Financing Innovation under Ambiguity and Skewness

April 1, 2022

Abstract

This paper extends the financing of a growth option framework to analyze the fostering effect of debt on innovation per recent empirical evidence. The main novelty of our paper is to extend the recursive multiple-prior utility framework with jump ambiguity to characterize the valuation of innovation returns, which are ambiguous and skewed. We utilize detection error probability to determine the size of the equivalent priors based on the amount of information available to investors. We specify the EBIT process as a double exponential jump-diffusion process to obtain analytical solutions. We first show that ambiguity can shape agents’ preferences for exploration or exploitation. We characterize the EBIT process of an exploration project with infrequent big jumps and that of an exploitation project with frequent small jumps. We show that for an exploration project and an exploitation project with identical NPV, if initiated immediately, agents will prefer the exploitation project when ambiguity elevates. Debt can accelerate project initiation and increase project value due to the associated net tax benefits. This fostering effect is more substantial for explorations with negative ex ante skewness, novel projects, and firms with lower agency costs. The impact of debt depends on innovation type and the relative concern for diffusion ambiguity or jump ambiguity.

Key Words— Innovation, growth option, capital structure, jump ambiguity.

JEL code— G32.
1 Introduction

Investing in innovation is challenging because its outcomes are highly ambiguous, and its returns are extremely skewed (Kerr and Nanda, 2015; Scherer and Harhoff, 2000). A well-functioning financial market is essential in driving economic growth by spurring innovation (e.g., Hsu et al., 2014). A traditional view of financial markets’ role is to allocate capital to firms with the highest potential to implement new processes and to commercialize new technologies, and much of the focus is on how financial constraints for entrepreneurs can hamper innovation activities (e.g., Beck et al., 2008).

However, a recent view is that financial markets might affect innovation by shaping the nature of the R&D being undertaken (Hall and Lerner, 2010). Earlier studies show that internal cash flow and external equity are the primary funding sources for innovation (e.g., Brown et al., 2009; Acharya and Xu, 2017). However, recent evidence shows that public equity financing can adversely impact the rate and trajectory of innovation, despite the fact that equity can better align the interests of investors and innovators and is more liquid.\(^1\) In contrast, growing evidence shows that debt is an important funding source, even for innovating startups.\(^2\) Furthermore, debt can increase the rate and novelty of innovation.\(^3\)

Given the vital role of innovation in the economy, it is critical to understand why debt financing can promote innovation. Prior studies (e.g., Aghion et al., 2013) primarily view the use of debt as a second-best outcome under asymmetric information, which follows that the inputs to innovation are hard to measure (e.g., Holmstrom, 1989) and the outcomes are hard to contract ex ante (e.g., Aghion and Tirole, 1994). Hence, the associated agency conflicts drive the adverse effects of equity financing on innovation because executives may enjoy their quiet lives by reducing the rate and risk of R&D projects when facing governance pressures from diverse external shareholders (e.g., Holmström, 1999). In contrast, banks can benefit from their superior monitoring abilities (Diamond, 1984), thereby partially alleviating managerial incentive constraints. Nevertheless, this view is more relevant to publicly traded firms owned by dispersed shareholders. This kind of agency conflicts should be less binding in private firms with more concentrated ownership and especially in small and mid-size enterprises where the founders hold the majority ownership.

Instead of resorting to the second-best, the first novelty of our study is to postulate that

---

\(^1\)For example, Bernstein (2015) finds that innovation novelty falls after firms file for initial public offerings (IPOs).

\(^2\)For instance, Mann (2018) documents that it is common for innovating firms to borrow and use patents as collateral. Hochberg et al. (2018) document a widespread use of loans to finance technology startups, even in early stages of development. Davis et al. (2020) highlight the role of venture debt for startups.

\(^3\)For instance, Chava et al. (2013) and Robb and Robinson (2014) find that firms innovate more and take riskier projects when the cost of debt falls due to deregulations in the banking industry.
the use of debt is a first-best outcome under perfect information. To characterize the investment under uncertainty nature of innovation and its interaction with the funding choice, our benchmark model integrates the financing of a growth option framework introduced by Sundaresan and Wang (2007) and Sundaresan et al. (2015) with the agency framework of Morellec et al. (2012).\textsuperscript{4} In this benchmark, an entrepreneur decides the timing to initiate an innovation project and its funding source. The entrepreneur can finance the project with external equity or an optimal mix of equity and debt where the optimality trades off the tax benefits of debt and the bankruptcy cost. As in Morellec et al. (2012), agency cost manifests itself through partial managerial ownership and private benefit.

The main novelty of our study is to introduce jump ambiguity to characterize innovation returns, which are ambiguous and skewed. While the literature offers alternative ways to model ambiguity and ambiguity aversion, to adapt to our benchmark model, we extend the original recursive multiple priors utility (RMPU) framework of Chen and Epstein (2002) by introducing jump ambiguity.\textsuperscript{5} Specifically, we model agents’ prior about the EBIT from the project as a set of equivalent Lévy processes based on the mathematical foundations of Quenez and Sulem (2013, 2014). In contrast to the original framework, where the multiple equivalent priors only differ in the drift component, our generalized framework allows both the drift and jump components to differ across the priors. To gain further insights, we specify the jump size distribution to be of the double exponential type such that corresponding option values have closed-form solutions (Kou and Wang, 2004). In this sense, we can explicitly characterize the ambiguous normal increments and rare increments, and we have full control over the skewness and kurtosis of the EBIT distribution. To assess the effect of learning on (resolving) ambiguity, we utilize relative entropy and detection error probability introduced by Anderson et al. (2003) and further developed by Aït-Sahalia and Matthys (2019) to determine the level of ambiguity that agents are concerned with based on their relevant histories of data.

Before we present our main results, it is necessary to sketch the structure of the priors. Recall that under drift ambiguity (κ-ignorance of Chen and Epstein, 2002), the set of equivalent priors can be generated by distorting the reference measure with a set of density generators. Consequently, the state process under an equivalent measure differs from that under the reference measure by a drift term according to Girsanov’s theorem. In contrast, a density generator has two components in the case of jump ambiguity, one for the Brownian motion and one for

\textsuperscript{4}In the agency framework of Morellec et al. (2012), there is no information asymmetry. The main effect of the agency conflict is to lower the optimal leverage.

\textsuperscript{5}In the original RMPU framework, agents are endowed with a set of priors characterized by equivalent Itô diffusions (in the sense of Girsanov’s theorem) and agents maximizing expected utility under the worst-case prior. However, this framework is inadequate for analyzing innovation because the increments of diffusions are symmetric and normally distributed, and it is widely acknowledged that the returns from R&D processes are highly skewed (Scherer and Harhoff, 2000).
the Poisson random measure. Hence, the state process under an equivalent measure differs by both the drift and jump terms, according to Girsanov’s theorem for Lévy processes. Now the drift term is affected by both components of the density generator. In the case of the double exponential jump-diffusion process, the jump component differs in both the jump size distribution and jump intensity.

Next, the worst-case pricing measures for equity investors and debt investors coincide and have the following characteristics. The drift is adjusted downward due to both the diffusion and jump distortions. The interpretation is similar to that in Nishimura and Ozaki (2007). For the jump part, the conditional mean positive jump size is the smallest possible, and the conditional mean negative jump size is the largest possible in absolute value. This follows from the fact that both equity investors and debt investors receive cash flows from the project. Hence the worst-case scenario is when positive jumps have the smallest possible probability of occurring and mean magnitude, and negative ones have the opposite. Next, the jump intensity is intermediate because investors perceive the lowest possible arrival rate of positive jumps and the largest possible arrival rate of negative jumps.

Our first main result is that ambiguity can shape agents’ preferences for exploration or exploitation. We characterize the EBIT of an exploration project with infrequent “big” jumps and that of an exploitation project with frequent “small” jumps. We show that ambiguity increases the optimal investment boundary, lowers the Arrow-Debreu (AD) price for investment and project value, and these effects are stronger for exploration than exploitation. Hence, given an exploration project and an exploitation project with identical NPV when initiated immediately, agents facing high ambiguity are more likely to choose the latter. The reason is that the optimal investment boundary is lower while the associated AD price and project value are higher for the latter.

Importantly, we demonstrate that financing innovation with an optimal mix of debt and equity is the first-best outcome under complete information. We begin by showing that debt financing can promote innovation in general by accelerating the project’s initiation and increasing its value with the net tax benefits of debt relative to all equity financing. In other words, at any EBIT level, the project value under optimal financing always dominates that under equity financing, hence for any predetermined NPV target, the required EBIT threshold is always lower under optimal financing, making investment happen sooner. Next, we show that the investment acceleration and value-enhancing effects of debt are stronger for exploration projects, especially those perceived to have negatively skewed returns \textit{ex ante}. The result on innovation type follows the fact that the proportional net tax benefits are higher for exploration projects, and the result on \textit{ex ante} skewness follows the fact that return variance and kurtosis drops faster

\footnote{Precisely, the AD price for investment is the Laplace transform of the optimal stopping time for investment.}
when ambiguity increases under positive *ex ante* skewness. In addition, we show that debt financing is more beneficial to relatively novel projects characterized by limited available information, and this result follows the fact that debt’s investment acceleration effect is stronger for projects with less available information. Next, the added value from debt may further depend on the type of innovation and the relative concern for diffusion ambiguity or jump ambiguity, as well as the relative magnitudes of the diffusion increments and jump increments. Lastly, agency concerns can lower the added value from debt, because agency cost lowers the optimal leverage.

Moreover, our results bring new challenges to the classic statement that real option value increases in the uncertainty of the project value dynamics. This classic statement is derived under the assumption that the project value dynamics follow geometric Brownian motions (e.g., McDonald and Siegel, 1986). In this case, the only uncertainty is diffusion risk. When we consider only drift ambiguity like Nishimura and Ozaki (2007), the classic statement needs to be restricted to the return variance of the project value dynamics, because ambiguity lowers the project value. Under jump ambiguity, the classic statement needs to be restricted to the variance of the continuous component of the project value dynamics, because both jump risk and jump ambiguity can contribute to return variance of the project value dynamics and affect project values in both ways.

The primary contribution of our paper is to the growing literature that analyzes the effects of debt on innovation. A close paper is Geelen et al. (2021) who also develop a model under perfect information. Their model features both incumbents and entrants, focusing on the implications at the aggregate level. The primary economic mechanism of debt is two-fold in their case. On the one hand, too much debt can hamper innovation by incumbents due to debt overhang; on the other hand, moderate debt can stimulate innovation by entrants due to debt’s net tax benefits. Their aggregation mechanism leads to the stimulating effect dominating the hampering effect; hence, debt fosters innovation and growth at the aggregate level. In contrast, our model can be viewed as an in-depth analysis of the entrant’s decision compared to their model. Similarly, the fostering effect of debt also stems from the net tax benefits of debt in our model. However, different from them, we further examine the interactions of debt with ambiguity and skewness, the defining features of innovation outcomes and returns. Moreover, we highlight that the fostering effect of debt depends on the type of innovation (exploration vs. exploitation), the relative novelty of innovation, the relative concern of diffusion ambiguity or jump ambiguity, and agency cost.

Our paper also contributes to the literature analyzing the effects of ambiguity on real option values. Nishimura and Ozaki (2007) extends the standard irreversible investment model by assuming the EBIT generated by the investment to be ambiguous with the RMPU preference.
of Chen and Epstein (2002). Miao and Wang (2011) further distinguish between the ambiguity about continuation values and terminal payoffs and show the relative degrees of them can either accelerate or delay option exercise under a general assumption about the state process. Schröder (2011) uses $\alpha$-maxmin expected utility as the preference to study an irreversible investment problem. Flor and Hesel (2015) examines the effects of ambiguous project value and cost under a two-dimensional setting using the robust framework of Anderson et al. (2003). We differ from these studies by studying explicitly the role of jump ambiguity and highlight that real option value increases in return variance only when the change in the variance is due to an increase in the variance of the continuous component of return.

Another contribution of our paper is the characterization of jump ambiguity. Drechsler (2013), Aït-Sahalia and Matthys (2019), and Jin et al. (2021) study the dynamic intertemporal portfolio choice problem under the assumption that asset prices follow an ambiguous jump-diffusion process. Different from us, Drechsler (2013) specifies the jump size distribution as normal distribution or gamma distribution, Aït-Sahalia and Matthys (2019) choose beta distribution, and Jin et al. (2021) considers normal distribution and one-sided power law distribution. Next, Aït-Sahalia and Matthys (2019) find that the jump intensity is lower and the conditional jump size is larger under the worst-case measure. Benefiting from our specification of the double-exponential distribution for the jump size, we further show that the lowered jump intensity is due to fewer positive jumps and the larger jump size is due to negative jumps.

2 Financing of innovation under ambiguity aversion

2.1 Characterize ambiguous innovation returns

An entrepreneur has access to an innovation project. Once started, it will generate an EBIT flow $X(t) := X(t, \omega)$ defined on a probability space $(\Omega, \mathcal{F}, Q^0)$ endowed with a standard complete filtration $\mathcal{F} = \{\mathcal{F}_t | t \geq 0\}$. Under the reference risk neutral measure $Q^0$, $X(t)$ follows a geometric Lévy process:

$$dX(t)/X(t^-) = \mu dt + \sigma dW(t) + \int_{\mathbb{R}} \nu(t, u) \tilde{N}(dt, du),$$

where $W(t)$ is a standard Brownian motion, $\tilde{N}(dt, du)$ is a compensated Poisson random measure given by

$$\tilde{N}(dt, du) = N(dt, du) - \nu(du) dt,$$

with $\nu(du)$ being the Lévy measure, and $\nu(t, u)$ is a square-integrable predictable process with respect to $\nu(du)$. Furthermore, we assume that $X(0) = x > 0$, $\mu, \sigma$, and $r$ are constants with $\mu < r$, the risk-free rate.
In this paper, we specify the jump component as a compound Poisson process with intensity \( \lambda < +\infty \) and a double exponential jump size distribution. In another word, \( X(t) \) is a double exponential jump-diffusion process, first introduced by Kou (2002). Hence, we can write the jump component explicitly as

\[
\iota(t, u) = e^u - 1, \quad \text{and} \quad \nu(du) = \lambda f(du),
\]

where the jump size density \( f \) is

\[
f_u = p\eta_1 e^{-\eta_1 u} 1_{u \geq 0} + q\eta_2 e^{\eta_2 u} 1_{u < 0}, \quad \eta_1 > 1, \eta_2 > 0, \quad p, q \geq 0, \quad p + q = 1.
\]

This process’s primary merit is the availability of analytical solutions for perpetual options with the American exercise style, as shown by Kou and Wang (2004). This property is key to our analyses of the investment and financing problems. Its second merit is its explicit control over positive jumps and negative jumps, and this control will generate a rich set of equivalent priors, as we show later.

The entrepreneur does not have the perfect knowledge of \( X(t) \), and he is averse to this ambiguity. To characterize ambiguity aversion, we extend the recursive multiple-prior utility framework of Chen and Epstein (2002) by allowing ambiguous jump intensity and size distribution (or the Lévy measure). We utilize the general results from Quenez and Sulem (2013, 2014), who provide the comparison theorem for backward differential equations (BSDEs) under Lévy process and the general results for related optimal stopping problems under ambiguity.

Let \( \Theta \) denote the space of density generators. For each \( \theta \in \Theta \), let \( Z^\theta(t) \) be the solution of the (forward) SDE:

\[
dZ^\theta(t) = Z^\theta(t^-) \left[ -\theta_W(t) dW(t) - \int_\mathbb{R} \theta_N(t, u) d\tilde{N}(dt, du) \right], \quad \theta = (\theta_W, \theta_N) \in \Theta,
\]

for \( t \in [0, T] \) and \( Z^\theta(0) = 1 \). For \( \theta_W(t) \), we allow it to be in \([-\kappa, \kappa] \) for \( 0 < \kappa < \infty \). Chen and Epstein (2002) term this specification as \( \kappa \)-ignorance. Based on the technical requirement in Quenez and Sulem (2013, 2014), we make the following specification for \( \theta_N(t, u) \):

\[
\theta_N(t, u) = 1 - e^{\theta_{N,1}(t) u} 1_{u \geq 0} - e^{\theta_{N,2}(t) u} 1_{u < 0},
\]

\[
\theta_{N,1}(t) \in [-M_1, 0], \quad \theta_{N,2}(t) \in [0, M_2], \quad \text{and} \quad M_1, M_2 > 0.
\]

Our specification for \( \theta_N(t, u) \) naturally follows the jump size distribution given by (2.3). As it will become clear later, we distort positive jumps and negative jumps separately.
We define an equivalent pricing measure \( Q^\theta \sim Q^0 \) on \( \mathcal{F}_T \) for \( \theta \in \Theta \) as

\[
E^\theta[1_A] = E[1_A Z^0_T], \quad A \in \mathcal{F}_T.
\]

Hence, we have under \( Q^\theta \)

\[
dW^\theta(t) = dW(t) + \theta_W(t)dt
\]

being a Brownian motion and

\[
\tilde{N}^\theta(dt,du) = \tilde{N}(dt,du) + \theta N(t,u)\nu(du)dt = N(dt,du) - (1 - \theta N(t,u))\nu(du)dt
\]

being a compensated Poisson random measure by Girsanov’s theorem.\(^7\) Therefore, under \( Q^\theta \), \( X(t) \) is given by

\[
dX(t)/X(t^-) = \left( \mu - \theta_W(t)\sigma - \int R(e^u - 1)\theta N(t,u)\nu(du)dt + \sigma dW^\theta(t) + \int R(e^u - 1)\tilde{N}^\theta(dt,du) \right)
\]

Notably, our way of distorting the reference measure can potentially lower the significance of either positive or negative jumps, depending on the monotonicity of the value function with respect to the state variable. In other words, the probability and the absolute value of the conditional mean size of positive (negative) jumps are the largest under the reference measure, and these jump characteristics can be smaller under the worst-case measure. Intuitively, if the value function is increasing in the state variable, then the worst-case scenario should be characterized with the smallest conditional probability and mean size of positive jumps and vice versa. To see this, \((2.7)\) indicates that \( N^\theta(dt,du) \) has a Lévy measure \( \nu^\theta(du) = (1 - \theta N(t,u))\nu(du) \) under \( Q^\theta \), and we can express it as

\[
\nu^\theta(du) = \lambda^\theta f^\theta u(du),
\]

where

\[
\lambda^\theta_t = \lambda \int R(1 - \theta N(t,u))f_u du = \lambda \left( \frac{p\eta_1}{\eta_1 - \theta N,1(t)} + \frac{q\eta_2}{\eta_2 + \theta N,2(t)} \right),
\]

and

\[
f^\theta u,t = \frac{(1 - \theta N(t,u))f_u}{\int R(1 - \theta N(t,u))f_u du}
\]

\[
= p^\theta_t(\eta_1 - \theta N,1(t))e^{(\eta_1 - \theta N,1(t))u}1_{u \geq 0} + q^\theta_t(\eta_2 + \theta N,2(t))e^{(\eta_2 + \theta N,2(t))u}1_{u < 0},
\]

\(^7\)See, for example, Chapter 1.4 of Øksendal and Sulem (2019).
with
\[ p_\theta^q = \frac{p \eta_1(\eta_2 + \theta_{N,2}(t))}{p \eta_1(\eta_2 + \theta_{N,2}(t)) + q \eta_2(\eta_1 - \theta_{N,1}(t))} \quad \text{and} \quad q_\theta^p = \frac{q \eta_2(\eta_1 - \theta_{N,1}(t))}{p \eta_1(\eta_2 + \theta_{N,2}(t)) + q \eta_2(\eta_1 - \theta_{N,1}(t))}. \]

In Table 1, we summarize the Lévy measures for \( N^\theta(dt,du) \) for \( \theta_{N,1}(t) \in [-M_1, 0] \) and \( \theta_{N,2}(t) \in [0, M_2] \).

[Table 1 goes about here]

### 2.2 The set of admissible model specifications

Since each \( \theta \in \Theta \) corresponds to a possible model that an agent considers, the natural question is what should be the domain for \( \Theta \) or the set of possible models. Intuitively, if a model \( Q^\theta \) deviates slightly from the reference model \( Q^0 \), then the agent will find it difficult to distinguish the two by observing their trajectories. On the other hand, if a model deviates substantially from its reference, its trajectories will also differ, making it easier to detect. Hence, the set of considered models should consist of statistically indistinguishable ones from the reference model.

To measure the distances between pairs of probability measures, we utilize the relative entropy advocated by Anderson et al. (2003). Given an alternative measure \( Q^\theta \) and the reference measure \( Q^0 \), we can write the growth in entropy of \( Q^\theta \) relative to \( Q^0 \) over the time interval \([t, t + \Delta t]\) as

\[
G(t, t + \Delta t) = \mathbb{E}^\theta_t \left[ \ln \left( \frac{Z^0(t + \Delta t)}{Z^0(t)} \right) \right], \quad \mathcal{R}(\theta) = \lim_{\Delta t \to 0} \frac{G(t, t + \Delta t)}{\Delta t} \quad t \geq 0. \tag{2.12}
\]

Hence, following Drechsler (2013) and Aït-Sahalia and Matthys (2019), we characterize the set of admissible model misspecification as

\[
\{ \theta(t) = (\theta_W(t), \theta_N(t)) \mid \mathcal{R}(\theta(t)) \leq h, \text{ for } t, h \geq 0 \} \tag{2.13}
\]

where \( h \) is a constant, defining an upper bound on the set of possible models. Intuitively, as \( h \) approaches zero, the set of considered models shrinks, implying that the agent gains confidence in the reference model. In contrast, a larger \( h \) indicates that the agent considers more models, especially those statistically farther away from the reference model, suggesting higher ambiguity. Hence, throughout the paper, we refer to \( h \) the total amount of ambiguity under consideration.

Next, the independence of the diffusion and jump components (per Itô-Lévy Decomposition) implies that the two contribute to relative entropy growth in an additive way, or \( \mathcal{R}(\theta(t)) = \mathcal{R}(\theta_W(t)) + \mathcal{R}(\theta_N(t)) \). Therefore, following Aït-Sahalia and Matthys (2019), we use \( h_W \) and
\(h_N = h - h_W\) to regulate the contribution of the diffusion and jump components to total relative entropy. In another word, we can use \(h_W\) or \(h_N\) to control the agent’s concern over diffusion ambiguity or jump ambiguity.

The key to the above is the determination of \(h\). Based on the log of the Radon-Nikodym derivative process \(\zeta^\theta(t) = \ln\left(Z^\theta(t)\right)\), Anderson et al. (2003) provide a statistical tool, detection-error probability, to quantify the amount of ambiguity that seems plausible to the agent. Intuitively, if the right model is \(Q^0\) and we have a history of the state process with length \(T - t\), this test statistic shows the likelihood that an agent will mistakenly assume that the data come from \(Q^\theta\) instead of \(Q^0\). Formally, this test statistic is defined as

\[
\pi(t, T; h) = \frac{1}{2} \left[ Q^0 \left\{ \zeta^\theta(T) > 0 \right\} F_{I_t} + Q^\theta \left\{ \zeta^\theta(T) < 0 \right\} F_{I_t} \right].
\]  

The first term on the right-hand side inside the bracket denotes the probability that an agent will falsely reject the right model \(Q^0\) for \(Q^\theta\) based on a history of the state process with length \(T - t\). Conversely, the second term is the probability that an agent will erroneously reject the right model \(Q^\theta\) for \(Q^0\). We provide more details of the calibration of (2.14) in Section 3.2.

### 2.3 Financing innovation with only equity

Since our focus is to examine the effects of ambiguity and skewness on innovation decisions and their interaction with the financing structure, we maintain the standard assumptions in the financing of growth options framework (e.g., Sundaresan and Wang, 2007; Sundaresan et al., 2015). We assume that the innovation project is irreversible. We assume that the entrepreneur cannot finance the project entirely with cash and must rely on external financing, solely equity or a mix of equity and debt. To generate richer implications, we introduce agency frictions based on Morellec et al. (2012), who assume that the entrepreneur can divert a fraction \(\xi \in [0, 1)\) of the free cash flow for his private benefit.

First, we consider the case where the entrepreneur finances the project entirely with equity. Let \(\delta \in [0, 1]\) denote his share percentage in the firm, once started. Then the value of his claim at the investment time \(\tau_I\) under an arbitrary measure \(Q^\theta\) for \(\theta \in \Theta\) is

\[
V_e(\tau_I, \theta) = \delta \left[ \mathbb{E}^\theta_{\tau_I} \left[ \int_{\tau_I}^{T} e^{-r(T-\tau_I)} (1-\xi)(1-\phi)X(t)dt \right] - I \right] + \mathbb{E}^\theta_{\tau_I} \left[ \int_{\tau_I}^{T} e^{-r(T-\tau_I)} \xi(1-\phi)X(t)dt \right].
\]  

Here \(\phi \in (0, 1)\) denotes the corporate tax rate, and \(0 < T \leq \infty\) (see the discussions in the next paragraph). From now on, we use \(\xi_\phi = 1 - (1-\xi)(1-\phi)\) for notational simplicity. In (2.15), the first term represents the value of his equity holdings, and the second term is the present value of his private benefits. At time 0, he decides the optimal timing \(\tau_I^*\) to start the project, under the
worst-case pricing measure $Q^{\theta^*}$. Hence his value function at time 0 is

$$V_e(0, \tau^*_I, \theta^*) = \sup_{\tau_I \in [0, T]} \inf_{Q^\theta} \mathbb{E}_0^\theta \left[ e^{-r \tau_I} V_e(\tau_I, \theta) \right], \text{ for } T \leq +\infty. \quad (2.16)$$

where $V_e(\tau_I, \theta)$ is given by (2.15).

It is necessary to discuss a few technical details of the optimal stopping problem under ambiguity, given by (2.15) and (2.16). First of all, Quenez and Sulem (2014) prove that the nonlinear expectation in (2.16) admits the minimax relation. Thus, we can first evaluate the minimum expectation problem to find the worst-case measure $Q^{\theta^*}$ and then solve the optimal stopping problem for $\tau_I^* I$ under $Q^{\theta^*}$. Since the minimum expectation problem involves the application of the Girsanov theorem and the comparison theorem for BSDEs under Lévy processes, we specify an arbitrary finite horizon $T > 0$. Since $Q^{\theta^*}$ is valid for any $T$ in our case, we can evaluate the optimal stopping problem under $Q^{\theta^*}$ for the infinite horizon by taking the limit for $T$.

Second, Quenez and Sulem (2013) show that the minimum expectation is dynamically consistent:

$$\inf_Q \mathbb{E}_0^Q \left[ e^{-r \tau_I} V_e(\tau_I, \theta) \right] = \inf_{Q'} \mathbb{E}_0^{Q'} \left[ \inf_{Q''} \mathbb{E}_0^{Q''} \left[ e^{-r \tau_I} V_e(\tau_I, \theta) \right] \right].$$

Dynamic consistency provides analytical convenience in that the worst-case density generators for $[0, \tau_I]$ and $[\tau_I, T]$ coincide the one for $[0, T]$, meaning that we can evaluate the minimum expectation period by period. Hence, it suffices to begin with the following:

$$V_e(0, \tau_I, \theta^*) = \inf_Q \mathbb{E}_0^Q \left[ e^{-r \tau_I} V_e(\tau_I, \theta) \right], \text{ for any } \tau_I. \quad (2.17)$$

Proposition 1. The density generator that gives the minimum expectation in (2.17) is $\theta^* = (\theta^*_W, \theta^*_N) = (\kappa, 1 - e^{-M_1 u} I_{u \geq 0} - I_{u < 0})$ for all $t \in [0, T]$.

In Appendix A, we provide the proof, which relies on dynamic consistency and the comparison theorem for BSDEs under Lévy process.

An immediate implication is that under $Q^{\theta^*}, X(t)$ follows

$$dX(t)/X(t^-) = \mu^{\theta^*} dt + \sigma dW^{\theta^*}(t) + \int R (e^u - 1) \tilde{N}^{\theta^*} (dt, du) \quad (2.18)$$

where

$$\mu^{\theta^*} = \mu - \kappa \sigma - \int R (e^u - 1)(1 - e^{-M_1 u} I_{u \geq 0} - I_{u < 0}) \nu(du) = \mu - \kappa \sigma - \left( \frac{\lambda p}{\eta_1 - 1} - \frac{\lambda p \eta_1}{(\eta_1 + M_1 - 1)(\eta_1 + M_1)} \right).$$

The above indicates that under the worst-case measure, the drift is adjusted downward by $\kappa \sigma + \lambda p/(\eta_1 - 1) - \lambda p \eta_1/(\eta_1 + M_1 - 1)(\eta_1 + M_1)$, where the first term comes from the diffusion
component and the last two terms stem from the jump component. The implication and interpretation of this result is similar to that of Nishimura and Ozaki (2007), except the absence of jump ambiguity in their setting.

For the jump part, the results are richer. First, the conditional mean log positive jump size is the smallest possible \(1/(q_1 + M_1)\), and the conditional mean log negative jump size is the largest possible in absolute value \(-1/q_2\). Second, the probability of positive jumps is the smallest possible \(p \eta_1/(p \eta_1 + q(q_1 + M_1))\), and the probability of negative jumps is the highest possible \(q(q_1 + M_1)/(p \eta_1 + q(q_1 + M_1))\). Third, the jump intensity is neither the highest nor the lowest but \(\lambda^* = \lambda(p \eta_1/(q_1 + M_1) + q)\). One can refer to Table 1 for the range of these parameters. The interpretations follow the fact that investors receive cash flows from the project, making the value function increase in the \(X(t)\) argument. The worst-case scenario is when positive jumps have the smallest probability and mean magnitude, and the opposite is true for negative jumps. Next, the jump intensity is intermediate because investors perceive the lowest possible arrival rate of positive jumps and the largest possible arrival rate of negative jumps.

Next, under \(\theta^*\), so long as \(Q^{\theta^*}(\tau_1^* < +\infty) = 1\), we have

\[
V_e^\infty(0, \tau_1^*, \theta^*) = \sup_{\tau \geq 0} \mathbb{E}^{\theta^*}[e^{-rt}V_e(\tau_1, \theta^*)] = \lim_{T \to \infty} \sup_{\tau_1 \in [0, T]} \mathbb{E}^{\theta^*}[e^{-rt}V_e(\tau_1, \theta^*)],
\]  

for \(V_e(\tau_1, \theta^*)\) given by (2.15). The above holds for our requirements of the parameters. Hence, we can study this infinite horizon problem to benefit from the analytical tractability offered by the double exponential jump-diffusion process for optimal stopping problems.

**Proposition 2.** Let \(\eta_1^* = \eta_1 + M_1\) and \(\eta_2^* = \eta_2\). The infinite horizon value function \(V_e^\infty(0, \tau_1^*, \theta^*)\) as in (2.19) has the following solution

\[
V_e^\infty(0, \tau_1^*, \theta^*) = A_0 X_1^\star \left[ c_{1,1} \left( \frac{X_1^\star}{\lambda_1} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_1^\star}{\lambda_1} \right)^{-\beta_2} \right] - \delta I \left[ c_{1,0} \left( \frac{X_1^\star}{\lambda_1} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_1^\star}{\lambda_1} \right)^{-\beta_2} \right] \quad (2.20)
\]

where

\[
X_1^\star = \frac{\delta I}{A_0 (\beta_1 - 1)(\beta_2 - 1)} \frac{\eta_1^* - 1}{\eta_1^*}, \quad A_0 = \frac{\delta (1 - \xi) + \xi (1 - \phi)}{r - \mu^{\theta^*}}, \quad X(0) = x. \quad (2.21)
\]

Here, \(\beta_1, \beta_2, \beta_3, \beta_4\), satisfying \(-\infty < -\beta_4 < -\eta_2^* < -\beta_3 < 0 < \beta_1 < \eta_1^* < \beta_2 < \infty\), are the four roots of the equation \(G(\beta) = r\), with

\[
G(\beta) = \frac{1}{2} \sigma^2 \beta^2 + \left[ \mu^\theta - \frac{1}{2} \sigma^2 - \lambda^* \left( \frac{p^* \eta_1^*}{\eta_1^* - 1} + \frac{q^* \eta_2^*}{\eta_2^* + 1} - 1 \right) \right] \beta + \lambda^* \left( \frac{p^* \eta_1^*}{\eta_1^* + \beta} + \frac{q^* \eta_2^*}{\eta_2^* + \beta} - 1 \right). \quad (2.22)
\]
where

$$
\mu^0 = \mu - \kappa \sigma - \lambda p/(\eta_1 - 1) + \lambda p\eta_1/(\eta_1^* - \eta_1^*), \quad \lambda^* = \lambda(p\eta_1/\eta_1^* + q),
$$

$$
p^* = \frac{p\eta_1}{p\eta_1 + q\eta_1^*}, \quad q^* = \frac{q\eta_1^*}{p\eta_1 + q\eta_1^*}.
$$

$$
c_{1,0} = \frac{\eta_1^* - \beta_1 \beta_2}{\beta_2 - \beta_1 \eta_1^*}, \quad c_{2,0} = \frac{\beta_2 - \eta_1^* \beta_1}{\beta_2 - \beta_1 \eta_1^*}, \quad c_{1,1} = \frac{\eta_1^* - \beta_1 \beta_2 - 1}{\beta_2 - \beta_1 \eta_1^* - 1}, \quad \text{and} \quad c_{2,1} = \frac{\beta_2 - \eta_1^* \beta_1 - 1}{\beta_2 - \beta_1 \eta_1^* - 1}.
$$

(2.23)

The multiplier $\beta_1 \beta_2 (\eta_1^* - 1)/((\beta_1 - 1)(\beta_2 - 1))$ in (2.21) characterizes the optimal investment threshold. An increase in $\eta_1^*$ could affect $X_I^*$ through two mechanisms, $(\eta_1^* - 1)/\eta_1^*$ and $\beta_1 \beta_2 /((\beta_1 - 1)(\beta_2 - 1))$. While the first term increases with $\eta_1^*$, the second term is less obvious.

### 2.4 Innovation under optimal financing

Under optimal financing for the project, the entrepreneur can issue debt at $\tau_I$ to benefit from its net tax benefit. We also follow Morellec et al. (2012) for the modeling of debt contract under agency concerns. The entrepreneur can issue a consol at a coupon $C$, which lowers the tax by $\phi C$. Default could happen at $\tau_D$, and we consider optimal default, standard in the literature (Sundaresan and Wang, 2007; Sundaresan et al., 2015). Default can lead to either liquidation or reorganization. In the case of liquidation (reorganization), a proportional $\alpha_L$ ($\alpha_R$) of the asset value is lost with $0 < \alpha_R < \alpha_L < 1$. Due to the higher cost of liquidation, a surplus $\alpha_L - \alpha_R$ of the cash flow can accrue to equity holders and debt holders if they settle default by reorganization.\(^8\)

Here, we use $\theta \in [0, 1]$ to denote the surplus share of equity holders.

#### 2.4.1 The optimal default policy under ambiguity

Because of $Q^0 \{\tau_1^* < +\infty\} = 1$ in the case of all equity financing and dynamic consistency, we can first solve the entrepreneur’s optimal default decision. Right after the issuance of debt at $\tau_I$, the value of his total claim under an arbitrary measure $Q^\theta$ is

$$
\Pi(\tau_I, \tau_D, \theta) = \delta E(\tau_I, \tau_D, \theta) + \Gamma(\tau_I, \tau_D, \theta), \quad \text{for} \quad \tau_I \leq \tau_D < T < \infty.
$$

(2.24)

The levered equity value $E(\tau_I, \tau_D, \theta)$ is given by

$$
E(\tau_I, \tau_D, \theta) = \mathbb{E}_{\tau_I} \left[ e^{-r(T-\tau_I)}(1-\xi_\phi)(X(t) - C) dt + \theta(\alpha_L - \alpha_R) \int_{\tau_D}^T e^{-r(t-\tau_I)}(1-\xi_\phi)X(t) dt \right].
$$

\(^8\)Fan and Sundaresan (2000) provide more detailed illustrations on this point.
The first term represents the dividend stream while the firm is solvent, and the second term is the share of the surplus accrue to all shareholders due to reorganization. The value of the managerial rent \( \Gamma(\tau_I, \tau_D, \theta) \) follows

\[
\Gamma(\tau_I, \tau_D, \theta) = \mathbb{E}^\theta_{\tau_I} \left[ \int_{\tau_I}^{T_D} e^{-r(t-\tau_I)} \xi(1-\phi)(X(t) - C)dt + \theta(\alpha_L - \alpha_R) \int_{\tau_D}^{T} e^{-r(t-\tau_I)} \xi(1-\phi)X(t)dt \right].
\]

The first term represents the value of the diverted free cash flows upon default, and the second term is the diverted surplus accrue to all shareholders due to reorganization. Since both \( E(\tau_I, \tau_D, \theta) \) and \( \Gamma(\tau_I, \tau_D, \theta) \) represent the cash flows that the entrepreneur collects after the issuance of debt, he decides the optimal default policy \( \tau_D^* \) to maximize the present value of these cash flows under the worst-case scenario:

\[
\Pi(\tau_I, \tau_D^*, \theta^*) = \sup_{\tau_D \in [\tau_I, T]} \inf_{Q^\theta} \left\{ \delta E(\tau_I, \tau_D, \theta) + \Gamma(\tau_I, \tau_D, \theta) \right\} \text{ for } T \leq +\infty, \quad (2.25)
\]

Using the same approach as that in Section 2.3, we begin with finding the minimum expectation density generator:

\[
\Pi(\tau_I, \tau_D, \theta^*) = \inf_{Q^\theta} \left\{ \delta E(\tau_I, \tau_D, \theta) + \Gamma(\tau_I, \tau_D, \theta) \right\}. \quad (2.26)
\]

**Proposition 3.** The density generator that gives the minimum expectation in (2.26) is \( \theta^* = (\theta^*_W, \theta^*_N) = (\kappa, 1 - e^{-M_i} u 1_{u \geq 0} - 1_{u < 0}) \) for all \( t \in [\tau_I, T] \).

The worst-case scenario for levered equity holders is the same for unlevered equity holders. This result follows that in both financing scenarios, equity holders collect net profit, whereas they collect net profit until bankruptcy (forever) in the former (latter) scenario. Consequently, in both scenarios, they are concerned with the prior that the EBIT profile is the least favorable.

Given that we have determined the worst-case prior for equity holders, we can find the optimal default timing \( \tau_D^* \). Same as in the investment case, under \( \theta^* \), so long as \( Q^\theta(\tau_D^* < +\infty) = 1 \), we have

\[
\Pi^\infty(\tau_I, \tau_D^*, \theta^*) = \sup_{\tau_D \in [\tau_I, \infty]} \Pi(\tau_I, \tau_D, \theta^*) = \lim_{T \to \infty} \sup_{\tau_D \in [\tau_I, T]} \Pi(\tau_I, \tau_D, \theta^*). \quad (2.27)
\]

The above holds for our requirements of the parameters. Hence, we can study this infinite horizon problem to benefit from the analytical tractability offered by the double exponential jump-diffusion process.

**Proposition 4.** Let \( \beta_3, \beta_4, \mu^{\theta^*}, \eta_1^*, \text{ and } \eta_2^* \) be the same as in Proposition 2. The infinite horizon
value function $\Pi^\infty(\tau_I, \tau_D^*, \theta^*)$ as in (2.27) has the following expression

$$
\Pi^\infty(\tau_I, \tau_D^*, \theta^*) = B_0 X(\tau_I) - B_1 C + B_1 C \left[ d_{1,0} \left( \frac{X_D}{X(\tau_I)} \right)^{\beta_3} + d_{2,0} \left( \frac{X_D}{X(\tau_I)} \right)^{\beta_4} \right]
- (B_0 - B_2) X_D \left[ d_{1,1} \left( \frac{X_D}{X(\tau_I)} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D}{X(\tau_I)} \right)^{\beta_4} \right]
$$

(2.28)

where

$$
B_0 = \frac{\delta(1 - \xi_\phi) + \xi(1 - \phi)}{r - \mu^{\theta^*}}, \quad B_1 = \frac{\delta(1 - \xi_\phi) + \xi(1 - \phi)}{r}, \quad B_2 = \frac{\theta(\alpha_L - \alpha_R)(\delta(1 - \xi_\phi) + \xi(1 - \phi))}{r - \mu^{\theta^*}},
$$

$$
d_{1,0} = \frac{\eta_2^* - \beta_3 \beta_4}{\beta_4 - \beta_3 \eta_2^*}, \quad d_{2,0} = \frac{\beta_4 - \eta_2^* \beta_3}{\beta_4 - \beta_3 \eta_2^*}, \quad d_{1,1} = \frac{\eta_2^* - \beta_3 \beta_4 + 1}{\beta_4 - \beta_3 \eta_2^* + 1}, \quad d_{2,1} = \frac{\beta_4 - \eta_2^* \beta_3 + 1}{\beta_4 - \beta_3 \eta_2^* + 1},
$$

(2.29)

and the optimal default policy is $\tau_D^* = \inf_t \{ X(t) \leq X_D^* \}$ with

$$
X_D^* = \frac{(r - \mu^{\theta^*})}{r(1 - \theta(\alpha_L - \alpha_R))(\beta_3 + 1)(\beta_4 + 1)\eta_2^*} C.
$$

(2.30)

### 2.4.2 The optimal financing choice

To determine the optimal financing choice, we need to find the debt value or the optimal coupon that maximizes the value of all the entrepreneur’s claims at the time of debt issuance. We assume that debt investors have the same preference and set of priors as equity investors. Furthermore, debt holders and equity holders know each other’s decision rules. Hence, when equity holders choose any coupon $C$ at $\tau_I$, debt holders’ valuation of debt is

$$
D^\infty(\tau_I, \tau_D^*, \theta^*) = \inf_Q \mathbb{E}^Q \left[ \int_{\tau_I}^{\tau_D^*} e^{-r(t-\tau_I)} C dt + (1 - \alpha_R - \theta(\alpha_L - \alpha_R)) \int_{\tau_D^*}^{\tau_I} e^{-r(t-\tau_I)} (1 - \xi_\phi) X(t) dt \right].
$$

(2.31)

Note that the value function is in the limiting sense in the above for notation simplicity because the limit is well-defined intuitively. From equity holders’ perspective, default happens almost surely. Hence it is natural that debt holders think the same.

**Proposition 5.** The density generator that gives the minimum expectation in (2.31) is $\theta^* = (\theta_W^*, \theta_N^*) = (\kappa, 1 - e^{-M_1 u} I_{u \geq 0} - I_{u < 0})$ for all $t \geq \tau_I$. Furthermore, the value of debt has the expression

$$
D^\infty(\tau_I, \tau_D^*, \theta^*) = \frac{C}{r} \left[ 1 - d_{1,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} - d_{2,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right] + \frac{a_0 X_D^*}{r - \mu^{\theta^*}} \left[ d_{1,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right],
$$

(2.32)

where $d_{i,j}$ is given by (2.29), $X_D^*$ is given by (2.31), and $a_0 = (1 - \alpha_R - \theta(\alpha_L - \alpha_R))(1 - \xi_\phi) > 0$.
The above result indicates that debt investors have the same pricing measure as the entrepreneur and external equity investors. This result is natural since debt investors become equity holders of the unlevered firm after reorganization. Hence, their worst-case scenario coincides with that of the entrepreneur and equity investors.

Now we can find the optimal capital structure by choosing the optimal coupon $C^*$ to maximize the value of the entrepreneur’s total claim at the investment time $\tau_I$

$$V_\infty^*(\tau_I) = \max_C \{\delta(E_\infty^*(\tau_I) + D_\infty^*(\tau_I) - I) + \Gamma_\infty^*(\tau_I)\} = \max_C \{\Pi_\infty^*(\tau_I) + \delta D_\infty^*(\tau_I) - \delta I\}. \quad (2.33)$$

In the absence of agency cost $\xi = 0$, the above problem is standard in that the entrepreneur maximizes total equity value at the time of debt issuance. When $\xi > 0$, the optimal coupon hence leverage deviate from the standard results. Unlike in the diffusion case (e.g., Sundaresan and Wang, 2007; Sundaresan et al., 2015; Morellec et al., 2012), the closed form solution for $C^*$ is unavailable for (2.33). Nevertheless, the following Proposition shows that $C^*$ is still a linear function of $X(\tau_I)$ like in the diffusion case.

**Proposition 6.** The $C^*$ that maximizes $V_\infty^*(\tau_I)$ in (2.33) has the form $C^* = \psi X(\tau_I)$ where $\psi$ is the unique positive solution to the following equation

$$B_3 = B_5 \psi^{\beta_3} + B_6 \psi^{\beta_4}, \quad (2.34)$$

where

$$B_3 = \frac{\delta \xi \phi - \xi (1 - \phi)}{r}, \quad B_4 = \frac{\delta \alpha_R (1 - \xi \phi) + \xi (1 - \phi) (1 - \theta (\alpha_L - \alpha_R))}{r - \mu^{\theta^*}};$$

$$B_5 = B_3 d_{1,0} (1 + \beta_3) \gamma^{\beta_3} + B_4 d_{1,1} (1 + \beta_3) \gamma^{1 + \beta_3}, \quad B_6 = B_3 d_{2,0} (1 + \beta_4) \gamma^{\beta_4} + B_4 d_{2,1} (1 + \beta_4) \gamma^{1 + \beta_4},$$

$\beta_i, d_{i,j}$ are given by (2.29), and $\gamma$ satisfies $X_D^* = \gamma C$ as in (2.30).

2.4.3 **The value of the growth option**

Using the above, we can write the total value of the entrepreneur’s claim at $\tau_I$ as

$$V_\infty^*(\tau_I, \tau_D^*, \theta^*) = B_7 X(\tau_I) - \delta I, \quad (2.35)$$

where

$$B_7 = B_0 + B_3 \psi \left[1 - d_{1,0} (\gamma \psi)^{\beta_3} - d_{2,0} (\gamma \psi)^{\beta_4}\right] - B_4 \gamma \psi \left[d_{1,1} (\gamma \psi)^{\beta_3} + d_{2,1} (\gamma \psi)^{\beta_4}\right]$$
for $B_0$ given in Proposition 4 and $B_3, B_4$ given in Proposition 6. Therefore, the value of the growth option under optimal financing at time 0 is

$$V^\infty_*(0, \tau^*_I, \theta^*) = \sup_{\tau^*_I} \inf_{\theta \in \Theta} e^{-r \tau^*_I} V^\infty_*(\tau^*_I, \tau^*_D, \theta^*).$$  \hfill (2.36)

Again, note that the above is in the limiting sense, since $\tau^*_I$ is finite almost surely.

**Proposition 7.** The density generator that gives the minimum expectation in (2.36) is

$$\theta^* = (\theta^*_W, \theta^*_N) = (\kappa, 1 - e^{-M_1 u} I_{u \geq 0} - I_{u < 0}) \text{ for all } t \in [0, \tau^*_I].$$

Furthermore, the value of the growth option has the expression

$$V^\infty_*(0, \tau^*_I, \theta^*) = B_7 X_I [c_{1,1} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_I}{x} \right)^{-\beta_2}] - \delta I [c_{1,0} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_I}{x} \right)^{-\beta_2}],$$  \hfill (2.37)

and the investment boundary satisfies

$$X^*_I = \frac{\delta I}{B_7 (\beta_1 - 1) (\beta_2 - 1)} \frac{\eta^*_1 - 1}{\eta^*_1},$$

In the above, $c_{i,j}, \beta_i$, and $\eta^*_i$ are the same as in Proposition 2, and $B_7$ is from (2.35).

### 3 Model implications

#### 3.1 Parameter choices

For contractual parameters, we largely follow the choice of Morellec et al. (2012). For example, we choose the liquidation cost $\alpha_L = 0.5$ and the renegotiation cost $\alpha_R = 0.02$.\footnote{The liquidation cost is higher than the typical value (0.25) used in the literature (e.g., Leland, 1998; Goldstein et al., 2001). This higher cost better suits the study of innovation, characterized by low tangibility and recovery rate in default.} We set shareholder bargaining power to $\theta = 0.5$ and private benefits to $\xi = 0.01$. However, we set the entrepreneur’s ownership to 51% instead of 7.41% in their setting because we focus on growth options instead of assets in place. It is normal to assume that an innovative entrepreneur wants to be the majority shareholder.

Motivated by the literature on innovation (e.g., Manso, 2011), we examine two types of innovation, exploration and exploitation, and their interactions with the financing structure. The characteristics of these two types inspire us to specify two sets of benchmark parameters for the EBIT process, (1) a small number of large jumps and (2) a moderate number of small jumps. We follow the first study (Chen and Kou, 2009) that model EBIT as a double exponential jump-diffusion process for the choice of the benchmark parameters, listed in Table 2. For case (1), the
benchmark jump arrival rate parameter $\lambda = 1/5$, with equal mean size for positive and negative jumps, $\eta_1 = \eta_2 = 3$. This benchmark choice implies that, on average, every five years, exploration can generate an outcome, which can be either a huge success or failure. For case (2), the benchmark jump arrival rate parameter $\lambda = 1$, with equal mean size for positive and negative jumps, $\eta_1 = \eta_2 = 8$. This choice implies that, on average, every year, exploitation can generate an outcome, which can be either a moderate success or failure. Furthermore, studies have documented that the risk-neutral skewness of individual stocks can be both positive and negative (e.g., Conrad et al., 2013), unlike that of the S&P 500 index. Therefore, we are further motivated to assess whether the symmetry of the log EBIT process under the reference measure could have any impact. This is why we set the mean jump sizes for both positive and negative jumps to be the same, unlike Chen and Kou (2009), and we control skewness with the conditional probability for positive jumps. Below, using the moment generating function of the log EBIT process $Y(t) = \ln(X(t))$, we list its first four central moments and report their values under different sets of parameters in Table 2.

Mean ($CM_1$): $E[Y(t)] = \mu - \frac{1}{2} \sigma^2 + \lambda \left(1 + p \frac{\eta_1}{\eta_1} - q \frac{\eta_2}{\eta_2} - \frac{p \eta_1}{\eta_1 - 1} - \frac{q \eta_2}{\eta_2 + 1}\right)$. (3.1)

Variance ($CM_2$): $E[(Y(t) - E[Y(t)])^2] = \sigma^2 + 2\lambda \left(p \frac{\eta_1}{\eta_1^2} + q \frac{\eta_2}{\eta_2^2}\right)$. (3.2)

Skewness ($CM_3$): $E[(Y(t) - E[Y(t)])^3]/CM_2^{3/2} = 6\lambda \left(p \frac{\eta_3}{\eta_1^3} - q \frac{\eta_3}{\eta_2^3}\right)/CM_2^{3/2}$. (3.3)

Kurtosis ($CM_4$): $E[(Y(t) - E[Y(t)])^4]/CM_2^2 = (3\sigma^4 + \frac{12\lambda^2 p^2 + 24\lambda p}{\eta_1^4} + \frac{12\lambda^2 q^2 + 24\lambda q}{\eta_2^4} + \frac{24\lambda^2 pq}{\eta_1^2 \eta_2^2} + \frac{12\lambda p \sigma^2}{\eta_1^2} + \frac{12\lambda q \sigma^2}{\eta_2^2})/CM_2^2$. (3.4)

Panel B of Table 2 lists the first four central moments of $Y(t)$ under our benchmark parameters. First, we notice that the kurtosis values of $Y(t)$ are much higher under Exploration, around 11.113, over that under Exploitation, around 4.154. This contrast is consistent with the idea that the payoff from exploration is characterized by more extreme values than that from exploitation. Second, within each case, by changing the probability of positive jumps $p$, we have symmetric, negatively skewed, and positively skewed increment distributions. Third, Exploration is associated with lower mean and higher variance values, yet the variance values are less sensitive to the change in $p$. 

[Table 2 goes about here]
3.2 Calibration of detection-error probability

As illustrated in Section 2.2, we must quantify the maximal amount of total ambiguity \( h = h_W + h_N \) under consideration to completely characterize the EBIT dynamics under the worst-case measure \( Q^0 \). For the double exponential jump-diffusion process, we can write the relative entropy growth \( \mathcal{R}(Z^{0*}(t)) \) as

\[
\mathcal{R}(Z^{0*}(t)) = \frac{1}{2} \kappa^2 + \left( \frac{1}{\eta_1} - \frac{2}{\eta_1^*} + \frac{1}{\eta_1^* + M_1} \right) \lambda^* p \eta_1 - \left( \frac{M_1}{(\eta_1^*)^2} - \frac{1}{\eta_1^*} + \frac{1}{\eta_1^* + M_1} \right) \lambda^* p^* \eta_1^*,
\]

(3.5)

where \( \eta_1^*, p^*, \) and \( \lambda^* \) are the same as in Proposition 2, and the derivations are in the Appendix B.2. It is clear that the relative entropy growth constraint will bind at the optimum, meaning that \( \mathcal{R}(Z^{0*}(t)) = h \). Moreover, given that the diffusion and jump components contribute to relative entropy growth in an additive way, we must have

\[
\frac{1}{2} \kappa^2 = h_W,
\]

and

\[
\left( \frac{1}{\eta_1} - \frac{2}{\eta_1^*} + \frac{1}{\eta_1^* + M_1} \right) \lambda^* p \eta_1 - \left( \frac{M_1}{(\eta_1^*)^2} - \frac{1}{\eta_1^*} + \frac{1}{\eta_1^* + M_1} \right) \lambda^* p^* \eta_1^* = h_N.
\]

In other words, we can recover the optimal values of \( \kappa \) and \( M_1 \) based on \( h \) and \( h_N \).\(^{10}\)

To find the optimal value of \( h \), we follow the recommendation of Anderson et al. (2003) and set \( h \) such that the detection-error probability is 10%.\(^{11}\) Specifically, for the detection-error probability expression given by (2.14), Aït-Sahalia and Matthys (2019) provide a way to calculate it based on the conditional Fourier transform of \( \zeta^{0*}(t) = \ln(Z^{0*}(t)) \). Given a sample of length \( n > 0 \), the detection-error probability is

\[
\pi(t, n; h) = \frac{1}{2} \left[ Q^0 \{ \zeta^{0*}(n) > 0 \} \mathcal{F}_t \right] + Q^0 \{ \zeta^{0*}(n) < 0 \} \mathcal{F}_t \}, \quad t \geq 0, \ n = mT
\]

(3.6)

\[
= \frac{1}{2} - \frac{1}{2\pi} \int_{\mathbb{R}^+} \Re \left[ \tilde{\zeta}^{0*}(u, t, n)/iu \right] - \Re \left[ \tilde{\zeta}_0^{0*}(u, t, n)/iu \right] \right] du.
\]

In the first line, \( T \) is the number of years, and \( m \) is the sampling frequency, provided that we use annualized parameter values. In the second line, \( i = \sqrt{-1} \), \( \Re(\cdot) \) denotes the real part of a complex number, and the two conditional Fourier transforms are

\[
\tilde{\zeta}^{0*}(u, t, n) = \mathbb{E}_t^Q \left[ e^{iu\zeta^{0*}(n)} \right], \quad \text{and} \quad \tilde{\zeta}_0^{0*}(u, t, n) = \mathbb{E}_t^{Q^0} \left[ e^{iu\zeta^{0*}(n)} \right].
\]

\(^{10}\)Aït-Sahalia and Matthys (2019) provide a formal treatment for the recovery of the ambiguity parameters based on the relative entropy growth constraint.

\(^{11}\)Notably, if this probability is 0.5 (0), then the two models become statistically indistinguishable (perfectly distinguishable).
Implementing the calibration process further requires us to specify \( h_w/h \) and \( n = mT \). For the baseline scenario, we consider \( h_w/h = 0.5 \), where the agent places equal concern for diffusion ambiguity and jump ambiguity. For the sample length, we assume that agents have access to Compustat to study historical EBIT data. Although there are sixty years of data available in Compustat, we choose thirty years \( T = 30 \) for our baseline since we consider growth options rather than assets in place, and remote observations might be less informative. Lastly, we choose the quarterly frequency \( m = 4 \), the highest frequency allowed by Compustat.

Figure 1 presents the calibration results for the detection error probabilities for both innovation scenarios. In all panels, black solid lines with triangles represent the results for exploration, and blue dashed lines with circles are for exploitation. In Panel (a), the detection error probability \( \pi(n; h) \) decreases as \( h \) increases, meaning that the probability of rejecting a wrong prior decreases when the set of prior enlarges. The critical value \( h^* \) satisfying \( \pi(n; h^*) = 0.1 \) is 0.0299 (0.0315) for exploration (exploitation). In addition, \( \pi(n; h) \) for exploitation is above that for exploration for all \( h \leq h^* \), meaning that agents consider a wider range of possible models for exploitation than for exploration. We can interpret this observation by the distinct nature of jumps in the processes for these two innovations. Jumps in the process for exploration are of lower frequency and larger magnitudes than those for exploitation. Hence it is less challenging to separate jumps from diffusion in exploration with sufficient historical data.

In Panels (b) to (g), we plot the associated parameters under the worst-case measure to provide additional insights for the calibration process. First, Panels (b) and (c) plot the ambiguity parameters \( \kappa^* \) and \( M_1^* \). The diffusion ambiguity parameter \( \kappa^* \) is the same in both exploration and exploitation since the relative entropy growth from the diffusion part is the same for the two cases. Next, the jump ambiguity parameter \( M_1^* \) is smaller for exploration for all \( h \in [0, h^*] \), meaning that agents are more confident with their identifications of jumps in exploration. In contrast, since jumps in exploitation are more frequent and of smaller magnitudes, agents find it more challenging to separate them from diffusion innovations. Second, Panels (d) to (g) show the parameters under the worst case measure. Notably, as total ambiguity \( h \) increases, the conditional mean log positive jump size \( 1/\eta_1^* \), the jump intensity \( \lambda^* \), the probability of positive jumps \( p^* \), and the ambiguity adjusted drift \( \mu^* \) all decrease. Notably, \( h \) lowers \( p^* \) and \( \mu^* \) to larger extents for exploration than exploitation, and the implication of this result will become clear in our analysis of the first four central moments of \( Y(t) \).

Figure 2 shows the first four central moments of \( Y(t) \) with respect to total ambiguity \( h \) for the two innovations. It is evident that exploration is characterized with a lower mean, a higher variance, a lower skewness, and a higher kurtosis than exploitation, and the discrepancy widens.
as $h$ increases. Ambiguity lowers both the mean and the variance of $Y(t)$, especially for exploration. Notably, although $Y(t)$ is symmetric for the two innovations under the reference measure, ambiguity makes $Y(t)$ negatively skewed, whereas the effect is stronger for exploration. For example, the skewness under exploration approaches -1.1 from 0 as $h$ approaches $h^*$ from 0, while the skewness under exploitation only approaches -0.2 from 0. Lastly, the kurtosis of $Y(t)$ for exploration is more than 11 under the reference measure, whereas the kurtosis for exploitation is around 4. This observation is not surprising since exploration is characterized by more extreme outcomes. The results on skewness and kurtosis indicate that ambiguity makes agents more concerned with extremely negative outcomes, especially for explorations.

To provide more insights on the effects of jump ambiguity, we plot the conditional jump size distributions of $Y(t)$ under the worst-case measures for each innovation type in Figure 3. Overall, the results indicate that ambiguity distorts the baseline symmetric distribution towards the negative domain. Under $Q^{\theta^*}$, the jump size distributions are negatively skewed with heavier left tails and lighter right tails, relative to the one under $Q^0$. This result suggests that the probability and size of negative jumps are larger under the worst-case measure than under the reference one. Comparing the results in Panels (a) and (b), we note that the distorting effects are stronger for Exploration than Exploitation. This observation is consistent with the fact that the jump size distribution under the reference measure for exploration has a higher dispersion and fatter tail. This comparison suggests that for explorations, ambiguity averse agents view negative jumps to be more likely and more severe.

3.3 The effects of ambiguity: baseline results

In Figure 4, we show the effects of ambiguity on investment value under the baseline parameters discussed above. In Panel (a), we plot the project values $V_i(0)$ under different innovation types and financing structures against total ambiguity $h$, where the subscript $e$ ($*$) denotes all equity (optimal) financing. In Panel (b) and (c), we plot the optimal investment boundary $X_{i}^{*}$ and the AD price for investment $E_{0}[e^{-rt_i}]$ with the superscript $i$ indicating equity or optimal financing. In Panels (d) to (e), we further plot the relative gain in the project value due to debt financing $V_*(0)/V_e(0) - 1 \ (%)$, and the relative differences in the optimal investment boundary and the AD prices.
Overall, ambiguity increases the investment boundaries, lowers the AD prices for investment and investment values, and the effects are stronger for exploration. For example, under the reference measure, the value of exploration is higher than exploitation's, even though the two have the same value if invested immediately, which is \((1 - \xi_{\phi})X(0)/(r - \mu) - I\). The difference in investment values is a product of the different investment boundaries and AD prices, where the former is higher for exploration and the latter is higher for exploitation. As ambiguity increases, the value of exploration drops faster than exploitation's. This result follows the responses of the optimal investment boundaries and the AD prices, where both are more sensitive to ambiguity for exploration. These results are not surprising since we have already shown in Figures 1 and 2 that ambiguity lowers the log EBIT skewness to larger extents for exploration. Since the growth option is call-like, its value decreases sharply with skewness. Hence, given an exploration project and an exploitation project with identical NPV when initiated now, agents facing high ambiguity are more likely to choose the latter. The reason is that the optimal investment boundary is lower while the associated AD price and project value are higher for the latter.

Next, it is important to see that the investment value under optimal financing always dominates that under equity financing for all ambiguity levels and both types of innovation. This difference is driven by the net tax benefit of debt, which makes total firm value exceed the value of its assets. This point is consistent with the fact that the optimal investment boundary under optimal financing is consistently below that under equity financing. In contrast, the AD price of investment is higher under optimal financing. Thanks to the net tax benefit, for any EBIT level, the investment value under optimal financing is higher than that under equity financing, hence for any required NPV threshold, the investment value under optimal financing can always meet the threshold earlier. Hence, these results suggest that access to debt financing can promote innovation in general by accelerating the project's initiation and increasing its value.

Notably, the results in Panels (d) to (f) indicate that the positive effect of debt financing on innovation is more prominent for exploration than exploitation. Specifically, Panel (d) shows that the relative gain from debt financing, \(V_*(0)/V_e(0) - 1\), increases in ambiguity for both exploration and exploitation, whereas the gain for the former increases more with ambiguity. For example, under the reference measure, the relative gain for exploration is around 11%, which is around 12% for exploitation. As ambiguity \(h\) approaches the critical level \(h^*\), the gain for exploration is around 24.5%, while that for exploitation is around 22%. To interpret this discrepancy, we can look at the relative differences for the investment boundary (Panel e) and the AD price (Panel f). As ambiguity increases, the absolute value of the relative difference for the
investment boundary decreases. Nevertheless, the absolute value of the difference for exploitation decreases at a faster rate than exploration. Next, the relative difference of the AD price closely resembles that for the investment value. For example, under the reference measure, the relative difference for exploration is around 10.8%, which is around 11.9% for exploitation. As ambiguity approaches the critical level $h^*$, the gain for exploration is around 24.4%, while that for exploitation is around 21.9%. Hence, we conclude that debt financing adds more value to exploration when ambiguity increases because the investment acceleration effect from debt is stronger for exploration when ambiguity is higher.

Lastly, it is necessary to examine the optimal capital structures for these two types of innovation. However, the analysis of the financing part is plagued by the fact that the optimal coupon $C^*$ is a function of $X(\tau^*_I)$ instead of $X^*_I$ (Proposition 6), whereas the two might not be the same due to the overshoot.\(^{12}\) Fortunately, leverage, $D/(D+E)$, and the AD price of default, $\mathbb{E}_{\tau_I^*}[e^{-r(\tau^*_D-\tau^*_I)}]$, do not depend on the value of $X(\tau^*_I)$, whereas the default boundary, $X^*_D = \gamma C^* = \gamma \psi X(\tau^*_I)$, can be presented as a ratio $X^*_D/X^*_I = \gamma \psi$. For other quantities like the optimal coupon $C^*$, the value of debt $D(\tau^*_I)$, and the value of levered equity $E(\tau^*_I)$, we compute the quantities as if $X(\tau^*_I) = X^*_I$ just for illustrative purposes.

Figure 5 shows the optimal capital structure decisions under the two types of innovation. At a first glance, the optimal leverage for exploration is always higher than that for exploitation. Considering the result that the optimal investment boundary is higher under exploration, we can interpret that this higher leverage leads to higher net tax benefits, which contribute to the accelerating effect of debt.\(^{13}\) To interpret this result, note that under the reference measure, debt value is higher for exploration (Panel e), while the same is true for (levered) equity value (Panel f). As total ambiguity increases, debt and equity values for exploration decrease at faster rates than those for exploitation, nevertheless, the rates of change are relatively consistent, which is also evident in the response of the optimal coupon (Panel d). Next, to better see the relation between levered equity and ambiguity, we can look at the (scaled) default boundary (Panel b) and the AD price of default (Panel c). Since the value of levered equity decreases with the AD price of default, which increases with ambiguity, then the value of levered equity decreases with ambiguity. Although lowered default boundary by itself could increase the value

\(^{12}\)This is not a problem under diffusion settings, because the trajectory of the state process is continuous.

\(^{13}\)Regarding the actual leverage level, it is not necessary to compare with the standard case (e.g., Leland, 1998; Goldstein et al., 2001) in that the modeling of default is different. In the Morellec et al. (2012) approach that we adopt, the optimal leverage should naturally be higher compared to the standard case because the bankruptcy cost is lower due to reorganization than liquidation. Nevertheless, our key results do not depend on the specific way to model default. In the earlier version of this draft, we assume liquidation and the results are qualitatively the same.
of levered equity, the lowered boundary does not offset the increased AD price at the same time.

3.4 Comparative statics

In this subsection, we present further comparative statics to the key characteristics of innovation. First, we examine whether the skewness of the EBIT process under the reference measure can affect innovation outcomes. Second, we study whether the novelty of the innovation project itself plays a role. Specifically, we analyze the effect of the size of agents’ information set available for calibrating the set of multiple priors. Third, we assess the implications of the relative concern for diffusion ambiguity or jump ambiguity. Lastly, we further substantiate the result that the fostering effect of debt on innovation stems from the net tax benefit channel by examining agency concerns.

3.4.1 The effect of ex ante skewness

Figure 6 shows the effects of ambiguity on the value of the innovation project conditional on the skewness of $Y(t)$ under the reference measure. In Panels (a) to (c), we plot the relative gain in the project value due to debt financing, and the project values under equity financing and debt financing against total ambiguity $h$ for exploration. In Panels (d) to (f), we plot the same for exploitation. Our computation only alters the probability of positive jumps under the reference measure and keeps the other parameters at their benchmark values. (see Table 2). We also set $h_W / h = 0.5$ and $n = 120$. To further interpret the results, we report all the relevant quantities at the critical total ambiguity level $h^*$ in Table 3.

[Figure 6 goes about here]

[Table 3 goes about here]

Overall, the results suggest that debt financing is most value-enhancing for exploratory projects with negative skewness ex ante. Under the reference measure and for both innovation projects, the project value is the highest (lowest) when $Y(t)$ is positively (negatively) skewed, the left end of the red dashed (black solid) line with the marker star (dot) in Panels (b), (c), (e), and (f). Since the growth option is call-like, its value is naturally the highest under positive skewness when holding everything else the same. Next, as ambiguity increases, the project values fall in all cases. Nevertheless, the project value falls the fastest under positive ex ante skewness and the slowest under negative skewness, reversing the orders of the magnitudes. Lastly, Panels (a) and (d) show that while the project value gain from leverage is the highest for exploratory projects with negative skewness, it is less so for exploitative projects. In the latter case, the gain
from leverage is the highest under negative skewness only for low ambiguity levels, suggesting nonlinear effects.

To further interpret the above results, we turn to Table 3 and see all the relevant quantities at the critical level of ambiguity $h^*$. Under negative ex ante skewness, the positive jump distortion parameter $M_{1}^*$ is the highest, meaning that investors view positive jumps as less significant when facing higher ambiguity. Recall that under the reference measure, the drift parameter $\mu = 0.02$ (Table 2), so it is interesting to see that the drift parameter decreases to the largest extent under positive skewness. This follows the expression for $\mu^*$ given by (2.18), which decreases in $p$. Moreover, the variance and kurtosis under positive skewness also decrease to the largest extent, and these relations can be seen from the related expressions, (3.2) and (3.4). Hence, these relations could shed light on the smallest project value under positive skewness at the maximum ambiguity level. Indeed, the comparison of the optimal investment boundary and the associated AD price of investment reveals consistent results. Similar to the benchmark case, the relative gain due to leverage is largely explained by the relative difference in the AD prices.

3.4.2 The effect of sample length

Table 4 shows the effects of ambiguity on the value of the innovation project conditional on the length of the sample available to investors for calibrating detection error probability. Hence, we can interpret the results here as the effects of innovation novelty in that a novel project should have less information available as the reference. We only alter the number of observations $n$ in our computation and keep the other parameters at their benchmark values (see Table 2). Specifically, we consider $n = 60, 120, 240$, corresponding to 15, 30, and 60 years of quarterly data.

The results show that the value of the innovation project increases when investors have more data available to narrow down the set of equivalent priors. For each innovation type, the relative entropy growth bound $h^*$ is the highest for $n = 60$, signifying the highest level of total ambiguity. It is consistent with the idea that more information can make agents more confident about their model. In this case, the value of the project is the lowest due to the highest investment boundary and the lowest AD price. We turn to the central moments for $Y(t)$ under the worst-case measures to interpret this result. For exploration, the four central moments increase with $n$ except for kurtosis, while the central moments all increase with $n$ for exploitation. These results suggest that higher mean and variance can increase real option values like in the diffusion case, while higher moments’ effects can be nonlinear. When skewness is moderately
negative, higher kurtosis could still increase real option value since the sizes of extreme positive jumps are still non-negligible. However, when skewness is sufficiently negative, higher kurtosis will decrease real option value since negative jumps dominate. Same as before, the relative gain from leverage is higher for exploration than for exploitation. For $n = 60$, the relative gain is 31% (27%) for exploration (exploitation), and it decreases to 18.9% (18.5%) for $n = 240$. This result is also explained mainly by the difference in the AD price of investment. These results collectively highlight that the availability of relevant information is key to reducing ambiguity and leading to higher valuation of innovation projects. Moreover, debt financing is more value improving for relatively novel projects, characterized by limited available information.

3.4.3 The effect of diffusion ambiguity share

Table 5 shows the effects of the relative concern of diffusion (jump) ambiguity on the value of the innovation project. In our computation, we only alter the share of diffusion ambiguity $h_\text{W}/h$ and keep the other parameters at their benchmark values (see Table 2). Specifically, we consider $h_\text{W}/h = 0.1, 0.5, 0.9$, meaning that agents consider diffusion ambiguity accounts for total ambiguity by 10%, 50%, and 90%.

[Table 5 goes about here]

Overall, the results suggest that both diffusion and jump ambiguities play important roles in determining the innovation project values and the gains from leverage, while the sensitivity to a particular type of ambiguity depends on the nature of the project. For exploration, the project value is the smallest (largest) under the case $h_\text{W}/h = 0.5 (h_\text{W}/h = 0.1)$, whereas the gain from leverage is the smallest (largest) under the case $h_\text{W}/h = 0.9 (h_\text{W}/h = 0.5)$. This result suggests that the value enhancing effect of debt is largely driven by the concern over jump ambiguity. For exploitation, the project value is the smallest (largest) under the case $h_\text{W}/h = 0.9 (h_\text{W}/h = 0.5)$, whereas the gain from leverage is the smallest (largest) under the case $h_\text{W}/h = 0.1 (h_\text{W}/h = 0.5)$. Hence, it implies that the effect of debt here comes mainly from the concern over diffusion ambiguity. These contrasts are consistent with the EBIT characteristics associated with each innovation type. Since we characterize exploration (exploitation) with infrequent (frequent) big (small) jumps, jump ambiguity naturally plays a more (less) prominent role in this case. In both cases, EBIT variation comes primarily from the diffusion component, hence it is not surprising to see that the effect of diffusion ambiguity is significant.

To provide more insights, we further present in Table 6 the results under alternative diffusive parameter ($\sigma$) values. In Panel A, we alter $\sigma$ to be 0.1 while keeping others the same as in Table 5; in Panel B, we choose $\sigma$ to be 0.3. Consequently, jumps are more important in the first case.
and less in the second.

The results in Table 6 highlight that the relative effects of diffusion versus jump ambiguity depend on the relative magnitudes of diffusion versus jumps under the reference measure. In Panel A, jump ambiguity exerts stronger negative effects on project values but leads to higher leverage gains, relative to Table 5. In Panel B, we observe the exact opposite. The results on other quantities are also worth discussing. It is not surprising to see that the maximum ambiguity parameters do not depend on \( \sigma \) because \( \sigma \) does not enter the entropy growth expression. However, the distribution of \( Y(t) \) becomes more extreme-valued under \( \sigma = 0.1 \), shown by the skewness and kurtosis values, while the opposite is true for \( \sigma = 0.3 \).

[Table 6 goes about here]

Importantly, the results in Tables 5 and 6 generate profound implications to the irreversible investment literature. In the standard case under pure diffusion (e.g., McDonald and Siegel, 1986), the investment option value increases in the variance of the log EBIT process because return variance is only determined by the diffusive volatility parameter, which increases both the investment boundary and the AD price. In Nishimura and Ozaki (2007), they show that drift ambiguity lowers the project value by increasing the boundary and lowering the AD price. However, in their setting, ambiguity does not affect the variance of the log EBIT process. In our case, since jump ambiguity further affects the variance, the traditional statement that real option value increases in “return standard deviation” is no longer valid under jump ambiguity. It is only valid when it is the diffusive parameter \( \sigma \) that increases the return standard deviation. This is evident by comparing the project values under alternative \( \sigma \) values.

3.4.4 The effect of agency cost

Table 7 shows the effects of agency cost on the values of innovation projects. We only alter the two agency cost parameters in our computation and keep the other parameters at their benchmark values (see Table 2). We first consider the case where \( \delta \) is 75% instead of 51%, meaning a lower agency cost. Second, we set \( \xi \) to be 5% instead of 1%, indicating a higher agency cost. We also add the results under the baseline parameters \((\delta, \xi) = (0.51, 0.01)\) for comparison. We set other parameters to their baseline values.

[Table 7 goes about here]

The results in Table 7 point out interactions among agency cost, ambiguity, the benefit of debt, and innovation type. For exploration, the relative gain due to debt is 23% under the baseline agency parameters, and it increases to 25% when \( \delta \) increases to 75% or decreases to 14%
when $\zeta$ amounts to 5%. For exploitation, the relative gain due to optimal financing is 22% under the baseline, and it increases to 23% when $\delta$ increases to 75% or decreases to 14% when $\zeta$ amounts to 5%. These findings suggest that the benefit due to debt financing in the presence of ambiguity decreases with agency cost. The interpretation of these results directly follows the responses of optimal leverage to changes in agency cost. Consistent with Morellec et al. (2012), optimal leverage hence the net tax benefit decreases with agency cost, resulting in lowered relative gain from optimal financing. Collectively, these findings reinforces the result that the fostering effect of debt financing on innovation stems from the net tax benefit mechanism.

3.4.5 Summary

In this subsection, we further establish several important results regarding the role of debt in promoting innovation. First of all, we show that access to debt financing benefits most exploratory projects, especially those concerned with extreme downside risk. Second, access to debt financing is most value-enhancing for relatively novel projects. Third, the relative importance of diffusion versus jump ambiguities depends on the nature of the project being exploratory or exploitative, and so does the added value of debt. Lastly, the innovation-promoting role of debt is more prominent in a firm with lower agency costs.

4 Conclusion

Inspired by the recent empirical literature emphasizing the importance of debt in promoting innovation, we develop a model to introduce a novel economic mechanism for the role of debt. The starting point of our model is the financing of a growth option framework introduced by Sundaresan and Wang (2007) and further developed by Sundaresan et al. (2015). We further incorporate into this framework the key characteristics of innovation highlighted by Kerr and Nanda (2015): ambiguity, skewed return, agency conflict, and intangibility. Specifically, we model the first two with jump ambiguity using the general framework of Quenez and Sulem (2013, 2014) and the last two based on Morellec et al. (2012).

Overall, our results demonstrate that ambiguity can shape agents’ preference for exploration or exploitation. We show that ambiguity increases the investment boundary, lowers the $AD$ price for investment and investment value, and these effects are stronger for exploration than exploitation. Hence, given an exploration project and an exploitation project with identical NPV when initiated now, agents facing high ambiguity are more likely to choose the latter. The reason is that the optimal investment boundary is lower while the associated $AD$ price and project value are higher for the latter.
Importantly, our results highlight that access to debt financing can promote innovation in general by accelerating the project's initiation and increasing its value through the optimal capital structure mechanism. Specifically, we show that the investment acceleration and value-enhancing effects of debt are stronger for exploration projects, especially those perceived to have negatively skewed returns \textit{ex ante}. In addition, we show that debt financing is more beneficial to relatively novel projects characterized by limited available information. Next, the added value from debt may further depend on the type of innovation and the relative concern for diffusion ambiguity or jump ambiguity. Lastly, agency concerns can lower the added value from debt.

References


A Appendix: Proofs

Proposition 1

The density generator that gives the minimum expectation in (2.17) is $\theta^* = (\theta_W^*, \theta_N^*) = (\kappa, 1 - e^{-M,t}1_{u\geq0} - 1_{u<0})$ for all $t \in [0, T]$.

Proof. We can express (2.17) as

$$V_e(0, \tau_I, \theta^*) = \inf_{Q^\theta} E_0^\theta \left[ \delta \left( \int_{\tau_I}^T e^{-r(t - \tau_I)} X(t) d t - e^{-r\tau_I} I \right) + \int_{\tau_I}^T e^{-r(t - \tau_I)} \xi(1 - \phi) X(t) d t \right], \text{ for any } \tau_I.$$ 

Since the above expectation is dynamically consistent, given our choice of $\Theta$, we can start with

$$V_e(0, \tau_I, \theta^*) = \inf_{Q^\theta} E_0^\theta \left[ \inf_{Q^\theta_{\tau_I}} \left[ \int_{\tau_I}^T e^{-r(t - \tau_I)} C_0 X(t) d t \right] - e^{-r\tau_I} \delta I \right], \text{ for any } \tau_I, \quad (A.1)$$

or

$$V_1(\tau_I, T, \theta^*) = \inf_{Q^\theta} V_1(\tau_I, T, \theta) = \inf_{Q^\theta_{\tau_I}} \left[ \int_{\tau_I}^T e^{-r(t - \tau_I)} C_0 X(t) d t \right], \quad (A.2)$$

where we use

$$C_0 = \delta(1 - \xi_\phi) + \xi(1 - \phi) > 0$$

to simplify the expressions.

Since under $Q^\theta, \{ V_1(t, T, \theta), \Sigma_1(t, \theta), K_1(t, u, \theta) \}$ is the solution to the following BSDE

$$-dV_1(t, T, \theta) = e^{-r(t - \tau_I)} C_0 X(t) d t - \Sigma_1(t, \theta) dW^\theta(t) - \int_{\mathbb{R}} K_1(t, u, \theta) N^\theta(dt, du), \quad (A.3)$$

for $t \in [\tau_I, T]$ and $V_1(T, T, \theta) = 0$. Then, by Girsanov, we have under $Q^\theta$

$$-dV_1(t, T, \theta) = \left( e^{-r(t - \tau_I)} C_0 X(t) - \Sigma_1(t, \theta) \theta W(t) - \int_{\mathbb{R}} K_1(t, u, \theta) \theta N(t, u) \nu(du) \right) dt$$

$$- \Sigma_1(t, \theta) dW(t) - \int_{\mathbb{R}} K_1(t, u, \theta) \tilde{N}(dt, du), \text{ for any } t \in [\tau, T], \text{ and } V_1(T, T, \theta) = 0. \quad (A.4)$$

Meanwhile, by Itô formula, denote $f^\theta(t, X(t)) = V_1(t, T, \theta)$, we have

$$dV_1(t, T, \theta) = f_t^\theta \theta d t + f_x^\theta(\mu^\theta X(t) d t + \sigma X(t) dW^\theta(t)) + \frac{1}{2} \sigma^2 X^2(t) f_{xx}^\theta \theta d t$$

$$+ \int_{\mathbb{R}} \left[ f_t^\theta(t, X(t^-)) + (e^u - 1) X(t^-) \right] - f^\theta(t, X(t)) \left( e^u - 1 \right) X(t^+) \nu(du) d t$$

$$+ \int_{\mathbb{R}} \left[ f_t^\theta(t, X(t^-)) + (e^u - 1) X(t^-) \right] - f^\theta(t, X(t^-)) \tilde{N}(dt, du). \quad (A.5)$$
Hence, we have

\[ \Sigma_1(t,\theta) = \sigma X(t) f^0_x(t, X(t)), \quad K_1(t, u, \theta) = f^0(t, e^u X(t^-)) - f^0(t, X(t^-)). \]  \hspace{1cm} (A.6)

Since \( V_1(t, T, \theta) = f^0(t, X(t)) \) increases in the \( x \) argument for all \( t \in [\tau_I, T] \), we have \( \Sigma_1(t, \theta) > 0 \) and \( -\Sigma_1(t, \theta) \theta_W(t) \geq -\Sigma_1(t, \theta) \kappa \) for \( \theta_W(t) \in [-\kappa, \kappa] \). Next, consider the \( dt \) term involving the integral in (A.4):

\[ I^0 = -\int_{R^+} K_1(t, u, \theta) \theta_N(t, u) v(du) = I^0_+ + I^0_-, \]  \hspace{1cm} (A.7)

where

\[ I^0_+ = -\int_{R^+} K_1(t, u, \theta) \theta_N(t, u) v(du), \quad I^0_- = -\int_{R^-} K_1(t, u, \theta) \theta_N(t, u) v(du). \]  \hspace{1cm} (A.8)

Since \( K_1(t, u, \theta) > (0) \) on \( R^+ (R^-) \) by the monotonicity of \( f^0 \) with \( x \), we have \( I^0_+ \geq \int_{R^+} K_1(t, u, \theta)(1-e^{-M_1 u}) v(du) \) and \( I^0_- \geq 0 \) for all \( t \).

Denote \( \theta^*(t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1-e^{-M_1 u} 1_{u \geq 0} - 1_{u < 0}). \) Given the above, the linear driver in (A.4) is the smallest under \( \theta^* \) for all \( (V_1(t, T, \theta), \Sigma_1(t, \theta), K_1(t, u, \theta)) \). Hence, by the comparison theorem for BSDE with jumps (Quenez and Sulem, 2013), \( V_1(\tau_I, T, \theta^*) \) is the smallest.

Next, since we can write \( V_1(\tau_I, T, \theta^*) = C_1 X(\tau_I) \), where \( C_1 > 0 \) is a constant, then \( V_e(0, \tau_I, \theta^*) \) is

\[ V_e(0, \tau_I, \theta^*) = \inf_{Q^0} \mathbb{E}_{0}^Q \left[ e^{-r \tau_I} C_1 X(\tau_I) - e^{-r \tau_I} \delta I \right]. \]  \hspace{1cm} (A.9)

Hence, under \( Q^0 \), \( (V_e(t, \tau_I, \theta), \Sigma_e(t, \theta), K_e(t, u, \theta)) \) is the solution to the following BSDE

\[ -dV_e(t, \tau_I, \theta) = -\Sigma_e(t, \theta) dW^\theta(t) - \int_{R^+} K_e(t, u, \theta) \tilde{N}^\theta(du, t), \quad t \in [0, \tau_I], \]  \hspace{1cm} (A.10)

and \( V_e(\tau_I, \tau_I, \theta) = C_1 X(\tau_I) - \delta I. \)

Again, by Girsanov, we have \( (V_e(t, \tau_I, \theta), \Sigma_e(t, \theta), K_e(t, u, \theta)) \) as the solution to the following BSDE under \( Q^0 \)

\[ -dV_e(t, \tau_I, \theta) = ( -\Sigma_e(t, \theta) \theta_W(t) - \int_{R^+} K_e(t, u, \theta) \theta_N(t, u) v(du) ) dt - \Sigma_e(t, \theta) dW(t) \]

\[ -\int_{R^-} K_e(t, u, \theta) \tilde{N}(dt, du), \quad t \in [0, \tau_I], \text{ and } V_e(\tau_I, \tau_I, \theta) = C_1 X(\tau_I) - \delta I. \]  \hspace{1cm} (A.11)

Similarly, by Itô formula, with \( V_e(t, \tau_I, \theta) = f^0(t, X(t)) \), we have the same for \( \Sigma_e(t, \theta) \) and \( K_e(t, u, \theta) \) as in (A.6).

Since (A.9) suggests that \( V_e(t, \tau_I, \theta) = f^0(t, X(t)) \) increases in the \( x \) argument for all \( t \in [0, \tau_I] \), we have \( \Pi(t, \theta) > 0 \) and \( -\Pi(t, \theta) \theta_W(t) \geq -\Pi(t, \theta) \kappa \) for \( \theta_W(t) \in [-\kappa, \kappa] \). Since \( K(t, u, \theta) > (0) \) on \( R^+ (R^-) \) by the monotonicity of \( f^0 \) with \( x \), we have \( I^0_+ \geq \int_{R^+} K(t, u, \theta)(1-e^{-M_1 u}) v(du) \) and
$I^0_t \geq 0$ for all $t$ for $I^0_s$ and $I^0_u$ defined in the same ways as (A.8).

Denote $\theta^*(t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1 - e^{-M_I u} 1_{u \geq 0} - 1_{u < 0})$ for $t \in [0, \tau_1]$. Given the above, the linear driver in (A.11) is the smallest under $\theta^*(t)$ for all $(V(t, \tau_1, \theta), \Pi(t, \theta), K(t, u, \theta))$. Hence, by the comparison theorem for BSDE with jumps, $V(0, \tau_1, \theta^*)$ is the smallest.

Taken together, $\theta^*(t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1 - e^{-M_I u} 1_{u \geq 0} - 1_{u < 0})$ is the minimum-expectation density generator for all $t \in [0, T]$. \hfill $\Box$

**Proposition 2**

Let $\eta^*_1 = \eta_1 + M_1$ and $\eta^*_2 = \eta_2$. The infinite horizon value function $V^\infty_e(0, \tau*I, \theta^*)$ as in (2.19) has the following solution

$$V^\infty_e(0, \tau*I, \theta^*) = A_0 X_I \left[ c_{1,1} \left( \frac{X^*_I}{x} \right)^{-\beta_1} + c_{2,1} \left( \frac{X^*_I}{x} \right)^{-\beta_2} \right] - \delta I \left[ c_{1,0} \left( \frac{X^*_I}{x} \right)^{-\beta_1} + c_{2,0} \left( \frac{X^*_I}{x} \right)^{-\beta_2} \right]$$

where

$$X_I^* = \frac{\delta I}{A_0 (\beta_1 - 1)(\beta_2 - 1)} \frac{\eta^*_1 - 1}{\eta^*_1}, \quad A_0 = \frac{\delta(1 - \xi \phi) + \xi(1 - \phi)}{r - \mu^\theta}, \quad X(0) = x.$$

Here, $\beta_1, \beta_2, \beta_3, \beta_4$, satisfying $-\infty < -\beta_4 < -\eta^*_2 < -\beta_3 < 0 < \beta_1 < \eta^*_1 < \beta_2 < \infty$, are the four roots of the equation $G(\beta) = r$, with

$$G(\beta) = \frac{1}{2} \sigma^2 \beta^2 + \left[ \mu^{\theta*} - \frac{1}{2} \sigma^2 - \lambda^* \left( \frac{p^* \eta^*_1}{\eta^*_1 - 1} + \frac{q^* \eta^*_2}{\eta^*_2 + 1} - 1 \right) \right] \beta + \lambda^* \left( \frac{p^* \eta^*_1}{\eta^*_1 - \beta} + \frac{q^* \eta^*_2}{\eta^*_2 + \beta} - 1 \right),$$

where

$$\mu^{\theta*} = \mu - \kappa \sigma - \lambda p'/(\eta_1 - 1) + \lambda p \eta_1 / ((\eta_1^*)^2 - \eta_1^*), \quad \lambda^* = \lambda (p \eta_1 / \eta_1^* + q),$$

$$p^* = \frac{p \eta_1}{p \eta_1 + q \eta_1^*}, \quad q^* = \frac{q \eta_1^*}{p \eta_1 + q \eta_1^*}.$$

$$c_{1,0} = \frac{\eta^*_1 - \beta_1 \beta_2}{\beta_2 - \beta_1 \eta^*_1}, \quad c_{2,0} = \frac{\beta_2 - \eta^*_1 \beta_1}{\beta_2 - \beta_1 \eta^*_1}, \quad c_{1,1} = \frac{\eta^*_1 - \beta_1 \beta_2 - 1}{\beta_2 - \beta_1 \eta^*_1 - 1}, \quad c_{2,1} = \frac{\beta_2 - \eta^*_1 \beta_1 - 1}{\beta_2 - \beta_1 \eta^*_1 - 1}.$$

**Proof.** Under $Q^{\theta*}$, $X(t)$ follows

$$dX(t)/X(t^-) = \mu^{\theta*} \, dt + \sigma \, dW^{\theta*}(t) + \int_{\mathbb{R}} \left( e^u - 1 \right) \tilde{N}^{\theta*}(dt, du), \quad (A.12)$$

where $\mu^{\theta*} = \mu - \kappa \sigma - \lambda p / (\eta_1 - 1) + \lambda p \eta_1 / ((\eta_1^*)^2 - \eta_1^*)$. 

34
By Itô formula, we have

\[ X(t) = X(0)e^{Y(t)}, \quad \text{where} \]

\[ Y(t) = (\mu^\theta - \frac{1}{2}\sigma^2)t + \sigma W(t) + t \int_R (u - e^u + 1)\bar{V}^\theta(du) + \int_0^t \int_R u\tilde{N}^\theta(ds, du). \quad (A.13) \]

Hence,

\[ \lim_{T \to \infty} \mathbb{E}_{\tau} \left[ \int_{\tau}^T e^{-rt} X(t) dt \right] = \lim_{T \to \infty} e^{-r\tau} X(\tau) \mathbb{E}_{\tau} \left[ \int_{\tau}^T e^{-r(t-\tau)} + Y(t-\tau) dt \right] \]

\[ = \lim_{T \to \infty} e^{-r\tau} X(\tau) \int_{\tau}^T \mathbb{E}_{\tau} \left[ e^{-r(t-\tau)} + Y(t-\tau) \right] dt = \lim_{T \to \infty} e^{-r\tau} X(\tau) \int_{\tau}^T e^{-(r-\mu^\theta)(t-\tau)} dt \]

\[ = \frac{e^{-r\tau} X(\tau)}{r - \mu^\theta}. \]

Then,

\[ V^\infty_e(0, \tau^*_I, \theta^*) = \sup_{\tau_I} \mathbb{E}_{0}^\theta \left[ e^{-r\tau_I} \left( A_0 X(\tau_I) - \delta I \right) \right], \quad (A.14) \]

where

\[ A_0 = \frac{\delta(1 - \xi \phi) + \xi(1 - \phi)}{r - \mu^\theta}. \]

The functional form inside the conditional expectation entails that the optimal stopping time is of the threshold type \( \tau^*_I = \inf_t \{X(t) \geq X_I^*\} \). However, due to the “overshooting” problem caused by jumps, it is also possible that \( X(\tau^*) > X_I^* \). Hence, we utilize the results from Kou and Wang (2003) for the following

\[ \mathbb{E}_{0}^\theta \left[ e^{-r\tau_I} \right] = c_{1,0} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_I}{x} \right)^{-\beta_2}, \quad (A.15) \]

and

\[ \mathbb{E}_{0}^\theta \left[ e^{-r\tau_I} X(\tau_I) \right] = X_I \left[ c_{1,1} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_I}{x} \right)^{-\beta_2} \right], \quad (A.16) \]

where \( \tau_I = \inf_t \{X(t) \geq X_I\} \) for some \( X_I \geq X(0) = x \), and \( \beta_i \) and \( c_{i,j} \) are given by (2.22) and (2.23). Hence,

\[ V^\infty_e(0, \tau_I, \theta^*) = A_0 X_I \left[ c_{1,1} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_I}{x} \right)^{-\beta_2} \right] - \delta I \left[ c_{1,0} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_I}{x} \right)^{-\beta_2} \right] \quad (A.17) \]

Therefore, we can utilize the “smooth pasting” principle, or Theorem 3.2 of Øksendal and Sulem (2019), that \( V^\infty_e(0, \tau^*_I, \theta^*) \) as a function of \( X(0) = x \) should be \( C^1 \) at \( X_I^* \) to be optimal to
find the free boundary $X_I^*$. Since
\[
\frac{\partial V_e^\infty}{\partial X} = A_1 X_I^{1-\beta_1} x^{\beta_1-1} + A_2 X_I^{1-\beta_2} x^{\beta_2-1} - A_3 X_I^{-\beta_1} x^{\beta_1-1} - A_4 X_I^{-\beta_2} x^{\beta_2-1},
\]
(A.18)
with
\[
A_1 = \beta_1 c_{1,1} A_0, \quad A_2 = \beta_2 c_{2,1} A_0, \quad A_3 = \beta_1 c_{1,0} \delta I, \quad \text{and} \quad A_4 = \beta_2 c_{2,0} \delta I,
\]
we have
\[
\frac{\partial V_e^\infty}{\partial x} \bigg|_{x=X_I} = A_1 + A_2 - A_3 X_I^{-1} - A_4 X_I^{-1}.
\]
(A.19)
At $X_I^*$, (A.19) should satisfy
\[
\frac{\partial V_e^\infty}{\partial x} \bigg|_{x=X_I^*} = A_0.
\]
(A.20)
Hence,
\[
X_I^* = \frac{A_3 + A_4}{A_1 + A_2 - A_0} = \frac{\delta I \beta_1 \beta_2}{A_0 (\beta_1 - 1) (\beta_2 - 1)} \frac{\eta^*_1 - 1}{\eta^*_1}.
\]
(A.21)

\[\square\]

**Proposition 3**

The density generator that gives the minimum expectation in (2.26) is $\theta^* = (\theta^*_W, \theta^*_N) = (\kappa, 1 - e^{-M_1 u_{t \geq 0} - 1_{u < 0}})$ for all $t \in [\tau_I, T]$.

**Proof.** Recall (2.26) is
\[
\Pi(\tau_I, \tau_D, \theta^*) = \inf_{Q^\theta} \{ \delta E(\tau_I, \tau_D, \theta) + \Gamma(\tau_I, \tau_D, \theta) \}
\]
(A.22)
where
\[
B_0 = \delta(1 - \xi_\phi) + \xi(1 - \phi), \quad B_1 = \delta(\alpha_L - \alpha_R) \{ \delta(1 - \xi_\phi) + \xi(1 - \phi) \}.
\]
Again, since the above conditional expectation is dynamically consistent, we have
\[
\inf_{Q^\theta} \mathbb{E}'_{\tau_I}\left[ B_0 \int_{\tau_I}^{T_D} e^{-r(t-\tau_I)} (X(t) - C) dt + B_1 \int_{\tau_D}^{T} e^{-r(t-\tau_I)} X(t) dt \right]
\]

36
First, like that in the proof of Proposition 1, we find \( \theta^*(t) \) for \( t \in [\tau_D, T] \). Denote

\[
V_0(\tau_D, T, \theta) = \inf_{Q^\theta} \mathbb{E}_{\tau_D} \left[ B_1 \int_{\tau_D}^T e^{-r(t)} X(t) \, dt \right].
\]

Since the above is almost the same as the value function given by (A.2), the same analysis yields that \( \theta^*(t) = (\kappa, 1 - e^{-M_1}u_{u \geq 0} - 1_{u < 0}) \) for all \( t \in [\tau_D, T] \).

Furthermore, it is clear that we can express \( V_0(\tau_D, T, \theta) \) as

\[
V_0(\tau_D, T, \theta) = B_1' e^{-r(\tau_D - t)} X(\tau_D)
\]

with some constant \( B_1' > 0 \). Hence, (A.22) becomes

\[
\Pi(\tau_I, \tau_D, \theta^*) = \inf_{Q^\theta} \mathbb{E}_{\tau_I} \left[ B_0 \int_{\tau_I}^{\tau_D} e^{-r(t)} (X(t) - C) \, dt + B_1' e^{-r(\tau_D - t)} X(\tau_D) \right]. \tag{A.23}
\]

Denote \( V_1(t, \tau_D, \theta) = \Pi(t, \tau_D, \theta) \) for \( t \in [\tau_I, \tau_D] \). Since under \( Q^\theta \), \( (V_1(t, \tau_D, \theta), \Sigma_1(t, \theta), K_1(t, u, \theta)) \) is the solution to the following BSDE

\[
-dV_1(t, \tau_D, \theta) = e^{-r(t)} B_0 (X(t) - C) \, dt - \Sigma_1(t, \theta) \, dW(t) - \int_{\mathbb{R}} K_1(t, u, \theta) \, \tilde{N}(d,t,u), \tag{A.24}
\]

for \( t \in [\tau_I, \tau_D] \), and \( V(\tau_D, \tau_D, \theta) = B_1' e^{-r(\tau_D - t)} X(\tau_D) \). Then by Girsanov, we have under \( Q^0 \)

\[
-dV_1(t, \tau_D, \theta) = \left( e^{-r(t)} B_0 (X(t) - C) - \Sigma_1(t, \theta) \theta_W(t) - \int_{\mathbb{R}} K_1(t, u, \theta) \theta_N(t,u) \, v(d,u) \right) \, dt
\]

\[
- \Sigma_1(t, \theta) \, dW(t) - \int_{\mathbb{R}} K_1(t, u, \theta) \tilde{N}(dt,d,u), \quad t \in [\tau_I, \tau_D], \tag{A.25}
\]

and \( V_1(\tau_D, \tau_D, \theta) = B_1' e^{-r(\tau_D - t)} X(\tau_D) \).

Similar to the proof of Proposition 1, by Itô formula, denote \( f^\theta(t, X(t)) = V_1(t, \tau_D, \theta) \), we have

\[
\Sigma_1(t, \theta) = \sigma X(t) f^\theta_X(t, X(t)), \quad K_1(t, u, \theta) = f^\theta(t, e^u X(t^-)) - f^\theta(t, X(t^-)). \tag{A.26}
\]

Since \( V_1(t, \tau_D, \theta) = f^\theta(t, X(t)) \) increases in the \( x \) argument for all \( t \in [\tau_I, \tau_D] \), we have \( \Sigma_1(t, \theta) > 0 \) and \(-\Sigma_1(t, \theta) \theta_W(t) \geq -\Sigma_1(t, \theta) \kappa \) for \( \theta_W(t) \in [-\kappa, \kappa] \). Next, consider the \( dt \) term involving the integral in (A.25):

\[
I^\theta = - \int_{\mathbb{R}} K_1(t, u, \theta) \theta_N(t,u) \, v(d,u) = I^\theta_+ + I^\theta_-, \tag{A.27}
\]

where

\[
I^\theta_+ = - \int_{\mathbb{R}^+} K_1(t, u, \theta) \theta_N(t,u) \, v(d,u), \quad I^\theta_- = - \int_{\mathbb{R}^-} K_1(t, u, \theta) \theta_N(t,u) \, v(d,u). \tag{A.28}
\]
Since $K_1(t,u,\theta) > (>)$ on $\mathbb{R}^+$ by the monotonicity of $f^0$ with $x$, we have $I^0_+ \geq -\int_{\mathbb{R}^+} K_1(t,u,\theta)(1-e^{-M_1u})v(du)$ and $I^0_- \geq 0$ for all $t$. Hence, the linear driver in (A.25) is the smallest under $\theta^*(t) = (\kappa, 1-e^{-M_1z}1_{u\geq0} - 1_{u<0})$ for all $(V_1(t,t,\theta), \Sigma_1(t,t,\theta), K_1(t,u,\theta))$. By the comparison theorem for BSDE with jumps, $V_1(\tau_I, \tau_D, \theta^*)$ is the smallest.

Collectively, we have shown $\theta^*(t) = (\kappa, 1-e^{-M_1u}1_{u\geq0} - 1_{u<0})$ for all $t \in [\tau_I, T]$. \hfill $\square$

### Proposition 4

Let $\beta_3, \beta_4, \mu^{0*}, \eta_1^*, \eta_2^*$ be the same as in Proposition 2. The infinite horizon value function $\Pi^{\infty}(\tau_I, \tau^*, \theta^*)$ as in (2.27) has the following expression

$$
\Pi^{\infty}(\tau_I, \tau_D^*, \theta^*) = B_0 X(\tau_I) - B_1 C + B_1 C [d_{1,0}\left(\frac{X_D}{X(\tau_I)}\right)^{\beta_3} + d_{2,0}\left(\frac{X_D}{X(\tau_I)}\right)^{\beta_4}] - (B_0 - B_2) X_D [d_{1,1}\left(\frac{X_D}{X(\tau_I)}\right)^{\beta_3} + d_{2,1}\left(\frac{X_D}{X(\tau_I)}\right)^{\beta_4}]
$$

where

$$
B_0 = \frac{\delta(1-\xi\phi) + \xi(1-\phi)}{r - \mu^{0*}}, \quad B_1 = \frac{\delta(1-\xi\phi) + \xi(1-\phi)}{r}, \quad B_2 = \frac{\theta(\alpha_L - \alpha_R)(\delta(1-\xi\phi) + \xi(1-\phi))}{r - \mu^{0*}}
$$

$$
d_{1,0} = \frac{\eta_2^* - \beta_3 \beta_4}{\beta_4 - \beta_3 \eta_2^*}, \quad d_{2,0} = \frac{\beta_4 - \eta_2^* \beta_3}{\beta_4 - \beta_3 \eta_2^*}, \quad d_{1,1} = \frac{\eta_2^* - \beta_3 \beta_4 + 1}{\beta_4 - \beta_3 \eta_2^* + 1}, \quad d_{2,1} = \frac{\beta_4 - \eta_2^* \beta_3 + 1}{\beta_4 - \beta_3 \eta_2^* + 1},
$$

and the optimal default policy is $\tau_D^* = \inf t\{|X(t)| \leq X_D^*\}$ with

$$
X_D^* = \frac{(r - \mu^{0*})}{r(1 - \theta(\alpha_L - \alpha_R))} \frac{\beta_3 \beta_4 (\eta_2^* + 1)}{\beta_3 + 1 (\beta_4 + 1) \eta_2^*} C.
$$

**Proof.** Recall that

$$
\Pi(\tau_I, \tau_D, \theta^*) = \mathbb{E}_{\tau_I}^\theta \left[ b_0 \int_{\tau_I}^{T_D} e^{-r(t-\tau_I)} X(t) dt + b_1 \int_{\tau_D}^{T} e^{-r(t-\tau_D)} X(t) dt \right],
$$

where

$$
b_0 = \delta(1-\xi\phi) + \xi(1-\phi), \quad b_1 = \theta(\alpha_L - \alpha_R)(\delta(1-\xi\phi) + \xi(1-\phi)).
$$

Under $Q^{\theta^*}$, $X(t)$ follows

$$
dX(t)/X(t^-) = \mu^{0*} dt + \sigma dW^{\theta^*}(t) + \int_{\mathbb{T}} (e^u - 1) \tilde{N}^{\theta^*}(dt, du),
$$

(A.29)

where $\mu^{0*} = \mu - \kappa \sigma - \lambda p l/(\eta_1 - 1) + \lambda p n \eta_1 /((\eta_1^*)^2 - \eta_1^*)$.
According to Dynkin’s formula, the first term of $\Pi(\tau_I, \tau_D, \theta^*)$ becomes

$$
\mathbb{E}_{\tau_I}^{\theta^*} \left[ b_0 \int_{\tau_I}^{\tau_D} e^{-r(t-\tau_I)} (X(t) - C) dt \right] = B_0 X(\tau_I) - b_1 C - \mathbb{E}_{\tau_I}^{\theta^*} \left[ e^{-r(\tau_D-\tau_I)} \left( B_0 X(\tau_D) - B_1 C \right) \right],
$$

where

$$
B_0 = \frac{b_0}{r - \mu^\theta} , \quad \text{and} \quad B_1 = \frac{b_0}{r}.
$$

The second term is

$$
\mathbb{E}_{\tau_I}^{\theta^*} \left[ b_1 \int_{\tau_D}^{T} e^{-r(t-\tau_I)} X(t) dt \right] = \mathbb{E}_{\tau_I} \left[ B_2' e^{-r(\tau_D-\tau_I)} X(\tau_D) \right],
$$

where

$$
B_2' = \frac{b_1}{r - \mu^\theta} \left( 1 - e^{-(r - \mu^\theta)(T - \tau_D)} \right)
$$

Then,

$$
\Pi^\infty(\tau_I, \tau_D^*, \theta^*) = \sup_{\tau_D \geq \tau_I} B_0 X(\tau_I) - B_1 C + \mathbb{E}_{\tau_I}^{\theta^*} \left[ e^{-r(\tau_D-\tau_I)} \left( B_1 C - (B_0 - B_2) X(\tau_D) \right) \right],
$$

(A.30)

given that $Q^{\theta^*} \{ \tau_D^* < \infty \} = 1$, and we have replaced

$$
B_2 = \lim_{T \to \infty} B_2' = \frac{b_1}{r - \mu^\theta}.
$$

It suffices to focus on the term that only involves $\tau_D$, and we term

$$
V^\infty(\tau_I, \tau_D^*, \theta^*) = \sup_{\tau_D \geq \tau_I} \mathbb{E}_{\tau_I}^{\theta^*} \left[ e^{-r(\tau_D-\tau_I)} \left( B_1 C - (B_0 - B_2) X(\tau_D) \right) \right].
$$

The functional form inside the conditional expectation entails that the optimal stopping time is of the threshold type $\tau_D^* = \inf_t \{ X(t) \leq X_D^* \}$. However, due to the “overshooting” problem caused by jumps, it is also possible that $X(\tau_D^*) < X_D^*$. Hence, we utilize the results from Kou and Wang (2003) for the following

$$
\mathbb{E}_{\tau_I}^{\theta^*} \left[ e^{-r(\tau_D-\tau_I)} \right] = d_{1,0} \left( \frac{X_D}{x} \right)^{\beta_3} + d_{2,0} \left( \frac{X_D}{x} \right)^{\beta_4},
$$

and

$$
\mathbb{E}_{\tau_I}^{\theta^*} \left[ e^{-r(\tau_D-\tau_I)} X(\tau_D) \right] = X_D \left[ d_{1,1} \left( \frac{X_D}{x} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D}{x} \right)^{\beta_4} \right],
$$

where $\tau_D = \inf_t \{ X(t) \leq X_D \}$ for some $X_D \leq X(\tau_I) = x$ and $d_{i,j}$ and $\beta_i$ are given by (2.29) and
(2.22). Hence,
\[
V^\infty(\tau_I, \tau_D, \theta^*) = B_1 C \left[ d_{1,0} \left( \frac{X_D}{x} \right)^{\beta_3} + d_{2,0} \left( \frac{X_D}{x} \right)^{\beta_4} \right] - (B_0 - B_2) X_D \left[ d_{1,1} \left( \frac{X_D}{x} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D}{x} \right)^{\beta_4} \right].
\]

Therefore, we can utilize the “smooth pasting” principle that \( V^\infty(\tau_I, \tau_D^*, \theta^*) \) as a function of \( X(\tau_I) = x \) should be \( C^1 \) at \( X_D^* \) to be optimal to find the free boundary \( X_D^* \). Since
\[
\frac{\partial V^\infty}{\partial x} \bigg|_{x = X_D^*} = -B_1 (\beta_3 d_{1,0} + \beta_4 d_{2,0}) C / X_D^* + (B_0 - B_2) (\beta_3 d_{1,1} + \beta_4 d_{2,1}),
\]
and
\[
\frac{\partial V^\infty}{\partial x} \bigg|_{x = X_D^*} = -(B_0 - B_2).
\]

We must have
\[
-B_1 (\beta_3 d_{1,0} + \beta_4 d_{2,0}) C / X_D^* + (B_0 - B_2) (\beta_3 d_{1,1} + \beta_4 d_{2,1}) = -(B_0 - B_2).
\]

Hence,
\[
X_D^* = \frac{B_1}{B_0 - B_2} \frac{\beta_3 d_{1,0} + \beta_4 d_{2,0}}{1 + \beta_3 d_{1,1} + \beta_4 d_{2,1}} C = \frac{r - \mu^*}{r (1 - \theta (\alpha_L - \alpha_R))} \frac{\beta_3 d_{1,1} + \beta_4 d_{2,1}}{(\beta_3 + 1) (\beta_4 + 1) \eta^*} C.
\]

\( \square \)

**Proposition 5**

The density generator that gives the minimum expectation in (2.31) is \( \theta^* = (\theta_W^*, \theta_N^*) = (\kappa, 1 - e^{-M_i} 1_{u \geq 0} - 1_{x < 0}) \) for all \( t \geq \tau_I \). Furthermore, the value of debt has the expression
\[
D^\infty(\tau_I, \tau_D^*, \theta^*) = \frac{C}{r} \left[ 1 - d_{1,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} - d_{2,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right] + \frac{a_0 X_D^*}{r - \mu^*} \left[ d_{1,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right],
\]
where \( d_{i,j} \) is given by (2.29), \( X_D^* \) is given by (2.31), and \( a_0 = (1 - \alpha_R - \theta (\alpha_L - \alpha_R)) (1 - \xi_{\phi}) > 0 \).

**Proof.** Recall that
\[
D^\infty(\tau_I, \tau_D^*, \theta^*) = \inf_{Q^\theta} \mathbb{E}^\theta \left[ \int_{\tau_I}^{\tau_D^*} e^{-r(t - \tau_I)} C d t + a_0 \int_{\tau_D^*}^{\infty} e^{-r(t - \tau_I)} X(t) d t \right],
\]
where
\[
a_0 = (1 - \alpha_R - \theta (\alpha_L - \alpha_R)) (1 - \xi_{\phi}) > 0.
\]

For the first part of the statement, the idea is the same as that of Proposition 1, i.e., to find the
minimum possible linear driver of the associated BSDEs. We begin with an arbitrary large horizon \( T \). Since the horizon consists of two parts, \([\tau_I, \tau^*_D]\) and \([\tau^*_D, T]\), we examine them separately as in Proposition 1.

Because of dynamic consistency, we first evaluate

\[
V_1(\tau^*_D, T, \theta^*) = \inf_{Q^0} a_0 E_{\tau^*_D} \left[ \int_{\tau^*_D}^{T} e^{-r t} X(t) \, dt \right].
\]  \hfill (A.35)

Since the functional form is the same as the one in (A.2), we can know that \( V_1 \) is the smallest under \( \theta^*(t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1 - e^{-M_1 z} \mathbf{1}_{u \geq 0} - \mathbf{1}_{u < 0}) \) using the same analysis as in Proposition 1. Hence,

\[
V_1(\tau^*_D, T, \theta^*) = \frac{a_0 X(\tau^*_D)}{r - \mu^{\theta^*}} \left( 1 - e^{-(r - \mu^{\theta^*})(T - \tau^*_D)} \right),
\]  \hfill (A.36)

where \( \mu^{\theta^*} = \mu - \kappa \sigma - \lambda p l (\eta_1 - 1) + \lambda \rho \eta_1 / (\eta_1^3 - \eta_1) \) for \( \eta_1^* = \eta_1 + M_1 \).

Consequently, we can write (2.31) for a finite \( T \) as

\[
D^T(\tau_I, \tau^*_D, \theta^*) = \inf_{Q^0} \mathbb{E}^0_{\tau_I} \left[ \frac{C}{r} \left( 1 - e^{-r (\tau^*_D - \tau_I)} \right) + e^{-r (\tau^*_D - \tau_I)} \frac{a_0 X(\tau^*_D)}{r - \mu^{\theta^*}} \left( 1 - e^{-(r - \mu^{\theta^*})(T - \tau^*_D)} \right) \right].
\]

Hence, denote \( V(t, \tau^*_D, \theta) = D^T(t, \tau^*_D, \theta) \), we have \( \{ V(t, \tau^*_D, \theta), \Sigma(t, \theta), K(t, u, \theta) \} \) be the solution of the following BSDE under \( Q^0 \)

\[
-dV(t, \tau^*_D, \theta) = -\Sigma(t, \theta) dW^\theta(t) - \int_{\mathbb{R}} K(t, u, \theta) \tilde{N}(d t, d u), \quad t \in [\tau_I, \tau^*_D],
\]

\[
V(\tau^*_D, \tau^*_D, \theta) = \frac{C}{r} \left( 1 - e^{-r (\tau^*_D - \tau_I)} \right) + e^{-r (\tau^*_D - \tau_I)} \frac{a_0 X(\tau^*_D)}{r - \mu^{\theta^*}} \left( 1 - e^{-(r - \mu^{\theta^*})(T - \tau^*_D)} \right).
\]

By Girsanov, we have \( \{ V(t, \tau^*_D, \theta), \Sigma(t, \theta), K(t, u, \theta) \} \) as the solution to the following BSDE under \( Q^0 \)

\[
-dV(t, \tau^*_D, \theta) = -\Sigma(t, \theta) w_W(t) - \int_{\mathbb{R}} K(t, u, \theta) \tilde{N}(t, u) v(d u) d t - \Sigma(t, \theta) dW(t) - \int_{\mathbb{R}} K(t, u, \theta) \tilde{N}(d t, d u),
\]

for \( t \in [\tau_I, \tau^*_D] \) with the same terminal condition at \( \tau^*_D \).

Since the above has the same form as (A.11), we can know that \( V \) is the smallest under \( \theta^*(t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1 - e^{-M_1 z} \mathbf{1}_{u \geq 0} - \mathbf{1}_{u < 0}) \) using the same analysis as in Proposition 1. Collectively, for all \( T \) and \( t \geq \tau_I \), \( D^T(t, \tau^*_D, \theta^*) \) is the smallest, so is \( D^\infty(\tau_I, \tau^*_D, \theta^*) \) for all \( t \geq \tau_I \).

Now we can evaluate \( D^\infty(\tau_I, \tau^*_D, \theta^*) \) as

\[
D^\infty(\tau_I, \tau^*_D, \theta^*) = \frac{C}{r} \left( 1 - e^{-\theta^*_D(\tau^*_D - \tau_I)} \right) + \frac{a_0}{r - \mu^{\theta^*}} \mathbb{E}^0_{\tau_I} \left[ e^{-r (\tau^*_D - \tau_I)} X(\tau^*_D) \right].
\]  \hfill (A.37)
Again, we use the results for the first passage probabilities to have

\[ D^\infty(\tau_I, \tau_D^*, \theta^*) = \frac{C}{r} \left[ 1 - d_{1,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} - d_{2,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right] + \frac{a_0 X_D^*}{r - \mu^{\theta^*}} \left[ d_{1,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right]. \]

\[ \square \]

**Proposition 6**

The \( C^* \) that maximizes \( V^\infty(\tau_I) \) in (2.33) has the form \( C^* = \psi X(\tau_I) \) where \( \psi \) is the unique positive solution to the following equation

\[ B_3 = B_5 \psi^{\beta_3} + B_6 \psi^{\beta_4}. \]

where

\[ B_3 = \frac{\delta \xi - \xi(1 - \phi)}{r}, \quad B_4 = \frac{\delta \alpha_R(1 - \xi) + \xi(1 - \phi) (1 - \theta(\alpha_L - \alpha_R))}{r - \mu^{\theta^*}}, \quad B_5 = B_3 d_{1,0} (1 + \beta_3) \gamma^{\beta_3} + B_4 d_{1,1} (1 + \beta_3) \gamma^{1 + \beta_3}, \quad B_6 = B_3 d_{2,0} (1 + \beta_4) \gamma^{\beta_4} + B_4 d_{2,1} (1 + \beta_4) \gamma^{1 + \beta_4}, \]

\( \beta_i, d_{i,j} \) are given by (2.29), and \( \gamma \) satisfies \( X^*_D = \gamma C \) as in (2.30).

**Proof.** Using the expressions for \( \Pi^\infty(\tau_I) \) and \( D^\infty(\tau_I) \) given by Propositions 4 and 5, we have

\[ V^\infty(\tau_I, C) = \Pi^\infty(\tau_I) + \delta D^\infty(\tau_I) \]

\[ = B_0 X(\tau_I) + \frac{\delta}{r} \left[ 1 - B_1 C(1 - \mathbb{E}_{\tau_I}(e^{-r(\tau_D^* - \tau_I)}) - (B_0 - B_2 - \frac{\delta a_0}{r - \mu^{\theta^*}}) \mathbb{E}_{\tau_I}[X(\tau_D^*) e^{-r(\tau_D^* - \tau_I)}]) \right] \]

\[ = B_0 X(\tau_I) + B_3 C \left[ 1 - \mathbb{E}_{\tau_I}(e^{-r(\tau_D^* - \tau_I)}) \right] - B_4 \mathbb{E}_{\tau_I}[X(\tau_D^*) e^{-r(\tau_D^* - \tau_I)}] \]

where

\[ B_3 = \frac{\delta}{r} - B_1 = \frac{\delta \xi - \xi(1 - \phi)}{r}, \quad B_4 = B_0 - B_2 - \frac{\delta a_0}{r - \mu^{\theta^*}} = \frac{\delta \alpha_R(1 - \xi) + \xi(1 - \phi) (1 - \theta(\alpha_L - \alpha_R))}{r - \mu^{\theta^*}}. \]

Note that we have omitted \( \delta I \) because it does not depend on \( C \).

Denote \( X^*_D = \gamma C \) as in the following

\[ V^\infty(\tau_I, C) = B_0 X(\tau_I) + B_3 C \left[ 1 - d_{1,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} - d_{2,0} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right] - B_4 X_D^* \left[ d_{1,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_3} + d_{2,1} \left( \frac{X_D^*}{X(\tau_I)} \right)^{\beta_4} \right]. \]
Then, the first order necessary condition for $C^*$ is

$$B_3 - B_3 d_{1,0}(1 + \beta_3) \left( \frac{\gamma C^*}{X(\tau_I)} \right)^{\beta_3} - B_3 d_{2,0}(1 + \beta_4) \left( \frac{\gamma C^*}{X(\tau_I)} \right)^{\beta_4} = 0,$$  \hspace{1cm} (A.38)

for all $X(\tau_I)$ values.

Denote $C^* = X(\tau_I) f(X(\tau_I))$. It suffices to show that $f(X(\tau_I))$ is a constant. Let $g(f(X(\tau_I)))$ be

$$g(f(X(\tau_I))) = B_3 - B_3 d_{1,0}(1 + \beta_3) \left( \gamma f(X(\tau_I)) \right)^{\beta_3} - B_3 d_{2,0}(1 + \beta_4) \left( \gamma f(X(\tau_I)) \right)^{\beta_4}$$

$$- B_4 d_{1,1} \gamma (1 + \beta_3) \left( \gamma f(X(\tau_I)) \right)^{\beta_3} - B_4 d_{2,1} \gamma (1 + \beta_4) \left( \gamma f(X(\tau_I)) \right)^{\beta_4}.$$ \hspace{1cm} (A.39)

Then, the derivative of $g$ with respect to the $f(X(\tau_I))$ argument is

$$g'(f(X(\tau_I))) = -B_4 d_{1,1} \beta_3 (1 + \beta_3) \gamma^{1+\beta_3} f(X(\tau_I))^{\beta_3-1} - B_4 d_{2,1} \beta_4 (1 + \beta_4) \gamma^{1+\beta_4} f(X(\tau_I))^{\beta_4-1}$$

$$- B_4 d_{1,1} \beta_3 (1 + \beta_3) \gamma^{1+\beta_3} f(X(\tau_I))^{\beta_3-1} - B_4 d_{2,1} \beta_4 (1 + \beta_4) \gamma^{1+\beta_4} f(X(\tau_I))^{\beta_4-1}.$$ \hspace{1cm} (A.40)

Note that if $B_3 > 0$, then $g'(f(X(\tau_I))) < 0$ for all $f(X(\tau_I)) > 0$ since $B_4 > 0$. Otherwise, it is not straightforward to determine the sign of $g'(f(X(\tau_I)))$. Fortunately, so long as $\delta$ is not unreasonably small and $\xi$ is not unreasonably large, we have $B_3 > 0$. In this case, $g(f(X(\tau_I)))$ is strictly monotonic in $\mathbb{R}^+$, meaning that $g(f(X(\tau_I)))$ has a unique root in $\mathbb{R}^+$ if it exists. Clearly, $g(0) > 0$ and $g(\infty) = -\infty$, then the root exists.

Thus, we have shown that $f(X(\tau_I)) = \psi$ or $C^* = \psi X(\tau_I)$, where $\psi$ satisfies

$$B_3 = (B_4 d_{1,1} (1 + \beta_3) \gamma^{1+\beta_3} + B_4 d_{2,1} (1 + \beta_3) \gamma^{1+\beta_3}) \psi^{\beta_3} + (B_3 d_{2,0} (1 + \beta_4) \gamma^{1+\beta_4} + B_4 d_{2,1} (1 + \beta_4) \gamma^{1+\beta_4}) \psi^{\beta_4}.$$

Furthermore, the above analysis suggests that $V^\infty_\psi(\tau_I)$ is concave in $C$, sufficient for $C^*$ to be value maximizing. \hfill\Box

**Proposition 7**

The density generator that gives the minimum expectation in (2.36) is $\theta^* = (\theta^*_W, \theta^*_N) = (\kappa, 1 - e^{-M t} 1_{u \geq 0} - 1_{u < 0})$ for all $t \in [0, \tau^*_I]$. Furthermore, the value of the growth option has the expression

$$V^\infty_\psi(0, \tau^*_I, \theta^*) = B_7 X_I \left[ c_{1,1} \left( \frac{X_I}{X} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_I}{X} \right)^{-\beta_2} \right] - \delta I \left[ c_{1,0} \left( \frac{X_I}{X} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_I}{X} \right)^{-\beta_2} \right],$$
and the investment boundary satisfies

$$X^*_I = \frac{\delta I - \beta_1 \beta_2 \eta^*_1 - 1}{B_I (\beta_1 - 1) (\beta_2 - 1)}.$$

In the above, $c_{i,j}$, $\beta_i$, and $\eta^*_i$ are the same as in Proposition 2, and $B_I$ is from (2.35).

**Proof.** Since the value function (2.36) is of the same type as the one in Proposition 1 Equation (A.9), we can use the same analysis to show that $\theta^* (t) = (\theta^*_W(t), \theta^*_N(t, u)) = (\kappa, 1 - e^{-M_0 u} \mathbf{1}_{u \geq 0} - \mathbf{1}_{u < 0})$ is the minimum-expectation density generator for all $t \in [0, \tau_I]$. Hence, we can write (2.36) as

$$V^\infty_*(0, \tau_I^*, \theta^*) = \sup_{\tau_I} \mathbb{E}^\theta_0 \left[ e^{-\tau_I (B_I X(\tau_I) - \delta I)} \right]. \quad (A.41)$$

Clearly, the functional form inside the conditional expectation operator indicates that $\tau_I^*$ is of the threshold type. Then, we follow the same procedure as before by evaluating the value function under an arbitrary boundary first. Let $\tau_I = \inf_t \{X(t) \geq X_I\}$. Then,

$$V^\infty_*(0, \tau_I, \theta^*) = \mathbb{E}^\theta_0 \left[ e^{-\tau_I (B_I X(\tau_I) - \delta I)} \right]$$

$$= B_I X_I \left[ c_{1,1} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,1} \left( \frac{X_I}{x} \right)^{-\beta_2} \right] - \delta I \left[ c_{1,0} \left( \frac{X_I}{x} \right)^{-\beta_1} + c_{2,0} \left( \frac{X_I}{x} \right)^{-\beta_2} \right],$$

where $X(0) = x$, $c_{i,j}$, and $\beta_i$ are the same as in Proposition 2.

Therefore, we can utilize the “smooth pasting” principle that $V^\infty_*(0, \tau_I^*, \theta^*)$ as a function of $x$ should be $C^1$ at $X^*_I$ to be optimal to find the free boundary $X^*_I$. Since

$$\frac{\partial V^\infty_*}{\partial x} = C'_1 X_I^{-\beta_1} x^{\beta_1 - 1} + C'_2 X_I^{-\beta_2} x^{\beta_2 - 1} - C'_3 X_I^{-\beta_1} x^{\beta_1 - 1} - C'_4 X_I^{-\beta_2} x^{\beta_2 - 1}, \quad (A.42)$$

with

$$C'_1 = \beta_1 c_{1,1} B_I, \quad C'_2 = \beta_2 c_{2,1} B_I, \quad C'_3 = \beta_1 \delta I c_{1,0}, \quad \text{and} \quad C'_4 = \beta_2 \delta I c_{2,0},$$

we have

$$\frac{\partial V^\infty_*}{\partial x} \bigg|_{x=X_I} = C'_1 + C'_2 - C'_3 X_I^{-1} - C'_4 X_I^{-1}. \quad (A.43)$$

At $X^*_I$, (A.43) should satisfy

$$\frac{\partial V^\infty_*}{\partial x} \bigg|_{x=X_I^*} = \frac{\partial V^\infty_*}{\partial x} \bigg|_{x=X_I^*} = B_I. \quad (A.44)$$

Hence,

$$X^*_I = \frac{C'_3 + C'_4}{C'_1 + C'_2 - B_I}. \quad (A.45)$$
The numerator simplifies to
\[ C_3' + C_4' = \frac{\delta I \beta_1 \beta_2}{\eta_1^*}, \]
and the denominator simplifies to
\[ C_1' + C_2' - B_7 = B_7 \frac{\beta_2 - 1)(\beta_1 - 1)}{\eta_1^* - 1}. \]
Hence,
\[ X_I^* = \frac{\delta I \beta_1 \beta_2}{B_7 (\beta_1 - 1)(\beta_2 - 1)} \frac{\eta_1^* - 1}{\eta_1^*}. \]  (A.46)
## B Detection-error probability

The key to this probability is the log of the Radon-Nikodym derivative $\zeta^0(t) = \ln(Z^0(t))$, where $Z^0(t) = dQ^0 / dQ^0$. Under $Q^0$, $Z^0(t)$ is

$$dZ^0(t)/Z^0(t) = -\theta_W(t)dW(t) - \int_{\mathbb{R}} \theta_N(t)\check{N}(dt,du).$$

Since $\theta^*(t) = \theta^*_W(t), \theta^*_N(t) = (\kappa, 1 - e^{-M_1 \eta_1} \mathbb{1}_{u \geq 0})$ for all $t$, we have

$$dZ^0*(t)/Z^0*(t) = -\kappa dW(t) - \int_{\mathbb{R}^+} (1 - e^{-M_1 \eta_1})\check{N}(dt,du). \quad (B.1)$$

Hence, by Itô formula, we have

$$Z^0*(t) = \exp\left\{ - \int_0^t \kappa dW(s) - \frac{1}{2} \kappa^2 t + \int_0^t \int_{\mathbb{R}^+} [\ln(1 - (1 - e^{-M_1 \eta_1})) + 1 - e^{-M_1 \eta_1}] \nu(du)ds \right\}$$

$$= \exp\left\{ - \kappa W(t) - \frac{1}{2} \kappa^2 t - \int_0^t \int_{\mathbb{R}^+} M_1 u \check{N}(ds,du) - t \int_{\mathbb{R}^+} (M_1 u - 1 + e^{-M_1 \eta_1}) \lambda p \eta_1 e^{-\eta_1 u}du \right\}$$

$$= \exp\left\{ - \kappa W(t) - \frac{1}{2} \kappa^2 t - \int_0^t \int_{\mathbb{R}^+} M_1 u \check{N}(ds,du) - t \lambda p \eta_1 \left( \frac{M_1}{\eta_1^2} - \frac{1}{\eta_1} + \frac{1}{\eta_1 + M_1} \right) \right\} \quad (B.2)$$

Thus,

$$\zeta^0*(t) = - \left( \frac{\kappa^2}{2} + \lambda p \eta_1 \left( \frac{M_1}{\eta_1^2} - \frac{1}{\eta_1} + \frac{1}{\eta_1 + M_1} \right) \right) t - \kappa W(t) - \int_0^t \int_{\mathbb{R}^+} M_1 u \check{N}(ds,du). \quad (B.3)$$

### B.1 Relative entropy growth

Given an alternative measure $Q^\theta$ and the reference measure $Q^0$, we can write the growth in entropy of $Q^\theta$ relative to $Q^0$ over the time interval $[t, t + \Delta t]$ as

$$G(t, t + \Delta t) = \mathbb{E}_t^{Q^\theta} \left[ \ln \left( \frac{Z^0(t + \Delta t)}{Z^0(t)} \right) \right], \quad \mathbb{R}(Z^\theta_t) = \lim_{\Delta t \to 0} \frac{G(t, t + \Delta t)}{\Delta t}, \quad t \geq 0. \quad (B.4)$$

To calculate the above, it suffices to write $Z^0(t)$, especially $Z^0*(t)$, under $Q^\theta$ or $Q^0$. Using
Girsanov, we can write (B.1) under \(Q^\theta^*\) as

\[
d Z^{\theta^*}(t)/Z^{\theta^*}(t^-) = -\kappa (dW^{\theta^*}(t) - \kappa dt) - \int_{\mathbb{R}^+} (1 - e^{-M_1 u}) \left( \tilde{N}^{\theta^*}(dt, du) - (1 - e^{-M_1 u}) v(du) dt \right)
\]

\[
= \left( \kappa^2 + \int_{\mathbb{R}^+} (1 - e^{-M_1 u})^2 v(du) \right) dt - \kappa dW^{\theta^*}(t) - \int_{\mathbb{R}^+} (1 - e^{-M_1 u}) \tilde{N}^{\theta^*}(dt, du)
\]

\[
= \alpha_{\theta^*} dt - \kappa dW^{\theta^*}(t) - \int_{\mathbb{R}^+} (1 - e^{-M_1 u}) \tilde{N}^{\theta^*}(dt, du).
\]

(B.5)

with

\[
\alpha_{\theta^*} = \kappa^2 + \left( \frac{1}{\eta_1} - \frac{2}{\eta_1 + M_1} + \frac{1}{\eta_1 + 2M_1} \right) \lambda \rho \eta_1,
\]

(B.6)

and \(W^{\theta^*}(t)\) and \(\tilde{N}^{\theta^*}(dt, du)\) being standard Brownian motion and compensated Poisson random measure under \(Q^{\theta^*}\). Here, \(\tilde{N}^{\theta^*}(dt, du)\) has a Lévy measure \(\nu^{\theta^*}(du)\) given by

\[
\nu^{\theta^*}(du) = \lambda^* f_u^{\theta^*}
\]

(B.7)

where

\[
\lambda^* = \lambda (p\eta_1 / (\eta_1 + M_1) + q),
\]

(B.8)

\[
f_u^{\theta^*} = p_L \eta_1^* e^{-\eta_1^* u} 1_{u \geq 0} + q_M \eta_2^* e^{\eta_2^* u} 1_{u < 0},
\]

(B.9)

where \(p_L\) (\(q_M\)) denotes the lower (upper) bound for \(p\) (\(q\)), and one can refer to Table 1.

Hence, under \(Q^{\theta^*}\) and by Itô formula, we have

\[
Z^{\theta^*}(t) = \exp \left( \alpha_{\theta^*} t - \frac{1}{2} \kappa^2 t - \kappa W^{\theta^*}(t) + \int_0^t \int_{\mathbb{R}^+} \ln(1 - (1 - e^{-M_1 u})) \tilde{N}^{\theta^*}(ds, du) \right)
\]

\[
+ \int_0^t \int_{\mathbb{R}^+} \left( \ln(1 - (1 - e^{-M_1 u})) + 1 - e^{-M_1 u} \right) \nu^{\theta^*}(du)
\]

\[
= \exp \left( (\alpha_{\theta^*} t - \frac{1}{2} \kappa^2 t - \kappa W^{\theta^*}(t) - \int_0^t \int_{\mathbb{R}^+} (M_1 u - 1 + e^{-M_1 u}) \lambda^* p_L \eta_1^* e^{-\eta_1^* u} du \right)
\]

\[
- \int_0^t \int_{\mathbb{R}^+} M_1 u \tilde{N}^{\theta^*}(ds, du)
\]

\[
= \exp \left( (\alpha_{\theta^*} t - \frac{1}{2} \kappa^2 t - \kappa W^{\theta^*}(t) - \left( \frac{M_1}{\eta_1^*} - \frac{1}{\eta_1^* + M_1} \right) t \right)
\]

\[
- \int_0^t \int_{\mathbb{R}^+} M_1 u \tilde{N}^{\theta^*}(ds, du)
\]

\[
= \exp \left( \alpha_{\theta^*} t - \kappa W^{\theta^*}(t) - \int_0^t \int_{\mathbb{R}^+} M_1 u \tilde{N}^{\theta^*}(ds, du) \right).
\]

(B.10)

47
\[ \alpha_{\theta^*} = \frac{1}{2} \kappa^2 + \left( \frac{1}{\eta_1} - \frac{2}{\eta_1 + M_1} + \frac{1}{\eta_1 + 2M_1} \right) \lambda p \eta_1 - \left( \frac{M_1}{\eta_1^*} - \frac{1}{\eta_1^* + M_1} \right) \lambda^* p \eta_1^*. \] (B.11)

Thus,
\[ \ln \left( \frac{Z^{\theta^*}(t + \Delta t)}{Z^{\theta^*}(t)} \right) = \alpha_{\theta^*} \Delta t - \kappa \left( W^{\theta^*}(t + \Delta t) - W^{\theta^*}(t) \right) - \int_t^{t+\Delta t} \int_{\mathbb{R}^+} M_1 u \tilde{N}^{\theta^*}(ds, du), \] (B.12)
\[ G(t, t + \Delta t) = \mathbb{E}_t^Q \left[ \ln \left( \frac{Z^{\theta^*}(t + \Delta t)}{Z^{\theta^*}(t)} \right) \right] = \alpha_{\theta^*} \Delta t, \] (B.13)
and
\[ \Re(Z_{t}^{\theta^*}) = \lim_{\Delta t \to 0} \frac{G(t, t + \Delta t)}{\Delta t} = \alpha_{\theta^*}. \] (B.14)

Therefore, the constraint becomes
\[ \Re(Z_{t}^{\theta^*}) \leq h. \] (B.15)

Here the value of \( \eta \) is to be determined by detection-error probabilities. For example, we set \( h \) such that the detection-error probability is at least 0.1. Furthermore, we can decompose \( h = h_W + h_N \), meaning that we restrict the robustness concerns for diffusion ambiguity and jump ambiguity explicitly. In this regard, \( \kappa^* \) and \( M_1^* \) will depend on \( h_W \) and \( h_N \), respectively.

### B.2 Detection-error probabilities

The detection-error probability is defined as
\[ \pi(t, n; \eta) = \frac{1}{2} \left[ Q^0 [\zeta^{\theta^*}(n) > 0 | \mathcal{F}_t] + Q^{\theta^*} [\zeta^{\theta^*}(n) < 0 | \mathcal{F}_t] \right], \quad t \geq 0, n = mT \] (B.16)
where \( T \) is the number of years and \( m \) is the sampling frequency, and \( \eta \) denotes the upper bound for total robustness concern. Maenhout (2006) and Aït-Sahalia and Matthys (2019) provide a way to calculate this probability based on the characteristic functions of \( \zeta^{\theta^*}(t) \) under \( Q^0 \) and \( Q^{\theta^*} \). That is
\[ \pi(t, n; \eta) = \frac{1}{2} - \frac{1}{2 \pi} \int_0^\infty \left( \Re \left[ \frac{\phi_{\theta^*}(u, t, n)}{iu} \right] - \Re \left[ \frac{\phi_0(u, t, n)}{iu} \right] \right) du, \] (B.17)
where \( i = \sqrt{-1} \) and \( \Re(\cdot) \) denotes the real part of a complex number. Next, we can use the Lévy-Khintchine formula to calculate the two characteristic functions.
Nevertheless, explicitly, we can write

$$\phi_0(u, t, n) = \mathbb{E}_t^{Q^0}[e^{iu\phi^0(n)}]$$

$$= E_t^{Q^0}\left[\exp \left( iu \left( \alpha_0(n-t) - \kappa(W(n) - W(t)) - \int_t^n \int_{\mathbb{R}^+} M_1 v \tilde{N}(ds, dv) \right) \right) \right]$$

$$= \exp \left( iu \alpha_0(n-t) - \frac{\kappa^2 u^2}{2} (n-t) + (n-t) \int_{\mathbb{R}^+} (e^{-iM_1 uv} - 1 + iM_1 uv) v(dv) \right)$$

$$= \exp \left( (n-t) \left( iu \alpha_0 - \frac{\kappa^2 u^2}{2} + \int_{\mathbb{R}^+} (e^{-iM_1 uv} - 1 + iM_1 uv) \lambda p \eta_1 e^{-\eta_1 v} dv \right) \right)$$

$$= \exp \left( (n-t) \left( iu \alpha_0 - \frac{\kappa^2 u^2}{2} + \lambda p \eta_1 \left( \frac{1}{iM_1 + \eta_1} - \frac{1}{\eta_1} + \frac{1}{\eta_1 + M_1} \right) \right) \right),$$

where

$$\alpha_0 = - \left( \frac{\kappa^2}{2} + \lambda p \eta_1 \left( \frac{M_1}{\eta_1^2} - \frac{1}{\eta_1} + \frac{1}{\eta_1 + M_1} \right) \right).$$

For $\phi^{\theta^*}(u, t, n)$ we have

$$\phi^{\theta^*}(u, t, n) = \mathbb{E}_t^{Q^{\theta^*}}[e^{iu\phi^{\theta^*}(n)}] = \mathbb{E}_t^{Q^0}[e^{i\alpha_n(n)} e^{\theta^*}(n)] = \mathbb{E}_t^{Q^0}[e^{(i+1)\phi^{\theta^*}(n)}]$$

$$= \mathbb{E}_t^{Q^0}\left[\exp \left( (i+1) \left( \alpha_0(n-t) - \kappa(W(n) - W(t)) - \int_t^n \int_{\mathbb{R}^+} M_1 v \tilde{N}(ds, dv) \right) \right) \right]$$

$$= \exp \left( (i+1) \alpha_0(n-t) + \frac{(i+1)^2 \kappa^2}{2} (n-t) \right) \mathbb{E}_t^{Q^0}\left[\exp \left( \int_t^n \int_{\mathbb{R}^+} -(iu+1) M_1 v \tilde{N}(ds, dv) \right) \right]$$

$$= \exp \left( (n-t) \left( (i+1) \alpha_0 + \frac{(i+1)^2 \kappa^2}{2} + \int_{\mathbb{R}^+} (e^{-(iu+1)M_1 v} - 1 + (iu+1) M_1 v) v(dv) \right) \right)$$

$$= \exp \left( (n-t) \left( (i+1) \alpha_0 + \frac{(i+1)^2 \kappa^2}{2} + \int_{\mathbb{R}^+} (e^{-(iu+1)M_1 v} - 1 + (iu+1) M_1 v) \lambda p \eta_1 e^{-\eta_1 v} dv \right) \right)$$

$$= \exp \left( (n-t) \left( (i+1) \alpha_0 + \frac{(i+1)^2 \kappa^2}{2} + \frac{1}{(iu + M_1 + \eta_1)} - \frac{1}{\eta_1} + \frac{(iu + 1) M_1}{\eta_1^2} \lambda p \eta_1 \right) \right). \tag{B.19}$$
## C Tables and figures

Table 1: Characteristics of the Lévy measure under multiple priors.

This table lists the characteristics of the Lévy measure under the reference measure $Q^0$ and alternative equivalent priors $Q^\theta$ for $\theta_{N,1}(t) \in [-M_1,0]$ and $\theta_{N,2}(t) \in [0,M_2]$.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>$Q^0$</th>
<th>$Q^\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump intensity</td>
<td>$\lambda$</td>
<td>$\lambda^\theta \in \left[ \lambda \left( \frac{pn_1}{\eta_1 + M_1} + \frac{qn_2}{\eta_2 + M_2} \right), \lambda \right]$</td>
</tr>
<tr>
<td>Conditional probability of positive jumps</td>
<td>$p^\theta \in \left[ \frac{pn_1}{p(n_1 + q(n_1 + M_1)), \frac{q(n_1 + M_1)}{q(n_1 + M_1)}} \right]$</td>
<td></td>
</tr>
<tr>
<td>Conditional probability of negative jumps</td>
<td>$q = 1 - p$</td>
<td>$p^\theta \in \left[ \frac{qn_2}{q(n_2 + M_2) + q(n_1 + M_1)}, \frac{q(n_1 + M_1)}{q(n_1 + M_1)} \right]$</td>
</tr>
<tr>
<td>Conditional mean log positive jump size</td>
<td>$\frac{1}{\eta_1}$</td>
<td>$\int_{\mathbb{R}^+} u f_1^\theta (du) \in \left[ \frac{1}{\eta_1 + M_1}, \frac{1}{\eta_1} \right]$</td>
</tr>
<tr>
<td>Conditional mean log negative jump size</td>
<td>$\frac{-1}{\eta_2}$</td>
<td>$\int_{\mathbb{R}^-} u f_1^\theta (du) \in \left[ \frac{-1}{\eta_2}, \frac{-1}{\eta_2 + M_2} \right]$</td>
</tr>
</tbody>
</table>
Table 2: **Benchmark parameter values.**

This table shows the benchmark parameter values for our numerical analyses and the central moments of \( Y(t) = \ln(X(t)) \) under these values. \( \alpha_L(\alpha_R) \) is the liquidation (renegotiation) cost. \( \theta \) is equity holders’ bargaining power in default. \( \xi \) is the fraction of the diverted free cash flow. \( \delta \) is the entrepreneur's ownership. \( \phi \) is the tax rate. \( r \) is the risk free rate. \( I \) is the cost of the project. \( x_0 \) is the initial EBIT value. The other process parameters can refer to Section 2.1.

A. Parameter values

| Parameter | Exploration | | | Exploitation | | | | |
|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| \( \alpha_L \) | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| \( \alpha_R \) | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| \( \theta \) | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| \( \xi \) | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| \( \delta \) | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| \( \phi \) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| \( r \) | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| \( I \) | 100 | 100 | 100 | 100 | 100 | 100 |
| \( x_0 \) | 3 | 3 | 3 | 3 | 3 | 3 |
| \( \mu \) | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| \( \sigma \) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| \( \lambda \) | 1/5 | 1/5 | 1/5 | 1 | 1 | 1 |
| \( \eta_1 \) | 3 | 3 | 3 | 8 | 8 | 8 |
| \( \eta_2 \) | 3 | 3 | 3 | 8 | 8 | 8 |
| \( p \) | 0.5 | 0.4 | 0.6 | 0.5 | 0.3 | 0.7 |

B. Central moments of \( Y(t) = \ln(X(t)) \)

| Moment | Exploration | | | Exploitation | | | | |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Mean   | -0.025 | -0.023 | -0.027 | -0.016 | -0.015 | -0.017 |
| Variance | 0.084 | 0.084 | 0.084 | 0.071 | 0.071 | 0.071 |
| Skewness | 0.000 | -0.362 | 0.362 | 0.000 | -0.246 | 0.246 |
### Table 3: Comparative statics: the effect of ex ante skewness

This table shows the related quantities under the worst case measure, where $Y(t) = \ln(X(t))$ under the reference measure has a skewness of negative ($-$ skew), zero (0 skew), and positive ($+$ skew). The optimal total relative entropy growth bound $h^*$ is chosen such that the detection error probability is 10% with $n = 120$ and $h_W/h = 0.5$. All the other quantities are evaluated under $h^*$. The central moments are for $Y(t)$. The process parameters can refer to Section 2.1. The specifications for exploration/exploitation and ex ante skewness are in Table 2. Here $X^t_0 (X^*_0)$, $AD^t_e (AD^*_e)$, $V_e(0)$ ($V_e(0)$) denote the investment boundary, the AD price for investment, and the investment value under equity (optimal) financing. $AD_D$ denotes the AD price for default. Leverage is $D/(D + E)$.

<table>
<thead>
<tr>
<th></th>
<th>Exploration</th>
<th></th>
<th></th>
<th>Exploitation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$-$ skew</td>
<td>0 skew</td>
<td>$+$ skew</td>
<td>$-$ skew</td>
<td>0 skew</td>
<td>$+$ skew</td>
</tr>
<tr>
<td>$h^*$</td>
<td>0.028</td>
<td>0.029</td>
<td>0.029</td>
<td>0.031</td>
<td>0.031</td>
<td>0.032</td>
</tr>
<tr>
<td>$\kappa^*$</td>
<td>0.168</td>
<td>0.170</td>
<td>0.171</td>
<td>0.176</td>
<td>0.177</td>
<td>0.178</td>
</tr>
<tr>
<td>$M_1^*$</td>
<td>2.173</td>
<td>1.834</td>
<td>1.607</td>
<td>2.346</td>
<td>1.726</td>
<td>1.421</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>-0.042</td>
<td>-0.048</td>
<td>-0.053</td>
<td>-0.033</td>
<td>-0.040</td>
<td>-0.045</td>
</tr>
<tr>
<td>$\lambda^*$</td>
<td>0.166</td>
<td>0.162</td>
<td>0.158</td>
<td>0.932</td>
<td>0.911</td>
<td>0.894</td>
</tr>
<tr>
<td>$p^*$</td>
<td>0.279</td>
<td>0.383</td>
<td>0.494</td>
<td>0.249</td>
<td>0.451</td>
<td>0.665</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.075</td>
<td>-0.079</td>
<td>-0.084</td>
<td>-0.065</td>
<td>-0.072</td>
<td>-0.077</td>
</tr>
<tr>
<td>Variance</td>
<td>0.070</td>
<td>0.068</td>
<td>0.065</td>
<td>0.066</td>
<td>0.064</td>
<td>0.063</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.328</td>
<td>-1.078</td>
<td>-0.781</td>
<td>-0.408</td>
<td>-0.195</td>
<td>0.048</td>
</tr>
<tr>
<td>$X^t_0$ ($)</td>
<td>16.697</td>
<td>17.272</td>
<td>17.838</td>
<td>16.190</td>
<td>16.885</td>
<td>17.511</td>
</tr>
<tr>
<td>$X^*_0$ ($)</td>
<td>15.546</td>
<td>16.094</td>
<td>16.633</td>
<td>15.077</td>
<td>15.737</td>
<td>16.331</td>
</tr>
<tr>
<td>$AD^t_e$ ($)</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.010</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>$AD^*_e$ ($)</td>
<td>0.006</td>
<td>0.005</td>
<td>0.004</td>
<td>0.012</td>
<td>0.009</td>
<td>0.007</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.132</td>
<td>0.107</td>
<td>0.096</td>
<td>0.309</td>
<td>0.206</td>
<td>0.159</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.164</td>
<td>0.132</td>
<td>0.118</td>
<td>0.374</td>
<td>0.251</td>
<td>0.194</td>
</tr>
<tr>
<td>$X^<em>_D/X(\tau^</em>_I)$</td>
<td>0.433</td>
<td>0.423</td>
<td>0.415</td>
<td>0.422</td>
<td>0.413</td>
<td>0.405</td>
</tr>
<tr>
<td>$AD_D$</td>
<td>0.609</td>
<td>0.618</td>
<td>0.625</td>
<td>0.589</td>
<td>0.601</td>
<td>0.611</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.572</td>
<td>0.568</td>
<td>0.566</td>
<td>0.565</td>
<td>0.563</td>
<td>0.561</td>
</tr>
</tbody>
</table>
Table 4: **Comparative statics: the effect of sample length**

This table shows the related quantities under the worst case measure, where the calibrations of the detection error probability are based on 15, 30, and 60 years of quarterly data, i.e., \( n = 60, 120, 240 \). The optimal total relative entropy growth bound \( h^* \) is chosen such that the detection error probability is 10% with \( h_W / h = 0.5 \). All the other quantities are evaluated under \( h^* \). The central moments are for \( Y(t) \). The process parameters can refer to Section 2.1. The specifications for exploration/exploitation are in Table 2, and \( Y(t) \) is symmetric in this table. Here \( X^e_I (X^*_I) \), \( AD^e_I (AD^*_I) \), \( V_e(0) (V_* (0)) \) denote the investment boundary, the AD price for investment, and the investment value under equity (optimal) financing. \( AD_D \) denotes the AD price for default. Leverage is \( D/(D+E) \).

<table>
<thead>
<tr>
<th>( h^* )</th>
<th>Exploration</th>
<th>Explloitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 60 )</td>
<td>0.054</td>
<td>0.061</td>
</tr>
<tr>
<td>( n = 120 )</td>
<td>0.029</td>
<td>0.031</td>
</tr>
<tr>
<td>( n = 240 )</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>( \kappa^* )</td>
<td>0.232</td>
<td>0.247</td>
</tr>
<tr>
<td>( \kappa^* )</td>
<td>0.170</td>
<td>0.177</td>
</tr>
<tr>
<td>( \kappa^* )</td>
<td>0.123</td>
<td>0.127</td>
</tr>
<tr>
<td>( M_1^* )</td>
<td>3.229</td>
<td>2.628</td>
</tr>
<tr>
<td>( \eta_1^* )</td>
<td>6.229</td>
<td>10.628</td>
</tr>
<tr>
<td>( \mu^* )</td>
<td>-0.067</td>
<td>-0.062</td>
</tr>
<tr>
<td>( \lambda^* )</td>
<td>0.148</td>
<td>0.876</td>
</tr>
<tr>
<td>( p^* )</td>
<td>0.325</td>
<td>0.429</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.097</td>
<td>-0.092</td>
</tr>
<tr>
<td>Variance</td>
<td>0.065</td>
<td>0.062</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.278</td>
<td>-0.256</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.260</td>
<td>3.938</td>
</tr>
<tr>
<td>( X^e_I ($) )</td>
<td>19.051</td>
<td>19.011</td>
</tr>
<tr>
<td>( X^*_I ($) )</td>
<td>17.785</td>
<td>17.759</td>
</tr>
<tr>
<td>( AD^e_I ($) )</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>( AD^*_I ($) )</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>( V_e(0) ($) )</td>
<td>0.011</td>
<td>0.029</td>
</tr>
<tr>
<td>( V_* (0) ($) )</td>
<td>0.014</td>
<td>0.037</td>
</tr>
<tr>
<td>( V_e / V_e - 1 )</td>
<td>30.767</td>
<td>27.007</td>
</tr>
<tr>
<td>( X^<em>_D / X(\tau_I^</em>) )</td>
<td>0.410</td>
<td>0.399</td>
</tr>
<tr>
<td>( AD_D )</td>
<td>0.650</td>
<td>0.643</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.565</td>
<td>0.559</td>
</tr>
</tbody>
</table>
Table 5: Comparative statics: the effect of diffusion ambiguity share

This table shows the related quantities under the worst case measure, where the calibrations of the detection error probability are based diffusion ambiguity shares $h_W/h = 0.1, 0.5, 0.9$. The optimal total relative entropy growth bound $h^*$ is chosen such that the detection error probability is 10% with $n = 120$. All the other quantities are evaluated under $h^*$. The central moments are for $Y(t)$. The process parameters can refer to Section 2.1. The specifications for exploration/exploitation are in Table 2, and $Y(t)$ is symmetric in this table. Here $X^* (X^*_I), AD^*_I (AD^*_I), V_e(0) (V_e(0))$ denote the investment boundary, the AD price for investment, and the investment value under equity (optimal) financing. $AD_D$ denotes the AD price for default. Leverage is $D/(D+E)$.

<table>
<thead>
<tr>
<th>Exploration</th>
<th>Exploitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h^*$</td>
<td>$h^*$</td>
</tr>
<tr>
<td>$h_W/h = 0.1$</td>
<td>$h_W/h = 0.5$</td>
</tr>
<tr>
<td>0.022</td>
<td>0.072</td>
</tr>
<tr>
<td>0.029</td>
<td>0.260</td>
</tr>
<tr>
<td>0.038</td>
<td>0.723</td>
</tr>
<tr>
<td>$\eta^*_I$</td>
<td>$\eta^*_I$</td>
</tr>
<tr>
<td>2.423</td>
<td>1.834</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>$\mu^*$</td>
</tr>
<tr>
<td>-0.031</td>
<td>-0.048</td>
</tr>
<tr>
<td>0.155</td>
<td>0.162</td>
</tr>
<tr>
<td>0.356</td>
<td>0.383</td>
</tr>
<tr>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>-0.061</td>
<td>-0.079</td>
</tr>
<tr>
<td>Variance</td>
<td>Variance</td>
</tr>
<tr>
<td>0.066</td>
<td>0.068</td>
</tr>
<tr>
<td>Skewness</td>
<td>Skewness</td>
</tr>
<tr>
<td>-1.188</td>
<td>-1.078</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>10.158</td>
<td>10.095</td>
</tr>
<tr>
<td>15.428</td>
<td>17.272</td>
</tr>
<tr>
<td>$X^*_I ($</td>
<td>$X^*_I ($</td>
</tr>
<tr>
<td>$AD^*_I ($</td>
<td>$AD^*_I ($</td>
</tr>
<tr>
<td>0.011</td>
<td>0.004</td>
</tr>
<tr>
<td>0.013</td>
<td>0.005</td>
</tr>
<tr>
<td>$V_e(0) ($</td>
<td>$V_e(0) ($</td>
</tr>
<tr>
<td>0.318</td>
<td>0.107</td>
</tr>
<tr>
<td>0.387</td>
<td>0.132</td>
</tr>
<tr>
<td>$V_e/V_e - 1$ (%)</td>
<td>$V_e/V_e - 1$ (%)</td>
</tr>
<tr>
<td>$X^<em>_D/X^</em>_I$</td>
<td>$X^<em>_D/X^</em>_I$</td>
</tr>
<tr>
<td>0.442</td>
<td>0.423</td>
</tr>
<tr>
<td>0.577</td>
<td>0.618</td>
</tr>
<tr>
<td>Leverage</td>
<td>Leverage</td>
</tr>
<tr>
<td>0.574</td>
<td>0.568</td>
</tr>
</tbody>
</table>
Table 6: **Comparative statics: the effect of diffusion ambiguity share conditional on $\sigma$**

This table shows the related quantities under the worst case measure conditional on the diffusive volatility parameter $\sigma$. The calibrations of the detection error probability are based diffusion ambiguity shares $h_W/h = 0.1, 0.5, 0.9$. The optimal total relative entropy growth bound $h^*$ is chosen such that the detection error probability is 10% with $n = 120$. All the other quantities are evaluated under $h^*$. The central moments are for $Y(t)$. The process parameters can refer to Section 2.1. The specifications for exploration/exploitation are in Table 2, and $Y(t)$ is symmetric in this table. Here $X^*_1$, $AD^*_I$, $V_e(0)$ ($V_*(0)$) denote the investment boundary, the AD price for investment, and the investment value under equity (optimal) financing. $AD_D$ denotes the AD price for default. Leverage is $D/(D + E)$.

Panel A: $\sigma = 0.1$

<table>
<thead>
<tr>
<th>Exploration</th>
<th></th>
<th>Exploitation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$h^*$</td>
<td>$h_W/h = 0.1$</td>
<td>$h_W/h = 0.5$</td>
<td>$h_W/h = 0.9$</td>
</tr>
<tr>
<td>$\kappa^*$</td>
<td>0.022</td>
<td>0.029</td>
<td>0.038</td>
</tr>
<tr>
<td>$M^*_1$</td>
<td>2.423</td>
<td>1.834</td>
<td>0.723</td>
</tr>
<tr>
<td>$\eta^*_1$</td>
<td>5.423</td>
<td>4.834</td>
<td>3.723</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>-0.024</td>
<td>-0.031</td>
<td>-0.026</td>
</tr>
<tr>
<td>$\lambda^*$</td>
<td>0.155</td>
<td>0.162</td>
<td>0.181</td>
</tr>
<tr>
<td>$p^*$</td>
<td>0.356</td>
<td>0.383</td>
<td>0.446</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.040</td>
<td>-0.047</td>
<td>-0.048</td>
</tr>
<tr>
<td>Variance</td>
<td>0.036</td>
<td>0.038</td>
<td>0.044</td>
</tr>
<tr>
<td>Skewness</td>
<td>-2.950</td>
<td>-2.603</td>
<td>-1.399</td>
</tr>
<tr>
<td>$X^*_1$ ($)</td>
<td>11.775</td>
<td>12.575</td>
<td>12.851</td>
</tr>
<tr>
<td>$X^*_I$ ($)</td>
<td>10.849</td>
<td>11.615</td>
<td>11.878</td>
</tr>
<tr>
<td>$AD^*_I$ ($)</td>
<td>0.005</td>
<td>0.004</td>
<td>0.013</td>
</tr>
<tr>
<td>$AD^*_D$ ($)</td>
<td>0.006</td>
<td>0.005</td>
<td>0.015</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.098</td>
<td>0.086</td>
<td>0.447</td>
</tr>
<tr>
<td>$V_*(0)$ ($)</td>
<td>0.130</td>
<td>0.112</td>
<td>0.541</td>
</tr>
<tr>
<td>$V_e/V_*(0)$</td>
<td>33.194</td>
<td>30.246</td>
<td>21.151</td>
</tr>
<tr>
<td>$X^<em>_I$ / $X$ ($\tau^</em>_I$)</td>
<td>0.526</td>
<td>0.505</td>
<td>0.496</td>
</tr>
<tr>
<td>$AD_D$</td>
<td>0.486</td>
<td>0.512</td>
<td>0.518</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.601</td>
<td>0.594</td>
<td>0.591</td>
</tr>
</tbody>
</table>

(To be continued next page.)
(Table 6 continued.)

Panel B: $\sigma = 0.3$

<table>
<thead>
<tr>
<th></th>
<th>Exploration</th>
<th></th>
<th>Exploitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h_\nu/h_\nu = 0.1$</td>
<td>$h_\nu/h_\nu = 0.5$</td>
<td>$h_\nu/h_\nu = 0.9$</td>
</tr>
<tr>
<td>$h^*$</td>
<td>0.022</td>
<td>0.029</td>
<td>0.038</td>
</tr>
<tr>
<td>$\kappa^*$</td>
<td>0.067</td>
<td>0.170</td>
<td>0.260</td>
</tr>
<tr>
<td>$M_1^*$</td>
<td>2.423</td>
<td>1.834</td>
<td>0.723</td>
</tr>
<tr>
<td>$\eta_1^*$</td>
<td>5.423</td>
<td>4.834</td>
<td>3.723</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>-0.037</td>
<td>-0.065</td>
<td>-0.079</td>
</tr>
<tr>
<td>$\lambda^*$</td>
<td>0.155</td>
<td>0.162</td>
<td>0.181</td>
</tr>
<tr>
<td>$p^*$</td>
<td>0.356</td>
<td>0.383</td>
<td>0.446</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.093</td>
<td>-0.121</td>
<td>-0.140</td>
</tr>
<tr>
<td>Variance</td>
<td>0.116</td>
<td>0.118</td>
<td>0.124</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.510</td>
<td>-0.470</td>
<td>-0.295</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.317</td>
<td>5.342</td>
<td>5.588</td>
</tr>
<tr>
<td>$X_1^e$ ($)</td>
<td>20.116</td>
<td>23.086</td>
<td>25.188</td>
</tr>
<tr>
<td>$AD_1^e$ ($)</td>
<td>0.017</td>
<td>0.006</td>
<td>0.004</td>
</tr>
<tr>
<td>$AD_1^*$ ($)</td>
<td>0.019</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.746</td>
<td>0.187</td>
<td>0.161</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.862</td>
<td>0.221</td>
<td>0.188</td>
</tr>
<tr>
<td>$V_e/V_e - 1$ (%)</td>
<td>15.541</td>
<td>17.818</td>
<td>16.634</td>
</tr>
<tr>
<td>$X_1^<em>/X(\tau_1^</em>)$</td>
<td>0.398</td>
<td>0.384</td>
<td>0.377</td>
</tr>
<tr>
<td>$AD_D$</td>
<td>0.663</td>
<td>0.704</td>
<td>0.727</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.561</td>
<td>0.557</td>
<td>0.555</td>
</tr>
</tbody>
</table>
Table 7: **Comparative statics: the effect of agency cost**

This table shows the related quantities under the worst case measure and different agency cost specifications. For each innovation case, we examine $\delta = 0.75$, and $\xi = 0.05$, while keeping the other at the benchmark level, and $\delta = 0.75$ and $\xi = 0.005$. The calibrations of the detection error probability are based on the diffusion ambiguity shares $h_W/h = 0.5$ and $n = 120$. The optimal total relative entropy growth bound $h^*$ is chosen such that the detection error probability is 10%. All the other quantities are evaluated under $h^*$. The central moments are for $Y(t)$. The process parameters can refer to Section 2.1. The specifications for exploration/exploitation are in Table 2, and $Y(t)$ is symmetric in this table. Here $X^e_I (X^*_I)$, $AD^e_I (AD^*_I)$, $V_e(0)$ ($V_*(0)$) denote the investment boundary, the AD price for investment, and the investment value under equity (optimal) financing. $AD_D$ denotes the AD price for default. Leverage is $D/(D + E)$.

<table>
<thead>
<tr>
<th>Exploration</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>$\xi$</td>
<td>$(\delta, \xi)$</td>
<td></td>
</tr>
<tr>
<td>$h^*$</td>
<td>0.029</td>
<td>0.029</td>
<td>0.029</td>
</tr>
<tr>
<td>$\kappa^*$</td>
<td>0.170</td>
<td>0.170</td>
<td>0.170</td>
</tr>
<tr>
<td>$M^*_I$</td>
<td>1.834</td>
<td>1.834</td>
<td>1.834</td>
</tr>
<tr>
<td>$\eta^*_I$</td>
<td>4.834</td>
<td>4.834</td>
<td>4.834</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>-0.048</td>
<td>-0.048</td>
<td>-0.048</td>
</tr>
<tr>
<td>$\lambda^*$</td>
<td>0.162</td>
<td>0.162</td>
<td>0.162</td>
</tr>
<tr>
<td>$p^*$</td>
<td>0.383</td>
<td>0.383</td>
<td>0.383</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.079</td>
<td>-0.079</td>
<td>-0.079</td>
</tr>
<tr>
<td>Variance</td>
<td>0.068</td>
<td>0.068</td>
<td>0.068</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.078</td>
<td>-1.078</td>
<td>-1.078</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.095</td>
<td>10.095</td>
<td>10.095</td>
</tr>
<tr>
<td>$X^e_I$ ($)</td>
<td>17.380</td>
<td>16.639</td>
<td>17.272</td>
</tr>
<tr>
<td>$X^*_I$ ($)</td>
<td>16.137</td>
<td>15.899</td>
<td>16.094</td>
</tr>
<tr>
<td>$AD^e_I$ ($)</td>
<td>0.004</td>
<td>0.005</td>
<td>0.004</td>
</tr>
<tr>
<td>$AD^*_I$ ($)</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.154</td>
<td>0.120</td>
<td>0.107</td>
</tr>
<tr>
<td>$V_*(0)$ ($)</td>
<td>0.193</td>
<td>0.137</td>
<td>0.132</td>
</tr>
<tr>
<td>$V_e/V_e - 1$ (%)</td>
<td>24.640</td>
<td>14.453</td>
<td>23.339</td>
</tr>
<tr>
<td>$X^<em>_D/X^</em>(\tau^*_I)$</td>
<td>0.431</td>
<td>0.334</td>
<td>0.423</td>
</tr>
<tr>
<td>$AD_D$</td>
<td>0.624</td>
<td>0.547</td>
<td>0.618</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.574</td>
<td>0.502</td>
<td>0.568</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exploitation</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>$\xi$</td>
<td>$(\delta, \xi)$</td>
<td></td>
</tr>
<tr>
<td>$h^*$</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
</tr>
<tr>
<td>$\kappa^*$</td>
<td>0.177</td>
<td>0.177</td>
<td>0.177</td>
</tr>
<tr>
<td>$M^*_I$</td>
<td>1.726</td>
<td>1.726</td>
<td>1.726</td>
</tr>
<tr>
<td>$\eta^*_I$</td>
<td>9.726</td>
<td>9.726</td>
<td>9.726</td>
</tr>
<tr>
<td>$\mu^*$</td>
<td>-0.040</td>
<td>-0.040</td>
<td>-0.040</td>
</tr>
<tr>
<td>$\lambda^*$</td>
<td>0.911</td>
<td>0.911</td>
<td>0.911</td>
</tr>
<tr>
<td>$p^*$</td>
<td>0.451</td>
<td>0.451</td>
<td>0.451</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.072</td>
<td>-0.072</td>
<td>-0.072</td>
</tr>
<tr>
<td>Variance</td>
<td>0.064</td>
<td>0.064</td>
<td>0.064</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.195</td>
<td>-0.195</td>
<td>-0.195</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.975</td>
<td>3.975</td>
<td>3.975</td>
</tr>
<tr>
<td>$X^e_I$ ($)</td>
<td>16.991</td>
<td>16.266</td>
<td>16.885</td>
</tr>
<tr>
<td>$X^*_I$ ($)</td>
<td>15.780</td>
<td>15.541</td>
<td>15.737</td>
</tr>
<tr>
<td>$AD^e_I$ ($)</td>
<td>0.007</td>
<td>0.008</td>
<td>0.007</td>
</tr>
<tr>
<td>$AD^*_I$ ($)</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>$V_e(0)$ ($)</td>
<td>0.297</td>
<td>0.228</td>
<td>0.206</td>
</tr>
<tr>
<td>$V_*(0)$ ($)</td>
<td>0.366</td>
<td>0.260</td>
<td>0.251</td>
</tr>
<tr>
<td>$V_e/V_e - 1$ (%)</td>
<td>23.029</td>
<td>13.633</td>
<td>21.827</td>
</tr>
<tr>
<td>$X^<em>_D/X^</em>(\tau^*_I)$</td>
<td>0.421</td>
<td>0.326</td>
<td>0.413</td>
</tr>
<tr>
<td>$AD_D$</td>
<td>0.607</td>
<td>0.527</td>
<td>0.601</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.569</td>
<td>0.497</td>
<td>0.563</td>
</tr>
</tbody>
</table>
Figure 1: **Calibration of detection error probabilities.** This figure plots calibration results for detection error probabilities $\pi(n; h)$, and the parameter values. Here $n = 120$ (quarters). $h$ denotes the total relative entropy growth bound. We set $h_W / h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 2: **Central moments under the worst-case measure.** This figure plots the first four central moments of $Y(t)$ based on the calibrated parameters plotted in Figure 1. $h$ denotes the total relative entropy growth bound. The calibration of $h^*$ uses $n = 120$ (quarters). We set $h_W/h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 3: **Jump size distributions under the worst-case measure.** This figure plots the probability density function for the jump size of $Y(t)$ under the reference measure $Q^0$ and under the worst-case measure $Q^0\theta^{*}$ where total ambiguity $h = h^{*}$, the maximum level. We set $h_W / h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 4: **The effects of ambiguity on growth option values.** In Panels (a) to (c), we plot the project value $V_i(0)$, the optimal investment boundary $X^*_i$, and the AD price of investment $AD^i$ against total ambiguity $h$ for $h \in [0, h^*]$, where $i$ denotes equity financing ($e$) and optimal financing ($\ast$). In Panels (d) to (f), we plot the corresponding relative differences. The calibration of $h^*$ uses $n = 120$ (quarters). We set $h_W/h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 4: **The effects of ambiguity on growth option values (Continued).** In Panels (a) to (c), we plot the project value $V_i(0)$, the optimal investment boundary $X^*_i$, and the AD price of investment $AD^*_i$ against total ambiguity $h$ for $h \in [0, h^*]$, where $i$ denotes equity financing ($e$) and optimal financing ($\ast$). In Panels (d) to (f), we plot the corresponding relative differences. The calibration of $h^*$ uses $n = 120$ (quarters). We set $h_W/h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 5: The effects of ambiguity on capital structure. This figure plots the optimal leverage $D/D + E$, the scaled optimal default boundary $X_D^*/X(\tau_I^*)$, the AD price of default, the optimal coupon $C^*$, the debt value, and equity value against total ambiguity $h$ for $h \in [0, h^*]$. In Panel (d) to (e), we plot the quantities at the optimal investment boundary $X_I^*$. The calibration of $h^*$ uses $n = 120$ (quarters). We set $h_W/h = 0.5$. We choose baseline Exploration and Exploitation parameters, Column (1) in Table 2 for the corresponding case.
Figure 6: **The effects of ambiguity and skewness on growth option values.** In Panels (a) to (c), we plot the gain in project value due to debt financing $V_*(0)/V_e(0) - 1$ (%), the project value under equity financing $V_e(0)$, and the project value under optimal financing $V_*(0)$ against total ambiguity $h$ for $h \in [0, h^*]$ for exploration. In Panels (d) to (f), we plot the same for exploitation. “0–”, “−”, and “+” Skew corresponds to the parameter values in Columns (1) to (3) in Table 2 for each innovation type. The calibration of $h^*$ is based on $h_W/h = 0.5$ and $n = 120$ (quarters).
Figure 6: The effects of ambiguity and skewness on growth option values (Continued). In Panels (a) to (c), we plot the gain in project value due to debt financing \( \frac{V_*(0)}{V_e(0)} - 1 \) (%), the project value under equity financing \( V_e(0) \), and the project value under optimal financing \( V_*(0) \) against total ambiguity \( h \) for \( h \in [0, h^*] \) for exploration. In Panels (d) to (f), we plot the same for exploitation. “0−”, “−”, and “+” Skew corresponds to the parameter values in Columns (1) to (3) in Table 2 for each innovation type. The calibration of \( h^* \) is based on \( h_W/h = 0.5 \) and \( n = 120 \) (quarters).