'''(w)$$

for all $w \in [0, w^m)$. From (35)–(36), it follows that $V_-'''(w^m) = 2(\rho - r)/(\lambda^2 \sigma^2) > 0$, and thus $V'' < 0$ and $V' > 0$ over an interval $(w^m - \varepsilon, w^m)$ for $\varepsilon > 0$ small enough. From (33), $V''(w) < 0$ whenever $rV(w) < \mu - (\rho - r)w$ and $V'(w) > 0$. Since $rV(w^m) = \mu - (\rho - r)w^m$ and V is strictly increasing over $(w^m - \varepsilon, w^m)$, it follows that these two conditions are satisfied over $(w^m - \varepsilon, w^m)$, and the first of these conditions remains satisfied for any $w < w^m$ as long as V' remains strictly positive over $[w, w^m)$. Now suppose that $V'(w) \leq 0$ for some $w < w^m - \varepsilon$, and let $\tilde{w} = \sup\{w < w^m - \varepsilon \,|\, V'(w) \leq 0\}$. Then $V'(\tilde{w}) = 0$, and for all $w \in (\tilde{w}, w^m)$, $V'(w) > 0$ and $rV(w) < \mu - (\rho - r)w$. It follows that $V''(w) < 0$ for any such w and thus, by (35), $V'(\tilde{w}) = -\int_{\tilde{w}}^{w^m} V''(w) \, dw > 0$, a contradiction. Hence $V' > 0$ and thus $V'' < 0$ over $[0, w^m)$, which concludes the proof. $\parallel$

Proof of Lemma 7. Since V is strictly concave over $[0, w^m]$ by Lemma 6, there exists a unique point $w_V^{h,r} \in [0, w^m)$ at which the mapping $w \mapsto V(w) - (\rho - r)hw/(1 + \rho h)$ reaches its maximum, and for h close enough to 0, $w_V^{h,r}$ is characterized by $V'(w_V^{h,r}) = (\rho - r)h/(1 + \rho h)$. Since $V'(w^m) = 0$ and V is continuously differentiable, $\lim_{h \downarrow 0} w_V^{h,r} = w^m$. In analogy with (13)–(14), define:

$$w_V^{h,m} = \frac{w_V^{h,r}}{1 + \rho h} + p\lambda(c_+^h - c_-^h), \tag{B.6}$$

$$w_V^{h,d} = w_V^{h,m} - \lambda(c_+^h - c_-^h). \tag{B.7}$$

It should be noted that $\lim_{h \downarrow 0} w_V^{h,m} = \lim_{h \downarrow 0} w_V^{h,d} = w^m$. In particular, $w_V^{h,d} > p\lambda(c_+^h - c_-^h)$ for h close enough to 0. By (1)–(2) and (B.6)–(B.7), one has:

$$w_V^{h,m} = w_V^{h,r} + p\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}) > w_V^{h,r}, \tag{B.8}$$

$$w_V^{h,d} = w_V^{h,r} - (1 - p)\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}) < w_V^{h,r} \tag{B.9}$$

for h close enough to 0. Moreover, using that $V'(w^m) = V''(w^m) = 0$, $V_-'''(w^m) = 2(\rho - r)/(\lambda^2 \sigma^2)$ and $V'(w_V^{h,r}) = (\rho - r)h/(1 + \rho h)$, a Taylor–Young expansion yields:

$$w^m = w_V^{h,r} + \lambda\sigma\sqrt{h} + o(\sqrt{h}). \tag{B.10}$$

Proceeding as in Subsection 2.3, it is straightforward to characterize $T^{h,c}V$ when h is close to 0, where $T^{h,c}$ is defined by (A.1)–(A.4). For any $w \geq 0$, define:

$$w_+^h(w) = (1 + \rho h)[w + (1 - p)\lambda(c_+^h - c_-^h)],$$

$$w_-^h(w) = (1 + \rho h)[w - p\lambda(c_+^h - c_-^h)].$$

Three cases must then be distinguished. When $w \in [w_V^{h,m}, \infty)$, $(w_+, w_-) = (w_V^{h,r}, w_V^{h,r})$ is optimal. Next, when $w \in [w_V^{h,d}, w_V^{h,m})$, $(w_+, w_-) = (w_V^{h,r}, w_-^h(w))$ is optimal. Finally, when $w \in [p\lambda(c_+^h -$

$c_-^h), w_V^{h,d}), (w_+, w_-) = (w_+^h(w), w_-^h(w))$ is optimal. For each $h > 0$ and $w \geq p\lambda(c_+^h - c_-^h)$, let $\Delta^h(w) = (1 + rh)[T^{h,c}V(w) - V(w)]$. We now establish that $\|\Delta^h\|_h = o(h)$, which implies the result.

Consider first the region $[w_V^{h,m}, \infty)$. For any w in this interval, one has:

$$T^{h,c}V(w) = \mu h + \frac{V(w_V^{h,r})}{1 + rh} - \frac{(\rho - r)hw_V^{h,r}}{(1 + rh)(1 + \rho h)}.$$

Multiplying by $1 + rh$ and subtracting $(1 + rh)V(w)$, we obtain that:

$$\Delta^h(w) = [\mu - rV(w_V^{h,r}) - (\rho - r)w_V^{h,r}]h + (1 + rh)[V(w_V^{h,r}) - V(w)] + o(h).$$

Since V attains its maximum at w^m, we have the following lower bound for $\Delta^h(w)$:

$$\Delta^h(w) \geq [\mu - rV(w_V^{h,r}) - (\rho - r)w_V^{h,r}]h + (1 + rh)[V(w_V^{h,r}) - V(w^m)] + o(h). \tag{B.11}$$

We show that the right-hand side of (B.11) is a $o(h)$. For the first term, this follows at once from the fact that $\lim_{h \downarrow 0} w_V^{h,r} = w^m$ and $rV(w^m) = \mu - (\rho - r)w^m$ by (33) and (35)–(36). For the second term, note that V is concave and that $w^m > w_V^{h,r}$ for all $h > 0$. Together with the definition of $w_V^{h,r}$, this implies that, for h close enough to 0,

$$\frac{|V(w_V^{h,r}) - V(w^m)|}{h} \leq \frac{V'(w_V^{h,r})(w^m - w_V^{h,r})}{h} = \frac{(\rho - r)(w^m - w_V^{h,r})}{1 + \rho h},$$

and the result follows since $\lim_{h \downarrow 0} w_V^{h,r} = w^m$. Next, from (B.8), $w \geq w_V^{h,m} > w_V^{h,r}$ for h close enough to 0, and thus we have the following upper bound for $\Delta^h(w)$:

$$\Delta^h(w) \leq [\mu - rV(w_V^{h,r}) - (\rho - r)w_V^{h,r}]h + o(h). \tag{B.12}$$

Proceeding as for the first term of the right-hand side of (B.11), one obtains that the right-hand side of (B.12) is a $o(h)$. It follows that:

$$\sup_{w \in [w_V^{h,m}, \infty)} \{|\Delta^h(w)|\} = o(h). \tag{B.13}$$

Consider next the region $[w_V^{h,d}, w_V^{h,m})$. For any w in this interval, one has:

$$T^{h,c}V(w) = \mu h + \frac{pV(w_V^{h,r}) + (1 - p)V(w_-^h(w))}{1 + rh} - \frac{(\rho - r)h[pw_V^{h,r} + (1 - p)w_-^h(w)]}{(1 + rh)(1 + \rho h)}.$$

Multiplying by $1 + rh$ and subtracting $(1 + rh)V(w)$, we obtain that:

$$\Delta^h(w) = \{\mu - rV(w) - (\rho - r)[pw_V^{h,r} + (1 - p)w_-^h(w)]\}h$$

$$+ p[V(w_V^{h,r}) - V(w)] + (1 - p)[V(w_-^h(w)) - V(w)].$$

Proceeding as for the region $[w_V^{h,m}, \infty)$, it is easy to check that:

$$\sup_{w \in [w_V^{h,d}, w_V^{h,m})} \{|\mu - rV(w) - (\rho - r)[pw_V^{h,r} + (1 - p)w_-^h(w)]|\}h = o(h).$$

We now consider the two remaining terms in $V(w_V^{h,r}) - V(w)$ and $V(w_-^h(w)) - V(w)$ that appear in the expression of $\Delta^h(w)$. The following argument is valid for both terms, so we shall only consider the first. A Taylor–Young expansion yields:

$$V(w_V^{h,r}) - V(w) = V'(w)(w_V^{h,r} - w) + \frac{1}{2}V''(w)(w_V^{h,r} - w)^2 + o((w_V^{h,r} - w)^2). \tag{B.14}$$

37

We treat each term on the right-hand side of (B.14) separately. For any $w \in [w_V^{h,d}, w_V^{h,m})$, (B.8)–(B.10) imply that:

$$|w_V^{h,r} - w| \leq \lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}), \tag{B.15}$$

$$|w^m - w| \leq \lambda\sigma(1 + \varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}). \tag{B.16}$$

Since V''', which is defined everywhere except at w^m, is bounded in absolute value by a positive constant K, and since $V''(w^m) = 0$ by (36), (B.16) implies that:

$$|V''(w)| \leq K|w^m - w| \leq \lambda\sigma(1 + \varepsilon_+ - \varepsilon_-)K\sqrt{h} + o(\sqrt{h}) \tag{B.17}$$

for all $w \in [w_V^{h,d}, w_V^{h,m})$. Finally, since $V'(w^m) = 0$ by (35), (B.16)–(B.17) imply that:

$$V'(w) = -\int_w^{w^m} V''(x)\,dx \leq \lambda^2\sigma^2(1 + \varepsilon_+ - \varepsilon_-)^2 Kh + o(h) \tag{B.18}$$

for all $w \in [w_V^{h,d}, w_V^{h,m})$. From (B.14)–(B.18), one has:

$$\sup_{w \in [w_V^{h,d}, w_V^{h,m})} \{|V(w_V^{h,r}) - V(w)|\} = o(h),$$

and a similar result holds for $V(w_-^h(w)) - V(w)$. It follows that:

$$\sup_{w \in [w_V^{h,d}, w_V^{h,m})} \{|\Delta^h(w)|\} = o(h). \tag{B.19}$$

Consider finally the region $[p\lambda(c_+^h - c_-^h), w_V^{h,d})$. For any w in this interval, one has:

$$T^{h,c}V(w) = \mu h + \frac{pV(w_+^h(w)) + (1-p)V(w_-^h(w))}{1+rh} - \frac{(\rho - r)h[pw_+^h(w) + (1-p)w_-^h(w)]}{(1+rh)(1+\rho h)}$$

Multiplying by $1 + rh$ and subtracting $(1 + rh)V(w)$, we obtain that:

$$\Delta^h(w) = [\mu - rV(w) - (\rho - r)w]h + p[V(w_+^h(w)) - V(w)] + (1-p)[V(w_-^h(w)) - V(w)] + o(h).$$

A Taylor–Young expansion yields:

$$\Delta^h(w) = [\mu - rV(w) - (\rho - r)w + \rho wV'(w)]h$$

$$+ \frac{1}{2}V''(w)\{p[w_+^h(w) - w]^2 + (1-p)[w_-^h(w) - w]^2\} \tag{B.20}$$

$$+ o([w_+^h(w) - w]^2) + o([w_-^h(w) - w]^2) + o(h).$$

By (B.9), $w_V^{h,d} < w_V^{h,r} < w^m$ for h close enough to 0. Hence V satisfies the differential equation in (33) at any $w \in [p\lambda(c_+^h - c_-^h), w_V^{h,d})$. One can therefore rewrite (B.20) as:

$$\Delta^h(w) = \frac{1}{2}V''(w)\{p[w_+^h(w) - w]^2 + (1-p)[w_-^h(w) - w]^2 - \lambda^2\sigma^2 h\}$$

$$\tag{B.21}$$

$$+ o([w_+^h(w) - w]^2) + o([w_-^h(w) - w]^2) + o(h).$$

Using (1)–(2) and the definitions of $w_+^h(w)$ and $w_-^h(w)$, one has:

$$w_+^h(w) - w = (1-p)\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}) \tag{B.22}$$

$$w_-^h(w) - w = -p\lambda\sigma(\varepsilon_+ - \varepsilon_-)\sqrt{h} + o(\sqrt{h}) \tag{B.23}$$

Using that $p(1-p)(\varepsilon_+ - \varepsilon_-)^2 = 1$ and that V'' is bounded, it follows from (B.21)–(B.23) that:

$$\sup_{w\in[p\lambda(c_+^h - c_-^h), w_V^{h,d})} \{|\Delta^h(w)|\} = o(h). \tag{B.24}$$

Collecting (B.13), (B.19) and (B.24) yields that $\|\Delta^h\|_h = o(h)$ by definition of the norm $\|\cdot\|_h$. The result follows. $\|$

APPENDIX C

In all this appendix, we fix a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. All the stochastic processes we shall consider have sample paths that are $\mathbb{P}$–almost surely right continuous with left limits. A real valued process of this kind can thus be considered as a random variable taking its values in the space $D(\mathbb{R}_+)$ of functions from $\mathbb{R}_+$ to $\mathbb{R}$ that are right continuous and have left limits, endowed with the standard Prokhorov metric δ (Jacod and Shiryaev (1987, Chapter VI, Section 1c)). Such a process $Y^h = \{Y_t^h\}_{t\geq 0}$ converges in law to a process $Y = \{Y_t\}_{t\geq 0}$ as h goes to 0 if the probability measure induced by Y^h over $D(\mathbb{R}_+)$ converges weakly to the probability measure induced by Y over $D(\mathbb{R}_+)$.

Proof of Proposition 4. Let $W_0 = w$ and suppose that $w \in [0, w^m]$. Since L is continuous with bounded variation, it follows from (40) and Itô's formula for continuous semimartingales (Karatzas and Shreve (1991, Chapter 3, Theorem 3.6)) that, for each $T \geq 0$,

$$e^{-\rho T\wedge\tau}W_{T\wedge\tau} = w + \int_0^{T\wedge\tau} e^{-\rho t}\lambda\sigma\, dZ_t - \int_0^{T\wedge\tau} e^{-\rho t}\, dL_t. \tag{C.1}$$

We first take expectations in (C.1) and then let T go to ∞. Since $W_\tau = 0$ by definition of τ, and since the second term on the right-hand side of (C.1) is a martingale, we obtain that:

$$w - \mathbb{E}^{(w,0)}\left[\int_0^\tau e^{-\rho t}\, dL_t\right] = 0,$$

from which (43) follows. Next, by (33)–(35), F satisfies:

$$rF(w) = \mu + \rho wF'(w) + \frac{\lambda^2\sigma^2}{2}F''(w) \tag{C.2}$$

for all $w \in [0, w^m]$, with boundary conditions:

$$F(0) = 0, \tag{C.3}$$

$$F'(w^m) = -1. \tag{C.4}$$

Since L is continuous with bounded variation, it follows from (40) and Itô's formula that, for each

$T \geq 0$,

$$e^{-rT \wedge \tau} F(W_{T \wedge \tau}) = F(w) + \int_0^{T \wedge \tau} e^{-rt} \left[-rF(W_t) + \rho W_t F'(W_t) + \frac{\lambda^2 \sigma^2}{2} F''(W_t) \right] dt$$

$$+ \int_0^{T \wedge \tau} e^{-rt} \lambda \sigma F'(W_t) \, dZ_t - \int_0^{T \wedge \tau} e^{-rt} F'(W_t) \, dL_t \qquad \text{(C.5)}$$

$$= F(w) - \int_0^{T \wedge \tau} e^{-rt} \mu \, dt + \int_0^{T \wedge \tau} e^{-rt} \lambda \sigma F'(W_t) \, dZ_t + \int_0^{T \wedge \tau} e^{-rt} \, dL_t,$$

where the second equality follows from (41) and (C.2), together with (42) and (C.4). We first take expectations in (C.5) and then let T go to ∞. Since $F(W_\tau) = F(0) = 0$ by definition of τ and (C.3), and since the third term on the right-hand side of (C.5) is a martingale, we obtain that:

$$F(w) - \mathbb{E}^{(w,0)} \left[\int_0^\tau e^{-rt} (\mu - dL_t) \, dt \right] = 0,$$

from which (44) follows. ‖

Proof of Proposition 5. Before proceeding with the proof, we recall some technical results on the Skorokhod problem (see Tanaka (1979)). We start with the deterministic case. Let $y \in D(\mathbb{R}_+)$ such that $y(0) \leq 0$. Then a pair $(x, k) \in D(\mathbb{R}_+)^2$ solves the Skorokhod problem associated to y if $x = y - k$, $x \leq 0$, and k is a non-decreasing function such that $k(0) = 0$ and $k(t) = \int_0^t 1_{\{x(s)=0\}} \, dk(s)$ for all $t \geq 0$. The interpretation is that k increases only at the points where x hits 0, which causes x to be reflected back at 0. It can be shown that there exists a unique solution to this problem, which is characterized by $k(t) = \sup_{s \in [0,t]} \{ \max\{y(s), 0\} \}$ for all $t \geq 0$. Now let $Y = \{Y_t\}_{t \geq 0}$ be a stochastic process and let $\mathcal{A}^Y$ be the filtration generated by Y. Then a pair $(X, K) = \{(X_t, K_t)\}_{t \geq 0}$ of $\mathcal{A}^Y$–adapted processes solves the Skorokhod problem associated to Y if for $\mathbb{P}$–almost every $\omega \in \Omega$, $(X(\omega), K(\omega))$ solves the Skorokhod problem associated to $Y(\omega)$. One has the following lemma.

Lemma C.1. *Let $\{y_n\}_{n \in \mathbb{N}}$ be a sequence of real numbers such that $y_0 \leq 0$, and let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence of real numbers such that $x_0 = y_0$ and:*

$$x_{n+1} = \min\{x_n + y_{n+1} - y_n, 0\} \qquad \text{(C.6)}$$

for all $n \in \mathbb{N}$. Then:

$$x_n = y_n - \max_{i \in [\![0,n]\!]} \{ \max\{y_i, 0\} \} \qquad \text{(C.7)}$$

for all $n \in \mathbb{N}$, where $[\![0, n]\!] = \{0, \dots, n\}$.

Proof. The proof proceeds by induction. Since $x_0 = y_0 \leq 0$, (C.7) holds for $n = 0$. Suppose next that (C.7) holds for some $n \in \mathbb{N}$. According to (C.6), two cases must be distinguished. Suppose first that $x_n + y_{n+1} - y_n \leq 0$. Then, by the induction hypothesis, $y_{n+1} \leq y_n - x_n = \max_{i \in [\![0,n]\!]} \{ \max\{y_i, 0\} \}$, which implies that $\max_{i \in [\![0,n]\!]} \{ \max\{y_i, 0\} \} = \max_{i \in [\![0,n+1]\!]} \{ \max\{y_i, 0\} \}$. By the induction hypothesis again, $x_{n+1} = x_n + y_{n+1} - y_n = y_{n+1} - \max_{i \in [\![0,n]\!]} \{ \max\{y_i, 0\} \} = y_{n+1} - \max_{i \in [\![0,n+1]\!]} \{ \max\{y_i, 0\} \}$, so that (C.7) holds for $n + 1$. The case $x_n + y_{n+1} - y_n > 0$ can be handled by similar considerations, and the result follows by induction. ‖

The intuitive interpretation of Lemma C.1 is that the sequence $\{x_n\}_{n \in \mathbb{N}}$ recursively defined by (C.6) is part of the solution to a discrete version of the Skorokhod problem associated to $\{y_n\}_{n \in \mathbb{N}}$. As shown by (C.7), the reflecting sequence is given by $\{ \max_{i \in [\![0,n]\!]} \{ \max\{y_i, 0\} \} \}_{n \in \mathbb{N}}$ in analogy with the solution to the continuous-time Skorokhod problem.

The proof now proceeds in three steps.

Step 1. For any fixed period length $h > 0$, consider a modification $\{\tilde{w}_{nh}^h\}_{n \in \mathbb{N}}$ of the utility process $\{w_{nh}^h\}_{n \in \mathbb{N}}$, that starts at the same value $w_0^h \leq w^{h,r}$, but whose dynamics is given by (23)–(24) over the whole interval $(-\infty, w^{h,r}]$. Using (1)–(2), this yields:

$$\tilde{w}_{(n+1)h}^h = \min\{(1 + \rho h)[\tilde{w}_{nh}^h + \lambda \sigma \varepsilon_n \sqrt{h}], w^{h,r}\}$$

for all $n \in \mathbb{N}$. Rearranging this expression, we obtain:

$$\frac{\tilde{w}_{(n+1)h}^h - w^{h,r}}{(1 + \rho h)^{n+1}} = \min\left\{\frac{\tilde{w}_{nh}^h - w^{h,r}}{(1 + \rho h)^n} + \frac{\rho w^{h,r} h}{(1 + \rho h)^{n+1}} + \frac{\lambda \sigma \varepsilon_n \sqrt{h}}{(1 + \rho h)^n}, 0\right\} \tag{C.8}$$

for all $n \in \mathbb{N}$. Now define a process $\{y_{nh}^h\}_{n \in \mathbb{N}}$ such that $y_0^h = w_0^h - w^{h,r}$ and:

$$y_{(n+1)h}^h = y_{nh}^h + \frac{\rho w^{h,r} h}{(1 + \rho h)^{n+1}} + \frac{\lambda \sigma \varepsilon_n \sqrt{h}}{(1 + \rho h)^n} \tag{C.9}$$

for all $n \in \mathbb{N}$. Applying Lemma C.1 path by path, it then follows from (C.8)–(C.9) that:

$$\frac{\tilde{w}_{nh}^h(\omega) - w^{h,r}}{(1 + \rho h)^n} = y_{nh}^h(\omega) - \max_{i \in [\![0,n]\!]}\{\max\{y_{ih}^h(\omega), 0\}\} \tag{C.10}$$

for all $(\omega, n) \in \Omega \times \mathbb{N}$. Define continuous-time processes $\tilde{W}^h = \{\tilde{W}_t^h\}_{t \geq 0}$ and $Y^h = \{Y_t^h\}_{t \geq 0}$ by setting $\tilde{W}_t^h = \tilde{w}_{[\![t/h]\!]h}^h$ and $Y_t^h = y_{[\![t/h]\!]h}^h$ for all $t \geq 0$. By construction, $\tilde{W}^h$ and Y^h have sample paths that are right continuous with left limits. The following continuous-time version of (C.10) holds:

$$\frac{\tilde{W}_t^h(\omega) - w^{h,r}}{(1 + \rho h)^{[\![t/h]\!]}} = Y_t^h(\omega) - \sup_{s \in [0,t]}\{\max\{Y_s^h(\omega), 0\}\} \tag{C.11}$$

for all $(\omega, t) \in \Omega \times \mathbb{R}_+$. Let $\mathcal{A}^{Y^h}$ be the filtration generated by Y^h, and consider the $\mathcal{A}^{Y^h}$–adapted processes $X^h = \{X_t^h\}_{t \geq 0}$ and $K^h = \{K_t^h\}_{t \geq 0}$ defined by $X_t^h = (1 + \rho h)^{-[\![t/h]\!]}(\tilde{W}_t^h - w^{h,r})$ and $K_t^h = \sup_{s \in [0,t]}\{\max\{Y_s^h(\omega), 0\}\}$ for all $t \geq 0$. Then (C.11) expresses the fact that the pair (X^h, K^h) solves the Skorokhod problem associated to Y^h.

We now study the limits in law of the processes X^h and Y^h as h goes to 0. Consider first Y^h. As explained in Subsection 3.2, the limit $W_0 = \lim_{h \downarrow 0} W_0^h$ is well defined and belongs to $[0, w^m]$. Moreover, by Corollary 1, $\lim_{h \downarrow 0} w^{h,r} = w^m$. Given (C.9) and the definition of Y^h, it follows then from Jacod and Shiryaev (1987, Chapter IX, Theorem 4.21) that the process Y^h converges in law to a diffusion process $Y = \{Y_t\}_{t \geq 0}$ such that $Y_0 = W_0 - w^m$ and:

$$dY_t = e^{-\rho t}(\rho w^m dt + \lambda \sigma dZ_t) \tag{C.12}$$

for all $t \geq 0$, where $Z = \{Z_t\}_{t \geq 0}$ is the standard Brownian motion in (39). Let $\mathcal{A}^Y$ be the filtration generated by Y. From Tanaka (1979, Theorem 4.1), there exists a unique continuous $\mathcal{A}^Y$–adapted process $X = \{X_t\}_{t \geq 0}$ starting at $W_0 - w^m$ and a unique non-decreasing $\mathcal{A}^Y$–adapted process $K = \{K_t\}_{t \geq 0}$ starting at 0 that jointly satisfy:

$$X_t = Y_t - K_t, \tag{C.13}$$

$$X_t \leq 0, \tag{C.14}$$

$$K_t = \int_0^t 1_{\{X_s = 0\}} \, dK_s \tag{C.15}$$

for all $t \geq 0$. That is, the pair (X, K) solves the Skorokhod problem associated to Y. Since Y^h converges in law to Y, it follows from Słomiński (1993, Proposition 4) that the triple (X^h, K^h, Y^h) converges in law to (X, K, Y) over the space $D(\mathbb{R}_+, \mathbb{R}^3)$ of functions from $\mathbb{R}_+$ to $\mathbb{R}^3$ that are right continuous and have left limits. Since the projection mapping is continuous, X^h converges in law to X, see Billingsley (1968, Theorem 5.1). Define the process $\tilde{W} = \{\tilde{W}_t\}_{t \geq 0}$ by $\tilde{W}_t = w^m + e^{\rho t} X_t$ for all $t \geq 0$, and consider the discounted processes $\tilde{W}^{h,\rho} = \{\tilde{W}_t^{h,\rho}\}_{t \geq 0}$ and $\tilde{W}^\rho = \{\tilde{W}_t^\rho\}_{t \geq 0}$ defined by $\tilde{W}_t^{h,\rho} = (1 + \rho h)^{-[\![t/h]\!]} \tilde{W}_t^h$ and $\tilde{W}_t^\rho = e^{-\rho t} \tilde{W}_t$ for all $t \geq 0$. Note that $\tilde{W}_t^{h,\rho} = X_t^h + (1 + \rho h)^{-[\![t/h]\!]} w^{h,r}$ and $\tilde{W}_t^\rho = X_t + e^{-\rho t} w^m$ for all $t \geq 0$. Since $\lim_{h \downarrow 0} w^{h,r} = w^m$ by Corollary 1, and since the mapping $t \mapsto (1 + \rho h)^{-[\![t/h]\!]}$ converges uniformly to the mapping $t \mapsto e^{-\rho t}$ as h goes to 0, it follows then from Billingsley (1968, Theorem 4.1) that $\tilde{W}^{h,\rho}$ converges in law to $\tilde{W}^\rho$.

Step 2. We now examine what happens when one kills the processes $\tilde{W}^h$ and $\tilde{W}$, respectively at $w^{h,l}$ and 0. Our treatment builds on Toldo (2006, Lemma 1.13 and Proposition 1.14). We shall need the following deterministic lemma.

Lemma C.2. *Consider the functions from $D(\mathbb{R}_+)$ to $\mathbb{R}_+ \cup \{\infty\}$ defined by:*

$$f^h(y) = \inf\{t > 0 \,|\, y(t) < (1 + \rho h)^{-[\![t/h]\!]} w^{h,l}\}, \quad f(y) = \inf\{t > 0 \,|\, y(t) < 0\} \qquad \text{(C.16)}$$

for all $(h, y) \in \mathbb{R}_{++} \times D(\mathbb{R}_+)$, with the convention $\inf\{\emptyset\} = \infty$. Let $y \in C(\mathbb{R}_+)$ be such that:

$$f(y) = \inf\{t > 0 \,|\, y(t) \leq 0\}. \qquad \text{(C.17)}$$

Suppose that as h goes to 0, $y^h \in D(\mathbb{R}_+)$ converges to y in the local uniform topology. Then $\lim_{h \downarrow 0} f^h(y^h) = f(y)$ in $\mathbb{R}_+ \cup \{\infty\}$.

Proof. For each $h > 0$, let $t^h = f^h(y^h)$. Extracting a subsequence if necessary, define $t = \lim_{h \downarrow 0} t^h$ in $\mathbb{R}_+ \cup \{\infty\}$. We first show that $t \geq f(y)$. If $t = \infty$, this is immediate. If $t < \infty$ then, since y^h converges to y in the local uniform topology and $y \in C(\mathbb{R}_+)$, $\lim_{h \downarrow 0} y^h(t^h) = y(t)$. For h close enough to 0, one has $t^h < \infty$ so that $y^h(t^h) \leq (1 + \rho h)^{-[\![t_h/h]\!]} w^{h,l}$ as $y^h \in D(\mathbb{R}_+)$. Since $\lim_{h \downarrow 0} w^{h,l} = 0$ by Lemma 4, it follows that $y(t) \leq 0$. By (C.17), one then has $t \geq f(y)$, as claimed. Now suppose that $t > f(y)$. Two cases must be distinguished. If $t = \infty$, then there exists $\varepsilon > 0$ such that for all h close enough to 0, $y^h(s) \geq (1 + \rho h)^{-[\![s/h]\!]} w^{h,l} > 0$ for all $s \in [0, f(y) + \varepsilon]$. However, by (C.16), there exists some $s \in [f(y), f(y) + \varepsilon]$ such that $y(s) < 0$. Thus y^h cannot converge uniformly to y over $[0, f(y) + \varepsilon]$, a contradiction. If $t < \infty$, then by (C.16) again, there exists some $s \in [f(y), f(y) + \varepsilon]$ such that $y(s) < 0$, where $\varepsilon \in (0, [t - f(y)]/2)$ is fixed. Take $\eta = \varepsilon \wedge [-y(s)]$. Since $\lim_{h \downarrow 0} t^h = t$ and y^h converges to y in the local uniform topology, it follows that $|t^h - t| < \eta$ and $|y^h(s) - y(s)| < \eta$ for h close enough to 0. Fix any such h. On the one hand, one has $t^h - s > t - \eta - s \geq t - f(y) - 2\varepsilon > 0$, so that $t^h > s$. On the other hand, one has $y^h(s) - (1 + \rho h)^{-[\![s/h]\!]} w^{h,l} < y(s) + \eta \leq 0$, so that $y^h(s) < (1 + \rho h)^{-[\![s/h]\!]} w^{h,l}$ and $t^h \leq s$, a contradiction. Thus $t \leq f(y)$, and the result follows. $\|$

Since $\tilde{W}^{h,\rho}$ converges in law to $\tilde{W}^\rho$, Skorokhod's representation theorem implies that there exists a probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ and processes $\bar{W}^{h,\rho} = \{\bar{W}_t^{h,\rho}\}_{t \geq 0}$ and $\bar{W}^\rho = \{\bar{W}_t^\rho\}_{t \geq 0}$ inducing the same probability measures over $D(\mathbb{R}_+)$ as $\tilde{W}^{h,\rho}$ and $\tilde{W}^\rho$, and such that $\lim_{h \downarrow 0} \bar{W}^{h,\rho} = \bar{W}^\rho$ $\bar{\mathbb{P}}$–almost surely in $D(\mathbb{R}_+)$. From Jacod and Shiryaev (1987, Chapter VI, Proposition 1.17), since $\bar{W}^\rho$ has sample paths that are $\bar{\mathbb{P}}$–almost surely continuous, one has $\lim_{h \downarrow 0} \bar{W}^{h,\rho} = \bar{W}^\rho$ $\bar{\mathbb{P}}$–almost surely for the local uniform topology. Moreover, since $\bar{W}^\rho$ is a diffusion process, the stopping times $\inf\{t > 0 \,|\, \bar{W}_t^\rho < 0\}$ and $\inf\{t > 0 \,|\, \bar{W}_t^\rho \leq 0\}$ coincide $\bar{\mathbb{P}}$–almost surely. Applying Lemma C.2 path by path, it then follows that

$\lim_{h\downarrow 0} (\bar{W}^{h,\rho}, f^h(\bar{W}^{h,\rho})) = (\bar{W}^\rho, f(\bar{W}^\rho))$ $\bar{\mathbb{P}}$–almost surely. Using the fact that $\bar{W}^{h,\rho}$ and $\bar{W}^\rho$ induce the same probability measures over $D(\mathbb{R}_+)$ as $\tilde{W}^{h,\rho}$ and $\tilde{W}^\rho$, it is easy to check that this implies in turn that $(\tilde{W}^{h,\rho}, f^h(\tilde{W}^{h,\rho}))$ converges in law to $(\tilde{W}^\rho, f(\tilde{W}^\rho))$ over the space $D(\mathbb{R}_+) \times \mathbb{R}_+$. Defining the stopping times $\tilde{\tau}^h = f^h(\tilde{W}^{h,\rho})$ and $\tilde{\tau} = f(\tilde{W}^\rho)$ and the killed processes $\tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h} = \{\tilde{W}^{h,\rho}_{t\wedge \tilde{\tau}^h}\}_{t\geq 0}$ and $\tilde{W}^\rho_{.\wedge \tilde{\tau}} = \{\tilde{W}^\rho_{t\wedge \tilde{\tau}}\}_{t\geq 0}$, one obtains that $\tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}$ converges in law to $\tilde{W}^\rho_{.\wedge \tilde{\tau}}$.

Step 3. We now deal explicitly with the liquidation region $[0, w^{h,l})$. Define the discounted version $W^{h,\rho} = \{W^{h,\rho}_t\}_{t\geq 0}$ of W^h by $W^{h,\rho}_t = (1 + \rho h)^{-[\![t/h]\!]} W^h_t$ for all $t \geq 0$. By Proposition 1, $W^{h,\rho}$ is a supermartingale. Moreover, from Step 2, $\tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}$ converges in law to $\tilde{W}^\rho_{.\wedge \tilde{\tau}}$. It follows then from Billingsley (1968, Theorem 4.1) that, in order to establish that $W^{h,\rho}$ converges in law to $\tilde{W}^\rho_{.\wedge \tilde{\tau}}$, it is sufficient to prove that $\delta(W^{h,\rho}, \tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h})$ converges to 0 in probability, where δ is the standard Prokhorov metric on $D(\mathbb{R}_+)$ (Jacod and Shiryaev (1987, Chapter VI, Section 1c)). By construction, the processes $W^{h,\rho}$ and $\tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}$ coincide up to time $\tilde{\tau}^h$. Using the definition of δ, one then has, for any $\varepsilon > 0$:

$$\mathbb{P}\Big[\delta(W^{h,\rho}, \tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}) > \varepsilon\Big] \leq \mathbb{P}\Big[\sup_{t\in[\tilde{\tau}^h,\infty)} \{|W^{h,\rho}_t - \tilde{W}^{h,\rho}_{\tilde{\tau}^h}|\} > \varepsilon\Big]$$

$$\leq \mathbb{P}\Big[\sup_{t\in[\tilde{\tau}^h,\infty)} \{W^{h,\rho}_t\} > \varepsilon\Big] + \mathbb{P}\Big[\tilde{W}^{h,\rho}_{\tilde{\tau}^h} > \varepsilon\Big]$$

$$\leq \frac{\mathbb{E}\Big[W^{h,\rho}_{\tilde{\tau}^h}\Big]}{\varepsilon} + \mathbb{P}\Big[W^{h,\rho}_{\tilde{\tau}^h} > \varepsilon\Big],$$

where the first inequality follows from the definition of δ and the last one from the fact that $W^{h,\rho}$ is a supermartingale that coincides with $\tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}$ up to time $\tilde{\tau}^h$. Since $W^{h,\rho}_{\tilde{\tau}^h} \in [0, (1 + \rho h)^{-[\![t/h]\!]} w^{h,l})$ by construction and $\lim_{h\downarrow 0} w^{h,l} = 0$ by Lemma 4, it follows that $\lim_{h\downarrow 0} \mathbb{P}[\delta(W^{h,\rho}, \tilde{W}^{h,\rho}_{.\wedge \tilde{\tau}^h}) > \varepsilon] = 0$ for all $\varepsilon > 0$, which establishes that $W^{h,\rho}$ converges in law to $\tilde{W}^\rho_{.\wedge \tilde{\tau}}$. We now show that this implies that W^h converges in law to $\tilde{W}_{.\wedge \tilde{\tau}}$, where $\tilde{W}_{.\wedge \tilde{\tau}} = \{\tilde{W}_{t\wedge \tilde{\tau}}\}_{t\geq 0}$ and $\tilde{W}$ is defined as in Step 1. For any $t \geq 0$ and $h > 0$, let $a(t) = e^{\rho t}$ and $a^h(t) = (1 + \rho h)^{[\![t/h]\!]}$. By construction, $\tilde{W}_{.\wedge \tilde{\tau}} = a\tilde{W}^\rho_{.\wedge \tilde{\tau}}$ and $W^h = a^h W^{h,\rho}$. It is easy to check that the function from $D(\mathbb{R}_+)$ to $D(\mathbb{R}_+)$ that maps any $y \in D(\mathbb{R}_+)$ into the product ay is continuous. Since $W^{h,\rho}$ converges in law to $\tilde{W}^\rho_{.\wedge \tilde{\tau}}$, it follows from Billingsley (1968, Theorem 5.1) that $aW^{h,\rho}$ converges in law to $a\tilde{W}^\rho_{.\wedge \tilde{\tau}} = \tilde{W}_{.\wedge \tilde{\tau}}$. Thus, by Billingsley (1968, Theorem 4.1), all we need to prove to establish that W^h converges in law to $\tilde{W}_{.\wedge \tilde{\tau}}$ is that $\delta(W^h, aW^{h,\rho}) = \delta(W^h, aW^h/a^h)$ converges to 0 in probability. Using the definition of δ, it is easy to see that a sufficient condition for this to occur is that $\delta_{lu}(W^h, aW^h/a^h)$ converges to 0 in probability, where δ_{lu} is the metric associated to the local uniform topology on $D(\mathbb{R}_+)$ and defined by $\delta_{lu}(y, y') = \sum_{n=1}^\infty (1/2^n) \min\{\sup_{t\leq n} |y(t) - y'(t)|, 1\}$ for all $(y, y') \in D(\mathbb{R}_+)^2$. This in turn follows at once from the fact that the processes W^h are uniformly bounded for h close to 0, and that the function a/a^h converges to 1 in the local uniform topology as h goes to 0. Hence, W^h converges in law to $\tilde{W}_{.\wedge \tilde{\tau}}$, as claimed. To complete the proof, consider the processes $W = \{W_t\}_{t\geq 0}$ and $L = \{L_t\}_{t\geq 0}$ defined by $W_t = \tilde{W}_{t\wedge \tilde{\tau}}$ and $L_t = \int_0^{t\wedge \tilde{\tau}} e^{\rho s} dK_s$ for all $t \geq 0$, where K is given by (C.15). Note that the events $\{X_t = 0\}$ and $\{W_t = w^m\}$ coincide and that $\tau = \inf\{t \geq 0 \,|\, W_t = 0\} = \inf\{t \geq 0 \,|\, \tilde{W}_t < 0\} = \tilde{\tau}$ $\mathbb{P}$–almost surely. Since K has bounded variation and $\tilde{W}_t = w^m + e^{\rho t} X_t$ for all $t \geq 0$, it follows from (C.12)–(C.15) and a straightforward application of Itô's formula that (W, L) is the unique continuous solution to the Skorokhod problem (40)–(42), as required. Hence the result. Using Słomiński (1993, Corollary 11), one can further show that the pair (R^h, W^h) converges in law to (R, W) over $D(\mathbb{R}_+, \mathbb{R}^2)$, where R^h and R are defined by (38)–(39). ∥

Proof of Corollary 2. Define $\tilde{W}^{\rho}_{\cdot \wedge \tilde{\tau}}$ and $W^{h,\rho}$ as in Steps 2 and 3 of the proof of Proposition 5, and define the discounted version $M^{h,\rho} = \{M_t^{h,\rho}\}_{t \geq 0}$ of M^h by $M_t^{h,\rho} = (1 + \rho h)^{-\llbracket t/h \rrbracket} M_t^h$ for all $t \geq 0$. By construction, $m_{nh}^h = (\prod_{i=0}^{n-1} \hat{x}_{ih}^h) \hat{m}_{nh}^h = (\prod_{i=0}^{n-1} \hat{x}_{ih}^h) \hat{w}_{nh}^h / \lambda^h$ for all $n \in \mathbb{N}$. It follows then from (A.11) that $M^{h,\rho}$ is a supermartingale, just as $W^{h,\rho}$. Since $\lim_{h \downarrow 0} \lambda^h = \lambda$, Step 2 of the proof of Proposition 5 implies that $W^{h,\rho}/\lambda^h$ converges in law to $\tilde{W}^{\rho}_{\cdot \wedge \tilde{\tau}}/\lambda$. Thus, by Billingsley (1968, Theorem 4.1), $M^{h,\rho}$ converges in law to $\tilde{W}^{\rho}_{\cdot \wedge \tilde{\tau}}/\lambda$ if $\delta(M^{h,\rho}, W^{h,\rho}/\lambda^h)$ converges to 0 in probability. Let $\tau^h = \inf\{t > 0 \,|\, W_t^{h,\rho} < (1 + \rho h)^{-\llbracket t/h \rrbracket} w^{h,l}\}$. By (27)–(29) and (A.10)–(A.12), $M^{h,\rho}$ coincides with $W^{h,\rho}/\lambda^h$ up to time τ^h. Using the definition of δ, one then has, for any $\varepsilon > 0$:

$$\mathbb{P}\left[\delta\left(M^{h,\rho}, \frac{W^{h,\rho}}{\lambda^h}\right) > \varepsilon\right] \leq \mathbb{P}\left[\sup_{t \in [\tau^h, \infty)}\left\{\left|M_t^{h,\rho} - \frac{W_t^{h,\rho}}{\lambda^h}\right|\right\} > \varepsilon\right]$$

$$\leq \mathbb{P}\left[\sup_{t \in [\tau^h, \infty)}\{M_t^{h,\rho}\} > \varepsilon\right] + \mathbb{P}\left[\sup_{t \in [\tau^h, \infty)}\left\{\frac{W_t^{h,\rho}}{\lambda^h}\right\} > \varepsilon\right]$$

$$\leq \frac{2\mathbb{E}\left[W_{\tau^h}^{h,\rho}\right]}{\varepsilon \lambda^h},$$

where the first inequality follows from the definition of δ and the last one from the fact that $M^{h,\rho}$ and $W^{h,\rho}/\lambda^h$ are supermartingales that coincide up to time τ^h. Since $W_{\tau^h}^{h,\rho} \in [0, (1 + \rho h)^{-\llbracket t/h \rrbracket} w^{h,l}]$ by construction and $\lim_{h \downarrow 0} w^{h,l} = 0$ by Lemma 4, it follows that $\lim_{h \downarrow 0} \mathbb{P}\left[\delta(M^{h,\rho}, W^{h,\rho}/\lambda^h) > \varepsilon\right] = 0$ for all $\varepsilon > 0$, which establishes that $M^{h,\rho}$ converges in law to $\tilde{W}^{\rho}_{\cdot \wedge \tilde{\tau}}/\lambda$. Proceeding as in Step 3 of the proof of Proposition 5, one can then verify that this implies that M^h converges in law to W/λ. ‖

Proof of Proposition 6. Let $W_0 = w$ and suppose that $w \in [0, w^m]$. Since L is continuous with bounded variation, it follows from (40) and Itô's formula that, for each $T \geq 0$,

$$e^{-rT \wedge \tau} W_{T \wedge \tau} = w + \int_0^{T \wedge \tau} e^{-rt}(\rho - r)W_t \, dt + \int_0^{T \wedge \tau} e^{-rt}\lambda\sigma W_t \, dZ_t - \int_0^{T \wedge \tau} e^{-rt} \, dL_t. \tag{D.1}$$

We first take expectations in (D.1) and then let T go to ∞. Since $W_\tau = 0$ by definition of τ, and since the third term on the right-hand side of (D.1) is a martingale, we obtain that:

$$w = \mathbb{E}^{(w,0)}\left[\int_0^\tau e^{-rt}[dL_t - (\rho - r)W_t dt]\right],$$

where $\mathbb{E}^{(w,0)}$ is the expectation operator induced by the process (W, L) starting at $(w, 0)$. Dividing by λ and using the fact that $M = W/\lambda$, this implies that:

$$m = \mathbb{E}^{(m,0)}\left[\int_0^\tau e^{-rt}\left[\frac{1}{\lambda} dL_t - (\rho - r)M_t dt\right]\right], \tag{D.2}$$

where $m = w/\lambda$ and where $\mathbb{E}^{(m,0)}$ is the expectation operator induced by the process (M, L) starting at $(m, 0)$. From (44) and (D.2), and using the definitions (48) and (52) of S and B together with the fact that S and B are deterministic functions of M, we obtain that:

$$F(w) + m = (1 - \lambda)\mathbb{E}^{(m,0)}\left[\int_0^\tau e^{-rt}\frac{1}{\lambda} dL_t\right] + \mathbb{E}^{(m,0)}\left[\int_0^\tau e^{-rt}[\mu - (\rho - r)M_t] dt\right]$$

$$= (1 - \lambda)\mathcal{S}(m) + \mathcal{D}(m),$$

which implies the result. $\|$

Proof of Proposition 7. At any date $t \geq 0$, $S_t = \mathcal{S}(M_t)$, where $\mathcal{S}$ is given by (49)–(51). Note that the homogenous equation (49) is simply (B.2) with $\lambda = 1$. A basis $(\mathcal{H}_0, \mathcal{H}_1)$ of the space of solutions to (49) is characterized by the initial conditions $\mathcal{H}_0(0) = 1$, $\mathcal{H}_0'(0) = 0$, $\mathcal{H}_1(0) = 0$ and $\mathcal{H}_1'(0) = 1$. As for H_1 in the proof of Lemma 5, one can show that $\mathcal{H}_1' > 0$ over $\mathbb{R}_+$. Since $\mathcal{S}(0) = 0$ and $\mathcal{S}'(w^m/\lambda) = 1$, one must have $\mathcal{S}(m) = \mathcal{H}_1(m)/\mathcal{H}_1'(w^m/\lambda)$ for all $m \in [0, w^m/\lambda]$ and thus $\mathcal{S}$ is strictly increasing over $[0, w^m/\lambda]$. To complete the proof, we need only to show that $\mathcal{H}_1$ is strictly concave. Using the fact that $\mathcal{H}_1$ is smooth, one obtains by differentiation of (49):

$$r\mathcal{H}_1'(m) = \rho\mathcal{H}_1'(m) + \rho m \mathcal{H}_1''(m) + \frac{\sigma^2}{2}\mathcal{H}_1'''(m)$$

for all $m \geq 0$. Since $\rho > r$ and $\mathcal{H}_1' > 0$ over $\mathbb{R}_+$, it follows that $\mathcal{H}_1'''(m) < 0$ at any point $m \geq 0$ where $\mathcal{H}_1''(m) = 0$. Hence $\mathcal{H}_1''$ is strictly decreasing at its zeros over $\mathbb{R}_+$. Since $\mathcal{H}_1(0) = 0$ and $\mathcal{H}_1'(0) = 1$ imply that $\mathcal{H}_1''(0) = 0$ by (49), one thus has $\mathcal{H}_1'' < 0$ over $\mathbb{R}_{++}$, and the result follows. $\|$

Proof of Proposition 8. At any date $t \in [0, \tau)$, $D_t/S_t = \mathcal{D}(M_t)/\mathcal{S}(M_t)$, where $\mathcal{S}$ and $\mathcal{D}$ are given by (49)–(51) and (53)–(55). For any $m \in [0, w^m/\lambda]$, let $W_{\mathcal{S},\mathcal{D}}(m) = \mathcal{S}(m)\mathcal{D}'(m) - \mathcal{D}(m)\mathcal{S}'(m)$. To establish the result, we show that $W_{\mathcal{S},\mathcal{D}} < 0$ over $(0, w^m/\lambda)$. It is easy to check that:

$$\rho m W_{\mathcal{S},\mathcal{D}}(m) = -\frac{\sigma^2}{2} W_{\mathcal{S},\mathcal{D}}'(m) - [\mu - (\rho - r)m]\mathcal{S}(m)$$

for all $m \in [0, w^m/\lambda]$. Since $\mu - (\rho - r)m > 0$ and $\mathcal{S}(m) > 0$ for all $m \in (0, w^m/\lambda)$, it follows that $W_{\mathcal{S},\mathcal{D}}'(m) < 0$ at any point $m \in (0, w^m/\lambda)$ where $W_{\mathcal{S},\mathcal{D}}(m) = 0$. Hence $W_{\mathcal{S},\mathcal{D}}$ is strictly decreasing at its zeros over $(0, w^m/\lambda)$, if any exists. Since $W_{\mathcal{S},\mathcal{D}}(0) = 0$, we therefore need only to check that $W_{\mathcal{S},\mathcal{D}}$ is strictly negative in a right neighborhood of 0. From (49)–(50), one has:

$$\mathcal{S}''(0) = 0,$$

$$\mathcal{S}'''(0) = \frac{2(r - \rho)}{\sigma^2}\mathcal{S}'(0),$$

Similarly, from (53)–(54), one has:

$$\mathcal{D}''(0) = -\frac{2\mu}{\sigma^2},$$

$$\mathcal{D}'''(0) = \frac{2(r - \rho)}{\sigma^2}[\mathcal{D}'(0) - 1].$$

As a result of this, Taylor-Young expansions yield:

$$\mathcal{S}(m) = \mathcal{S}'(0)m + o(m^2),$$

$$\mathcal{S}'(m) = \mathcal{S}'(0) + \frac{r - \rho}{\sigma^2}\mathcal{S}'(0)m^2 + o(m^2),$$

and similarly:

$$\mathcal{D}(m) = \mathcal{D}'(0)m - \frac{\mu}{\sigma^2}m^2 + o(m^2),$$

$$\mathcal{D}'(m) = \mathcal{D}'(0) - \frac{2\mu}{\sigma^2}m + \frac{(r - \rho)}{\sigma^2}[\mathcal{D}'(0) - 1]m^2 + o(m^2).$$

45

It follows that $W_{\mathcal{S},\mathcal{D}}(m) = -\mu\mathcal{S}'(0)m^2/\sigma^2 + o(m^2)$, which is strictly negative in a right neighborhood of 0. Hence $W_{\mathcal{S},\mathcal{D}}(m) < 0$ for all $m \in (0, w^m/\lambda)$, so that $\mathcal{D}/\mathcal{S}$ is strictly decreasing over this interval. Since S is a strictly increasing function of M by Proposition 7, the result follows. ‖

Proof of Proposition 9. At any date $t \in [0, \tau)$, $B_t = \mathcal{B}(M_t) = [\mathcal{S}(M_0) + M_t - M_0]/\mathcal{S}(M_t)$, where $\mathcal{S}$ is given by (49)–(51). Thus, for any $m \in (0, w^m/\lambda)$, $\mathcal{B}'(m) = \phi(m)/\mathcal{S}(m)^2$, where by definition $\phi(m) = \mathcal{S}(m) - \mathcal{S}'(m)[\mathcal{S}(M_0) + m - M_0]$. Since $\mathcal{S}$ is strictly concave over $[0, w^m/\lambda]$ by Proposition 7, and since $\mathcal{S}'(w^m/\lambda) = 1$ by (51), one has $\mathcal{S}' > 1$ over $[0, w^m/\lambda)$. From this and $\mathcal{S}(0) = 0$, it follows that $\mathcal{S}(M_0) - M_0 > 0$, and thus $\phi(0) < 0$. Similarly, one has $\mathcal{S}(w^m/\lambda) - w^m/\lambda > \mathcal{S}(M_0) - M_0$, and from this and $\mathcal{S}'(w^m/\lambda) = 1$, it follows that $\phi(w^m/\lambda) > 0$. Furthermore, ϕ is strictly increasing as $\phi'(m) = -\mathcal{S}''(m)[\mathcal{S}(M_0) + m - M_0] > 0$ for all $m \in (0, w^m/\lambda)$. Hence, there exists a unique $m \in (0, w^m/\lambda)$ such that $\phi(m) = 0$, and $\mathcal{B}$ is U-shaped, as claimed. Since S is a strictly increasing function of M by Proposition 7, the result follows. ‖

Proof of Proposition 10. By (62), Δ_t is a strictly increasing and strictly convex function of T_t, that vanishes at 0. Moreover, $T_t = \mathcal{T}(M_t)$ at any date $t \in [0, \tau]$, where $\mathcal{T}$ is given by (63)–(65). To establish the result, we show that $\mathcal{T} > 0$ and $\mathcal{T}'' > 0$ over $[0, w^m/\lambda]$, and that $\mathcal{T}' < 0$ over $[0, w^m/\lambda)$. The homogenous equation (63) is the same as (49). One can thus write:

$$\mathcal{T}(m) = \beta_0 \mathcal{H}_0(m) + \beta_1 \mathcal{H}_1(m)$$

for all $m \in [0, w^m/\lambda]$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ are defined as in the proof of Proposition 7. To pin down the coefficients β_0 and β_1, we impose the boundary conditions $\mathcal{T}(0) = 1$ and $\mathcal{T}'(w^m/\lambda) = 0$. This yields:

$$\mathcal{T}(m) = \mathcal{H}_0(m) - \frac{\mathcal{H}_0'(w^m/\lambda)}{\mathcal{H}_1'(w^m/\lambda)} \mathcal{H}_1(m). \tag{D.3}$$

We now prove that $\mathcal{T}' < 0$ over $[0, w^m/\lambda)$. Note from (D.3) that $\mathcal{T}(w^m/\lambda) > 0$ since the Wronskian $W_{\mathcal{H}_0,\mathcal{H}_1} = \mathcal{H}_0\mathcal{H}_1' - \mathcal{H}_1\mathcal{H}_0'$ of $\mathcal{H}_0$ and $\mathcal{H}_1$ is strictly positive. It follows then from (63) and (65) that $\mathcal{T}''(w^m/\lambda) = 2r\mathcal{T}(w^m/\lambda)/\sigma^2 > 0$. Since $\mathcal{T}'(w^m/\lambda) = 0$, this implies that $\mathcal{T}' < 0$ over an interval $(w^m/\lambda - \varepsilon, w^m/\lambda)$ for $\varepsilon > 0$ small enough. Now suppose that $\mathcal{T}'(m) \geq 0$ for some $m < w^m/\lambda - \varepsilon$, and let $\tilde{m} = \sup\{m < w^m/\lambda - \varepsilon \mid \mathcal{T}'(m) \geq 0\}$. Then $\mathcal{T}(\tilde{m}) > 0$ since $\mathcal{T}' < 0$ over $(\tilde{m}, w^m/\lambda)$ and $\mathcal{T}(w^m/\lambda) > 0$. Moreover, $\mathcal{T}'(\tilde{m}) = 0$ and $\mathcal{T}''(\tilde{m}) \leq 0$. However, by (63), these two last conditions imply that $\mathcal{T}(\tilde{m}) \leq 0$, a contradiction. Hence $\mathcal{T}' < 0$ over $[0, w^m/\lambda)$, as claimed. Since $\mathcal{T}(w^m/\lambda) > 0$, it follows that $\mathcal{T} > 0$ over $[0, w^m/\lambda]$. Using (63) again, we obtain that $\mathcal{T}'' > 0$ over $[0, w^m/\lambda]$, which concludes the proof. ‖

Proof of Proposition 11. Since $W = \lambda M$, we shall hereafter use M, rather than W, as the state variable. Define accordingly a mapping $\mathcal{F} : \mathbb{R}_+ \to \mathbb{R}$ by $\mathcal{F}(m) = V(\lambda m) - \lambda m$ for all $m \geq 0$. Note that, by construction, $\mathcal{F}$ is strictly concave over $[0, w^m/\lambda]$. Fix the initial wealth A of the entrepreneur. Then, since the entrepreneur is assumed to have all the bargaining power at date 0, the investment capacity of the firm is $\max_{m \in \mathbb{R}_+}\{\mathcal{F}(m)\} + A$. Moreover, for a fixed investment outlay I, the initial level of cash reserves is $\max\{m \in \mathbb{R}_+ \mid \mathcal{F}(m) \geq I - A\}$, and $\mathcal{F}$ is decreasing at this point. We need therefore only to check that, over $(0, w^m/\lambda)$, $\mathcal{F}$ is strictly increasing with respect to μ, and strictly decreasing with respect to λ and σ^2. From (33)–(36), $\mathcal{F}$ satisfies:

$$r\mathcal{F}(m) = \mu + \rho m\mathcal{F}'(m) + \frac{\sigma^2}{2}\mathcal{F}''(m) \tag{D.4}$$

for all $m \in [0, w^m/\lambda]$, with boundary conditions:

$$\mathcal{F}(0) = 0, \tag{D.5}$$

$$\mathcal{F}'\left(\frac{w^m}{\lambda}\right) = -\lambda, \tag{D.6}$$

$$\mathcal{F}''\left(\frac{w^m}{\lambda}\right) = 0. \tag{D.7}$$

Consider first $\mathcal{F}_\mu = \partial \mathcal{F}/\partial \mu$. From (D.4)–(D.7), $\mathcal{F}_\mu$ satisfies:

$$r\mathcal{F}_\mu(m) = 1 + \rho m \mathcal{F}_\mu'(m) + \frac{\sigma^2}{2}\,\mathcal{F}_\mu''(m) \tag{D.8}$$

for all $m \in [0, w^m/\lambda]$, with boundary conditions:

$$\mathcal{F}_\mu(0) = 0, \tag{D.9}$$

$$\mathcal{F}_\mu'\left(\frac{w^m}{\lambda}\right) = 0. \tag{D.10}$$

Using Itô's formula as in the proof of Proposition 4, together with the dynamics (45) of M, one can check from (D.8)–(D.10) that $\mathcal{F}_\mu(m) = \mathbb{E}\left[\int_0^\tau e^{-rt}\,dt\right]$ for all $m \in [0, w^m/\lambda]$, where $\mathbb{E}^{(m,0)}$ is the expectation operator induced by the process (M, L) starting at $(m, 0)$. Thus $\mathcal{F}_\mu(m) > 0$ for all $m \in (0, w^m/\lambda)$, and $\mathcal{F}$ is strictly increasing with respect to μ over $(0, w^m/\lambda)$.

Consider next $\mathcal{F}_\lambda = \partial \mathcal{F}/\partial \lambda$. From (D.4)–(D.7), $\mathcal{F}_\lambda$ satisfies:

$$r\mathcal{F}_\lambda(m) = \rho m \mathcal{F}_\lambda'(m) + \frac{\sigma^2}{2}\,\mathcal{F}_\lambda''(m) \tag{D.11}$$

for all $m \in [0, w^m/\lambda]$, with boundary conditions:

$$\mathcal{F}_\lambda(0) = 0, \tag{D.12}$$

$$\mathcal{F}_\lambda'\left(\frac{w^m}{\lambda}\right) = -1. \tag{D.13}$$

Comparing (D.11)–(D.13) with the boundary value problem (49)–(51) solved by $\mathcal{S}$, it is immediate that $\mathcal{F}_\lambda(m) = -\mathcal{S}(m)$ for all $m \in [0, w^m/\lambda]$. Thus $\mathcal{F}_\lambda(m) < 0$ for all $m \in (0, w^m/\lambda)$, and $\mathcal{F}$ is strictly decreasing with respect to λ over $(0, w^m/\lambda)$.

Consider finally $\mathcal{F}_{\sigma^2} = \partial \mathcal{F}/\partial \sigma^2$. From (D.4)–(D.7), $\mathcal{F}_{\sigma^2}$ satisfies:

$$r\mathcal{F}_{\sigma^2}(m) = \frac{1}{2}\,\mathcal{F}''(m) + \rho m \mathcal{F}_{\sigma^2}'(m) + \frac{\sigma^2}{2}\,\mathcal{F}_{\sigma^2}''(m) \tag{D.14}$$

for all $m \in [0, w^m/\lambda]$, with boundary conditions:

$$\mathcal{F}_{\sigma^2}(0) = 0, \tag{D.15}$$

$$\mathcal{F}_{\sigma^2}'\left(\frac{w^m}{\lambda}\right) = 0. \tag{D.16}$$

In analogy with $\mathcal{F}_\mu$, one can check from (D.14)–(D.16) that $\mathcal{F}_{\sigma^2}(m) = \mathbb{E}^{(m,0)}\left[\int_0^\tau e^{-rt}\mathcal{F}''(M_t)\,dt\right]/2$ for all $m \in [0, w^m/\lambda]$. Moreover, $\mathcal{F}$ is strictly concave over $[0, w^m/\lambda]$. Thus $\mathcal{F}_{\sigma^2}(m) < 0$ for all $m \in (0, w^m/\lambda)$, and $\mathcal{F}$ is strictly increasing with respect to σ^2 over $(0, w^m/\lambda)$. The result follows. ∥

Proof of Proposition 12. Since Δ_t is a strictly increasing function of T_t by (62), and $T_t = \mathcal{T}(M_t)$ at any date $t \in [0, \tau]$, we need only to check that, over $(0, w^m/\lambda)$, $\mathcal{T}$ is strictly decreasing with respect to μ, and strictly increasing with respect to λ. Note that $\mathcal{T}$ depends on μ and λ only through w^m/λ. We first determine how w^m/λ varies with μ and λ. Define H_1 as in the proof of Lemma 5, and $\mathcal{H}_1$ as in the proof of Proposition 7. From the proof of Lemma 5, w^m is the unique solution to the equation:

$$[\rho w H_1'(w) - r H_1(w)] \exp\left(\frac{\rho w^2}{\lambda^2 \sigma^2}\right) = \mu.$$

As a result of this, w^m/λ is the unique solution to the equation:

$$[\rho \lambda m H_1'(\lambda m) - r H_1(\lambda m)] \exp\left(\frac{\rho m^2}{\sigma^2}\right) = \mu. \tag{D.17}$$

Define $\mathfrak{H}_1(m) = H_1(\lambda m)$ for all $m \geq 0$. Then, since H_1 solves (B.2), $\mathfrak{H}_1$ solves (49), just as $\mathcal{H}_1$. Moreover, since $\mathfrak{H}_1(0) = \mathcal{H}_1(0) = 0$, and since $\mathcal{H}_1'(0) = 1$ and $\mathfrak{H}_1'(0) = \lambda$, one must have $\mathfrak{H}_1 = \lambda \mathcal{H}_1$. Thus (D.17) can be rewritten as:

$$[\rho m \mathcal{H}_1'(m) - r \mathcal{H}_1(m)] \exp\left(\frac{\rho m^2}{\sigma^2}\right) = \frac{\mu}{\lambda}. \tag{D.18}$$

The left-hand side of (D.18) is a strictly increasing and strictly convex function of m, that vanishes at 0. Since $\mathcal{H}_1$ is independent of μ and λ, it is immediate from (D.18) that w^m/λ is strictly increasing with respect to μ and strictly decreasing with respect to λ. Finally, let $\mathcal{T}_{w^m/\lambda} = \partial \mathcal{T}/\partial(w^m/\lambda)$. From (63)–(65), $\mathcal{T}_{w^m/\lambda}$ satisfies:

$$r \mathcal{T}_{w^m/\lambda}(m) = \rho m \mathcal{T}_{w^m/\lambda}'(m) + \frac{\sigma^2}{2} \mathcal{T}_{w^m/\lambda}''(m) \tag{D.19}$$

for all $m \in [0, w^m/\lambda]$, with boundary conditions:

$$\mathcal{T}_{w^m/\lambda}(0) = 0, \tag{D.20}$$

$$\mathcal{T}_{w^m/\lambda}'\left(\frac{w^m}{\lambda}\right) = -\mathcal{T}''\left(\frac{w^m}{\lambda}\right). \tag{D.21}$$

Comparing (D.19)–(D.21) with the boundary value problem (49)–(51) solved by $\mathcal{S}$, it is immediate that $\mathcal{T}_{w^m/\lambda}(m) = -\mathcal{T}''(w^m/\lambda)\,\mathcal{S}(m)$ for all $m \in [0, w^m/\lambda]$. Thus $\mathcal{T}_{w^m/\lambda}(m) < 0$ for all $m \in (0, w^m/\lambda)$, and $\mathcal{T}$ is strictly decreasing with respect to w^m/λ over $(0, w^m/\lambda)$. The result then follows from the fact that w^m/λ is strictly increasing with respect to μ, and strictly decreasing with respect to λ. ‖

Proof of Proposition 13. Since $S_t = \mathcal{S}(M_t)$ at any date $t \geq 0$, we need only to check that, over $(0, w^m/\lambda]$, $\mathcal{S}$ is strictly increasing with respect to μ, and strictly decreasing with respect to λ. As shown in the proof of Proposition 7, one has $\mathcal{S}(m) = \mathcal{H}_1(m)/\mathcal{H}_1'(w^m/\lambda)$ for all $m \in [0, w^m/\lambda]$, where $\mathcal{H}_1$ is a positive, strictly increasing and strictly concave function. Thus $\mathcal{S}$ is strictly increasing with respect to w^m/λ over $(0, w^m/\lambda]$. The result then follows from the fact that w^m/λ is strictly increasing with respect to μ, and strictly decreasing with respect to λ, as shown in the proof of Proposition 12. ‖

REFERENCES

ALBUQUERQUE, R., and HOPENHAYN, H. (2004), "Optimal Lending Contracts and Firm Dynamics", *Review of Economic Studies*, **71**, 285–315.

ASQUITH, P., BEATTY, A. and WEBER, J. (2005), "Performance Pricing in Debt Contracts", *Journal of Accounting and Economics*, **40**, 101–128.

BILLINGSLEY, P. (1968) *Convergence of Probability Measures* (New York, John Wiley & Sons).

BLACK, F. (1976), "Studies of Stock Market Volatility Changes", *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

BLACK, F. and SCHOLES, M. (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, **81**, 637–654.

BOLTON, P. and SCHARFSTEIN, D. (1990), "A Theory of Predation Based on Agency Problems in Financial Contracting", *American Economic Review*, **80**, 93–106.

BREALEY, R. and MYERS, S. (2000) *Principles of Corporate Finance*, 6th edition (New York: Irwin McGraw-Hill).

BULAN, L., SUBRAMANIAM, N. and TANLU, L. (2005), "On the Timing of Dividend Initiations", (Mimeo, Harvard Business School).

CLEMENTI, G.L. and HOPENHAYN, H. (2006), "A Theory of Financing Constraints and Firm Dynamics", *Quarterly Journal of Economics*, **121**, 229-265.

COLES, J.L., LEMMON, M.L. and MESCHKE, J.F., (2006), "Structural Models and Endogeneity in Corporate Finance: The Link between Managerial Ownership and Corporate Performance", (Mimeo, Arizona State University).

DEANGELO, H., DEANGELO, L. and STULZ, R.M. (2006), "Dividend Policy and the Earned/Contributed Capital Mix: A Test of the Life-Cycle Theory", *Journal of Financial Economics*, **81**, 227–254.

DEMARZO, P.M. and FISHMAN, M.J. (2003), "Optimal Long-Term Financial Contracting with Privately Observed Cash Flows" (Mimeo, Stanford University).

DEMARZO, P.M. and SANNIKOV, Y. (2006), "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model", *Journal of Finance*, forthcoming.

DEMSETZ, H. and VILLALONGA, B. (2001), "Ownership Structure and Corporate Performance", *Journal of Corporate Finance*, **7**, 209–233.

DENIS, D.K. and SHOME, D.K. (2005), "An Empirical Investigation of Corporate Asset Downsizing", *Journal of Corporate Finance*, **11**, 427–448.

DIAL, J. and MURPHY, K.J. (1995), "Incentives, Downsizing, and Value Creation at General Dynamics", *Journal of Financial Economics*, **37**, 261–314.

GREEN, E. (1987), "Lending and the Smoothing of Uninsurable Income", in E.C. Prescott and N. Wallace (eds.) *Contractual Arrangements for Intertemporal Trade*, (Minneapolis: University of Minnesota Press).

GROMB, D. (1999), "Renegotiation in Debt Contracts" (Mimeo, Massachusetts Institute of Technology).

HELLWIG, M.F. and SCHMIDT, K.M. (2002), "Discrete-Time Approximations of the Holmström-Milgrom Brownian-Motion Model of Intertemporal Incentive Provision", *Econometrica*, 70, 2225–2264.

HIMMELBERG, C.P., HUBBARD, R.G. and PALIA, D. (1999), "Understanding the Determinants of Managerial Ownership and the Link between Ownership and Performance", *Journal of Financial Economics*, **53**, 353–384.

HOLMSTRÖM, B. and MILGROM, P. (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica*, 55, 303–328.

HOLMSTRÖM, B. and TIROLE, J. (1997), "Financial Intermediation, Loanable Funds, and the Real Sector", *Quarterly Journal of Economics*, **112**, 663–691.

INNES, R. (1990), "Limited Liability and Incentive Contracting with Ex-Ante Action Choices", *Journal of Economic Theory*, **52**, 45–67.

JACOD, J. and SHIRYAEV, A.N. (1987) *Limit Theorems for Stochastic Processes* (Berlin, Heidelberg, New York: Springer-Verlag).

JEANBLANC-PICQUÉ, M. and SHIRYAEV, A.N. (1995), "Optimization of the Flow of Dividends", *Russian Mathematical Surveys*, **50**, 257–277.

KAPLAN, S.N. and STRÖMBERG, P. (2003), "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts", *Review of Economic Studies*, **70**, 281–316.

KAPLAN, S.N. and STRÖMBERG, P. (2004), "Characteristics, Contracts and Actions: Evidence from Venture Capitalist Analyses", *Journal of Finance*, **59**, 2177–2210.

KARATZAS, I. and SHREVE, S.E. (1991): *Brownian Motion and Stochastic Calculus*, 2nd edition (Berlin, Heidelberg, New York: Springer-Verlag).

LANDO, D. and MORTENSEN A. (2004), "On the Pricing of Step-Up Bonds in the European Telecom Sector", *Journal of Credit Risk*, **1**, 71–110.

MODIGLIANI, F. and MILLER, M. (1958), "The Cost of Capital, Corporation Finance, and the Theory of Investment", *American Economic Review*, **48**, 261–297.

MURPHY, K.J. (1999), "Executive Compensation", in O.C. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 3B (Amsterdam: North-Holland).

NELSON, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, **59**, 347–370.

QUADRINI, V. (2004), "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard", *Journal of Monetary Economics*, **51**, 713–751.

RADNER, R. and SHEPP, L. (1996), "Risk vs. Profit Potential: A Model for Corporate Strategy", *Journal of Economic Dynamics and Control*, **20**, 1373–1393.

SANNIKOV, Y. (2003), "A Continuous-Time Version of the Principal-Agent Problem" (Mimeo, Stanford University).

SHIM, I. (2006), "Dynamic Prudential Regulation: Is Prompt Corrective Action Optimal?" (BIS Working Paper No. 206, Bank for International Settlements, Basel).

SŁOMIŃSKI, L. (1993), "On Existence, Uniqueness and Stability of Solutions of Multidimensional SDE's with Reflecting Boundary Conditions", *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, **29**, 163–198.

SPEAR, S. and SRIVASTAVA, S. (1987), "On Repeated Moral Hazard with Discounting", *Review of Economic Studies*, **54**, 599–607.

TANAKA, H. (1979), "Stochastic Differential Equations with Reflecting Boundary Condition in Convex Regions", *Hiroshima Mathematical Journal*, **9**, 163–177.

TCHISTYI, A. (2006), "Security Design with Correlated Hidden Cash Flows: The Optimality of Performance Pricing" (Mimeo, New York University).

THOMAS, J. and WORRALL, T. (1990), "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem", *Journal of Economic Theory*, **51**, 367–390.

TOLDO, S. (2006), "Stability of Solutions of BSDEs with Random Terminal Time", *ESAIM: Probability and Statistics*, **10**, 141–163.

WELCH, I. (2004), "Capital Structure and Stock Returns", *Journal of Political Economy*, **112**, 106–131.