Submitted to *Operations Research*

manuscript OPRE-2016-08-453

Fear of the Market or Fear of the Competitor? Ambiguity in a Real Options Game

Tobias Hellmann

Norddeutsche Landesbank, Hannover, Germany, Tobias.hellmann@nordlb.de

Jacco J.J. Thijssen

The York Management School, University of York, United Kingdom, jacco.thijssen@york.ac.uk

In this paper we study an investment game between two firms with a first–mover advantage, where payoffs are driven by a geometric Brownian motion. At least one of the firms is assumed to be ambiguous over the drift, with maxmin preferences over a strongly rectangular set of priors. We develop a strategy and equilibrium concept allowing for ambiguity and show that equilibria can be preemptive (a firm invests at a point where investment is Pareto dominated by waiting) or sequential (one firm invests as if it were the exogenously appointed leader). Following the standard literature, the worst–case prior for an ambiguous firm in the follower role is obtained by setting the lowest possible trend in the set of priors. However, if an ambiguous firm is the first mover, then the worst–case drift can fluctuate between the lowest and the highest trends. This novel result shows that "worst–case prior" in a setting with drift ambiguity does not always equate to "lowest trend". As a consequence, preemptive pressure reduces. We show that this results in the possibility of firm value being increasing in the level of ambiguity. If only one firm is ambiguous, then the value of the non–ambiguous firm can be increasing in the level of ambiguity of the ambiguous firm.

Key words: Real Options, Knightian Uncertainty, Worst–Case Prior, Optimal Stopping, Timing Game

---

1. Introduction

Many, if not most, investment decisions taken by firms are characterized by substantial upfront sunk costs, (partial) irreversibility, and uncertainty over future cash flows (cf. Dixit and Pindyck (1994)). As has been well–recognized since Knight (1921), the uncertainty over future cash flows can seldom be captured by a unique probability measure; there is usually *ambiguity* over the correct probability measure. For example, when firms are contemplating entering a new market, or introducing a new technology, there is typically no historical data available from which the distribution of future demand can be inferred. Instead, managers have to rely on market research, expert opinion, etc., to gauge likely future demand for their products. This variety of opinion results in a set of probability distributions over future events, which may not easily be condensed into a single distribution. Therefore, managers need a theory of valuation of investment projects under ambiguity. In this paper paper we provide such a theory for the case where competing firms have a similar investment option, thereby adding a strategic dimension to the ambiguous business environment.

The effect of ambiguity on decision making has been studied extensively, most famously by Ellsberg (1961). The overwhelming conclusion of the experimental literature is that decision makers are *ambiguity averse*. The classical Ellsberg experiment uses two urns, both with 100 red or blue balls. For the first urn it is known that half the balls are red. For the second urn no such information is available. Most people prefer bets on the first urn with comparable bets on the second urn. This observation violates Savage's "sure thing principle"; in fact, it calls into doubt "a basic tenet of Bayesianism, namely, that all uncertainty can be expressed in a probabilistic way. Exhibiting preferences for known versus unknown probabilities is incompatible with this tenet" (Gilboa (2009, p. 134)).

The Ellsberg paradox is not really a paradox, because it does not result from a cognitive bias, but, rather, from lack of information. It is perfectly possible for decision makers to make consistent decisions under ambiguity. This has been shown by Gilboa and Schmeidler (1989), who incorporate an ambiguity aversion axiom into the subjective expected utility framework. They then show that a rational decision maker acts as if she maximizes expected utility over the worst–case prior within a (subjectively chosen) set of priors. Ambiguity models have been successful in the finance literature to explain phenomena such as selective participation, under–diversification, and portfolio inertia (see Epstein and Schneider (2010)).

In the corporate environment, ambiguity may arise in several contexts. For example, a firm's management may be considering introducing a new technology to the market, so that there might not be any relevant historical data that can guide decision making. Instead, management may have to rely on expert opinion. It is well–documented that expert opinion can wildly vary in such cases (see, e.g., Ball and Watt (2013)). It is still possible, of course, for management to act in a Bayesian way and construct a single prior from such varying opinions and there is a literature that provides ways of doing so (see, e.g., Gzyl et al. (2017)). However, it has been suggested in the risk management literature (see, e.g., Randall (2011)) that in cases where risks are not easy to reduce to a single probability measure managers should apply a "precautionary principle". The multiple prior maxmin utility of Gilboa and Schmeidler (1989) can be seen as an operationalization of this idea. It allows the decision maker to pick a "cautious" probability measure from a range of plausible ones in a way that is consistent with experimentally observed behavior.

In this paper, we use multiple–prior maxmin preferences to analyse investment timing decisions by ambiguity–averse decision makers in a duopoly. According to our theory, ambiguity aversion reduces pre-emptive pressure. In addition, firm value under ambiguity aversion can actually be increasing in ambiguity.

In addition, the value of an unambiguous firm can be increasing in the degree of ambiguity of an ambiguity averse firm. Hence, a firm that is not ambiguity averse benefits from its rival's ambiguity aversion.

Our theory stays close to the real options literature (cf. Dixit and Pindyck (1994)), the most important prediction of which (i.e. that higher volatility leads to later investment) is validated empirically; see, e.g., Bloom et al. (2007) and Bloom (2009), and references therein. In addition, a major survey reported in Graham and Harvey (2001) concludes that the real options approach is fairly popular among managers as a tool for investment appraisal.

In the real options literature it is common to assume that future cash flows evolve according to a (continuous–time) stochastic process, where cash flows grow at an expected rate augmented with shocks that follow a random walk. Incorporating ambiguity in such a setting is typically done by assuming that at any time t the expected growth rate is not known, but can take any value in a given set (this is referred to as *drift ambiguity*). The worst case in a monopolistic model is induced by the lowest possible expected growth rate (Nishimura and Ozaki (2007)). So, in the Gilboa and Schmeidler (1989) framework applied to investment timing problems, the presence of drift ambiguity leads the firm to act cautiously: by considering the lowest possible expected growth rate the firm values future cash flows assuming that nature will act malevolently. One could interpret this as "fear of the market".

When firms take investment decisions in a competitive environment, they are not only ambiguous about future cash flows, but also about the timing of their competitors' actions. After all, suppose that a firm has just invested in a new technology to obtain, say, a cost advantage, but that its competitor still has the option to invest as well. It is natural to assume that investment by the competitor lowers the first adopter's cash flows. It is similarly innocuous to assume that the competitor will make its investment decision when it expects the future cash flows to be high enough. This implies that, in expectation, the competitor will invest sooner when the expected growth rate of cash flows is higher. This, in turn, means that the worst case for the first adopter is represented by the earliest possible time, in expectation, that the competitor invests, which is represented by the highest possible expected growth rate. One can think of this as "fear of the competitor".

In this paper we investigate how these two diametrically opposed "fears" balance: what is *the* worst case at any given time when "fear of the market" suggests the lowest possible expected growth rate, but "fear of the competitor" suggests the highest possible expected growth rate? It turns out that we can compute the worst–case prior explicitly: it is *either* the lowest *or* the highest expected growth rate. The regions where each of these worst cases dominates the other can, as we show, be determined exactly.

We make contributions to two strands of literature, the first being the literature on decisions under ambiguity. Contrary to the standard literature, in Section 3 we use an analysis based on backward stochastic differential equations (BSDEs) and g–expectations, as introduced by Peng (1997), to study which of the two "fears" dominates. It turns out that for small values of the stochastic process, the worst case always corresponds to the lowest possible trend, whereas for higher values the highest possible trend *may* represent the worst case, depending on the underlying parameters. As a consequence, if firms are symmetric in terms of the set of priors, but asymmetric in, for example, the sunk costs of investment, then there will be times where firms use different priors to determine firm value. So, the worst–case prior is not only role–dependent (first mover or second mover), but also firm–dependent. As we follow the seminal contribution of Chen and Epstein (2002) this result is obtained under standard modeling assumptions.

Our second contribution is to the literature on real options games. We combine the single–firm ambiguity model of Nishimura and Ozaki (2007) with a standard real options game without ambiguity as analysed in, for example, Pawlina and Kort (2006) and Steg (2015, Section 4.1). This allows for a clear analysis of the consequences of ambiguity on equilibrium investment scenarios and firm value.

In Section 4 we show that equilibria in our game can be of the familiar two types. First, there may be *preemptive equilibria* in which both firms are willing to invest at a time where it is not optimal for at least one firm to do so. This type of equilibrium is familiar from the literature (e.g., Fudenberg and Tirole (1985), Weeds (2002), and Pawlina and Kort (2006)), but we embed a technique recently developed by Riedel and Steg (2017) to rigorously prove existence of this type of equilibrium. In particular, we extend their notion of strategy in such a way that firms are ambiguous about *when* action is taken, but not about *what* those actions are. This extension allows for close comparisons with the existing literature and ensures that our results are due to ambiguity only and not to our modeling of strategies.

In a preemptive equilibrium there are subgames where it is known (a.s.) ex ante which firm is going to invest first. This firm will invest at a point in time where its leader value exceeds its follower value, but where its competitor is indifferent between the two roles. There are also subgames where firms compete instantaneously for the leader role. In such subgames, the intensities with which firms try to “win” the leader role are such that any expected advantage of becoming the leader is competed away; so–called *rent equalization* (cf. Fudenberg and Tirole (1985)). Finally, there are subgames where both firms invest simultaneously.

A second type of equilibrium that can arise is a *sequential equilibrium*, in which one firm invests at the same time it would if it knew that the other firm will not preempt. Each game always has at least an equilibrium of one of these two types, which can not co–exist.

In Section 5 we provide some comparative statics of equilibrium strategies and firm values under various levels of ambiguity. We are particularly interested in the difference between an increase in ambiguity and a decrease in the trend in a model with a unique prior. Unlike the real options models analysed in the literature so far, there are cases in our model where the worst–case prior changes over time. While this does not change the types of equilibria that can be obtained, it does have an impact on the prevailing equilibrium for given parameterizations. We show that the preemption region in our model with symmetric ambiguity is always contained in the preemption region that one would obtain in a model under a unique prior with the lowest possible trend. This makes it more likely, in our model, that a sequential equilibrium emerges, rather than a preemptive one. So, ambiguity can reduce preemptive pressure. As we will show below, this is beneficial to the firm with the lower sunk costs of investment. In fact, we show that, counter–intuitively, this firm's equilibrium value may actually be increasing in ambiguity.

The remainder of the paper is organized as follows. In Section 2 we present our model of a real options game under ambiguity. The value functions of the different roles in this game are derived in Section 3.3, after which these functions are used to determine equilibria in Section 4. A numerical analysis of our model is given in Section 5. In Section 6 we discuss some of our modeling assumptions and possible extensions. We also provide a historic case to illustrate the applicability of the theory presented in this paper.

2. The Model

We follow Pawlina and Kort (2006) in considering two firms that are competing, for example to enter a new market. Uncertainty in the market is modeled on a (reference) probability space $(\Omega, \mathscr{F}, \mathsf{P})$ through a geometric Brownian motion $X = (X_t)_{t\geq 0}$ that solves the stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \tag{1}$$

where $(B_t)_{t\geq 0}$ is a Wiener process, and $\mu \in \mathbb{R}$ and $\sigma > 0$ are fixed parameters such that X admits a unique strong solution. Information is modeled by the P–augmented filtration $(\mathscr{F}_t)_{t\geq 0}$ generated by X. The sunk costs of investment are firm specific and given by $I_i > 0$, $i = 1, 2$.

The payoff streams (or cash flows) are given by processes $(D_{k\ell} X_t)_{t\geq 0}$, where $D_{k\ell}$, $k, \ell = 0, 1$, denotes a scaling factor if the firm's investment status is k ($k = 0$ if the firm has not invested and $k = 1$ if the firm has invested) and the investment status of the competitor is $\ell \in \{0, 1\}$. It is assumed that $D_{10} > D_{11} \geq D_{00} \geq D_{01} \geq 0$, and that there is a first mover advantage, i.e. $D_{10} - D_{00} > D_{11} - D_{01}$. In this paper we consider symmetric revenues, although the values for the $D_{k\ell}$s could easily be made firm–specific at the cost of more cumbersome notation. We provide a discussion of these assumptions in Section 6.

We assume that at least one firm is assumed to be ambiguous about the measure P. Following the recent literature on drift ambiguity in continuous time, we model the priors that firm i, $i = 1, 2$, considers using a set of density generators Θ_i. The resulting set of probability measures that constitutes the firm's set of priors is denoted by $\mathscr{P}^{\Theta_i}$. A process $(\theta_t)_{t\geq 0}$ is a *density generator* if the process $(M_t^\theta)_{t\geq 0}$, with

$$dM_t^\theta = -\theta_t M_t^\theta dB_t, \quad M_0^\theta = 1, \tag{2}$$

is a P–martingale. Such a process $(\theta_t)_{t\geq 0}$ generates a new measure Q via the Radon–Nikodym derivative $d\mathsf{Q}/d\mathsf{P} = M_\infty^\theta$. While we assume that both firms use the same reference prior P, the set of density generators can differ between firms.

Under some regularity conditions (see Chen and Epstein (2002)), the set of density generators is defined as

$$\Theta_i := \left\{ (\theta_t)_{t\geq 0} \,\middle|\, \theta_t(\omega) \in \Theta_{i,t}(\omega), d\mathsf{P}\text{–a.e., all } t \geq 0 \right\},$$

and the resulting set of measures $\mathscr{P}^{\Theta_i}$ is called *strongly rectangular*. For sets of strongly–rectangular priors is has been shown that (Chen and Epstein (2002)):

1. $\mathsf{P} \in \mathscr{P}^{\Theta_i}$;
2. all measures in $\mathscr{P}^{\Theta_i}$ are uniformly absolutely continuous with respect to P and are equivalent to P;
3. for every $\xi \in \mathscr{L}^2(\Omega, \mathscr{F}, \mathsf{P})$, there exists $\mathsf{P}_i^* \in \mathscr{P}^{\Theta_i}$ such that for all $t \geq 0$,

$$\mathsf{E}^{\mathsf{P}_i^*}[\xi|\mathscr{F}_t] = \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}[\xi|\mathscr{F}_t]. \tag{3}$$

The measure P_i^* above is called the *worst–case prior* and is generated by a specific density generator called the *upper–rim generator*, which is denoted by $\theta_i^* = (\theta_{i,t}^*)_{t\geq 0}$. Note that the worst–case prior is payoff and player–specific.

For a given density generator $\theta \in \Theta_i$ it follows immediately from Girsanov's theorem, that under the measure Q generated by θ, the process $(B_t^\theta)_{t\geq 0}$, defined by

$$B_t^\theta = B_t + \int_0^t \theta_s ds,$$

is a Q–Brownian motion and that, under Q, the process X follows the diffusion

$$dX_t = (\mu - \sigma\theta_t)X_t dt + \sigma X_t dB_t^\theta.$$

In the remainder we will assume that $\Theta_{i,t} = [-\kappa_i, \kappa_i]$, for all $t \geq 0$, for some $\kappa_i \geq 0$. Denote $\Delta_i = [\underline{\mu}_i, \overline{\mu}_i] = [\mu - \sigma\kappa_i, \mu + \sigma\kappa_i]$. This form of ambiguity is called κ*–ignorance* (cf. Chen and Epstein (2002)), and its advantage is that Θ_i is strongly rectangular so that the results stated above apply. Note that, if $\kappa_i = 0$, it holds that $\mathsf{P}_i^* = \mathsf{P}$. Note that κ–ignorance does not allow for learning; we will discuss the impact of this assumption in Section 6.

Both firms are assumed to be ambiguity averse in the sense of Gilboa and Schmeidler (1989), i.e., they maximise expected profits over the worst–case prior. Such preferences are typically called *muliptle–prior maxmin*. Gilboa and Schmeidler (1989) axiomatize such preferences in a static model, by relaxing some of the Savage axioms of subjective expected utility. An axiomatization for a dynamic setting (in discrete time) is pursued in Epstein and Schneider (2003). Our set up is, in essence, a continuous–time limit of the Epstein and Schneider (2003) model; see also Chen and Epstein (2002).

In the remainder, a special role is played by the density generator $\theta^{-\kappa_i}$ with $\theta_t^{-\kappa_i} = -\kappa_i$ for all $t \geq 0$. The measure that is generated by $\theta^{-\kappa_i}$ is denoted by $\mathsf{P}^{-\kappa_i}$. Under $\mathsf{P}^{-\kappa_i}$ it immediately follows that X follows a GBM with trend $\underline{\mu}_i$ and volatility σ. In fact, in the literature on κ–ambiguity in real options models it typically holds that $\theta_i^* = \theta^{-\kappa_i}$. We will see that in our model, there are payoff streams for which $\theta_i^* \neq \theta^{-\kappa_i}$.

Finally, the discount rate is assumed to be $r > \max\{\overline{\mu}_1, \overline{\mu}_2\}$ and to apply to both firms.

3. Value Functions of Firm Roles

In a timing game with two firms, each firm can play one of three roles. Firm i becomes the *leader* at time t if it is the first firm to invest at that time. The firm becomes the *follower* at time t if its competitor is the first to invest at that time. Finally, firm i becomes a *simultaneous* investor at time t if both firms invest at that time. In equilibrium, a careful balancing of the payoffs accruing from these roles is required. In this section we therefore derive the value functions of these different roles. It turns out that the derivation of the leader value in particular is of interest in its own right. In the following, we denote firm i's worst–case priors of the leader, follower and simultaneous investment values by $\mathsf{P}^*_{L,i}$, $\mathsf{P}^*_{F,i}$ and $\mathsf{P}^*_{M,i}$, respectively.

3.1. The follower value

Assume that firm j, $j=1,2$, becomes the leader at time $t \geq 0$. Then firm i, $i \neq j$, becomes the follower at that time. Firm i determines its time to invest by maximizing its value. This is achieved by solving the optimal stopping problem

$$F_i(x_t) := \sup_{\tau_i^F \geq t} \inf_{\mathsf{Q} \in \mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}} \left[\int_t^{\tau_i^F} e^{-r(s-t)} D_{01} X_s ds + \int_{\tau_i^F}^{\infty} e^{-r(s-t)} (D_{11} X_s - r I_i) ds \Big| \mathscr{F}_t \right]. \tag{4}$$

The solution to this problem is called the *follower value*.

If the set of priors $\mathscr{P}^{\Theta_i}$ is strongly rectangular, it turns out that problem (4) can be reduced to a standard optimal stopping problem and, hence, can be solved using standard techniques. This reduction is possible due to the following lemma, the proof of which is standard (cf. Nishimura and Ozaki (2007)) and is, thus, omitted.

Lemma 1 *Let $\mathscr{P}^{\Theta_i}$ be strongly–rectangular. Then*

$$F_i(x_t) = \sup_{\tau_i^F \geq t} E^{P^*_{F,i}} \left[\int_t^{\tau_i^F} e^{-r(s-t)} D_{01} X_s ds + \int_{\tau_i^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - e^{-r(\tau_i^F - t)} I_i \Big| \mathscr{F}_t \right], \tag{5}$$

where $P^*_{F,i} = P^{-\kappa_i}$.

Hence, for the follower problem of firm i, the worst case is always induced by the lowest possible drift $\underline{\mu}_i$. This observation makes sense, because the actions of the leader have no influence on the decision of the follower once the leader has invested. The problem therefore reduces to one of a "monopolistic" decision maker. In the language of the introduction, the follower is only exposed to "fear of the market" and not to "fear of the competitor."

Applying standard techniques (see, e.g., Dixit and Pindyck (1994)), we find that the follower value can be expressed as

$$F_i(x_t) = \begin{cases} \frac{D_{01} x_t}{r - \underline{\mu}_i} + \left(\frac{D_{11} - D_{01}}{r - \underline{\mu}_i} x_i^F - I_i \right) \left(\frac{x_t}{x_i^F} \right)^{\beta_1(\underline{\mu}_i)} & \text{if } x_t < x_i^F \\ \frac{D_{11} x_t}{r - \underline{\mu}_i} - I_i & \text{if } x_t \geq x_i^F, \end{cases} \tag{6}$$

where x_i^F is the *investment trigger* of firm i, i.e., the value of the process X, which, once crossed from below, makes investment (rather than waiting) the optimal decision. In this case

$$x_i^F = \frac{\beta_1(\underline{\mu}_i)}{\beta_1(\underline{\mu}_i)-1}\frac{r-\underline{\mu}_i}{D_{11}-D_{01}}I_i,$$

and $\beta_1(\mu)>1$ is the positive root of the quadratic equation

$$\mathscr{Q}(\beta;\mu)\equiv\frac{1}{2}\sigma^2\beta(\beta-1)+\mu\beta-r=0. \tag{7}$$

3.2. The simultaneous investment value

In a similar way one can argue that the upper–rim generator of $\mathscr{P}^{\Theta_i}$ for the value of simultaneous investment, denoted by M_i, satisfies $\theta_i^*=\theta^{-\kappa_i}$. Hence, under the worst–case prior the firm acts *as if* the trend is given by $\underline{\mu}$ and therefore that

$$\begin{aligned} M_i(x_t) &:= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^\infty e^{-r(s-t)}(D_{11}X_s-rI_i)ds\Big|\mathscr{F}_t\right] \\ &= \mathsf{E}^{\mathsf{P}^*_{M,i}}\left[\int_t^\infty e^{-r(s-t)}(D_{11}X_s-rI_i)ds\Big|\mathscr{F}_t\right] = \frac{D_{11}}{r-\underline{\mu}_i}x_t-I_i, \end{aligned}$$

where $\mathsf{P}^*_{M,i}=\mathsf{P}^{-\kappa_i}$.

3.3. The leader value

The standard techniques for computing value functions are not applicable any longer for determining the leader value. In our setting one needs to allow for the worst–case drift to change over time. Indeed, we will see that there are cases where $\mathsf{P}^*_{L,i}\neq\mathsf{P}^{-\kappa_i}$. In such cases there will be times t for which $\theta^*_{1,t}\neq\theta^*_{2,t}$, even if $\mathscr{P}^{\Theta_1}=\mathscr{P}^{\Theta_2}$.

If firm i becomes the leader at time $t\geq 0$ (implying that firm j follows at the stopping time τ_j^F), firm i's *leader value* is given by

$$L_i(x_t) := \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_sds+\int_{\tau_j^F}^\infty e^{-r(s-t)}D_{11}X_sds\Big|\mathscr{F}_t\right]-I_i. \tag{8}$$

In Theorem 1 below, which describes this leader value, two cases are distinguished. If the difference $D_{10}-D_{11}$ is sufficiently small, we find that the worst case is, as before, always induced by $\underline{\mu}_i$. If this condition is not satisfied, then the worst case is given by $\underline{\mu}_i$ for small values x_t up to a certain threshold x_i^*, where it jumps to $\overline{\mu}_i$. The intuition for this fact can already be derived from equation (8): the lowest trend $\underline{\mu}_i$ gives the minimal values for the payoff stream $(D_{kl}X_t)$. However, the higher the trend μ, the sooner the stopping time τ_j^F is expected to be reached. The higher payoff stream $(D_{10}X_t)$ is then sooner replaced by the lower one $(D_{11}X_t)$. If the drop in payoffs becomes sufficiently small, the former effect always dominates the latter. Then the worst case is given by $\underline{\mu}_i$ for each x_t. If the drop in payoffs is large enough, there may be values x_t where the latter effect dominates the former and where the worst–case prior is induced by $\overline{\mu}_i$ instead.

Theorem 1 *The worst case for the leader value function is always given by the lowest possible drift $\underline{\mu}_i$ ($\theta_i^* = \theta^{-\kappa_i}$) if, and only if, it holds that*

$$\frac{D_{10} - D_{11}}{D_{10}} \leq \frac{1}{\beta_1(\underline{\mu}_i)}. \tag{9}$$

In this case, the leader value function becomes

$$L_i(x_t) = \begin{cases} \frac{D_{10}x_t}{r-\underline{\mu}_i} + \left(\frac{x_t}{x_j^F}\right)^{\beta_1(\underline{\mu}_i)} \frac{D_{11}-D_{10}}{r-\underline{\mu}_i} x_j^F - I_i & \text{if } x_t < x_j^F \\ \frac{D_{11}x_t}{r-\underline{\mu}_i} - I_i & \text{if } x_t \geq x_j^F. \end{cases} \tag{10}$$

On the other hand, if $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu}_i)}$, then there exists a unique threshold $x_i^* \in (0, x_j^F)$ such that $\underline{\mu}_i$ is the worst case ($\theta_{i,t}^* = -\kappa_i$) on $\{X_t < x_i^*\}$ and $\overline{\mu}_i$ is the worst case ($\theta_{i,t}^* = +\kappa_i$) on $\{x_i^* \leq X_t < x_j^F\}$. Furthermore, in this case the leader value function is given by

$$L_i(x_t) = \begin{cases} \frac{D_{10}x_t}{r-\underline{\mu}_i} - \frac{1}{\beta_1(\underline{\mu}_i)} \frac{D_{10}x_i^*}{r-\underline{\mu}_i} \left(\frac{x_t}{x_i^*}\right)^{\beta_1(\underline{\mu}_i)} - I_i & \text{if } x_t < x_i^* \\ \frac{D_{10}x_t}{r-\overline{\mu}_i} + \frac{(x_i^*)^{\beta_2(\overline{\mu}_i)} x_t^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)} x_t^{\beta_2(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)} (x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)} (x_j^F)^{\beta_2(\overline{\mu}_i)}} \left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right) x_j^F & \\ + \frac{(x_j^F)^{\beta_1(\overline{\mu}_i)} x_t^{\beta_2(\overline{\mu}_i)} - (x_j^F)^{\beta_2(\overline{\mu}_i)} x_t^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)} (x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)} (x_j^F)^{\beta_2(\overline{\mu}_i)}} \left[\left(1 - \frac{1}{\beta_1(\underline{\mu}_i)}\right) \frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right] x_i^* - I_i & \text{if } x_i^* \leq x_t < x_j^F \\ \frac{D_{11}x_t}{r-\underline{\mu}_i} - I_i & \text{if } x_t \geq x_j^F, \end{cases} \tag{11}$$

where $\beta_2(\mu) < 0$ is the negative root of the quadratic equation $\mathscr{Q}(\beta;\mu) = 0$ (cf. (7)).

Before proving this theorem we point out some of its features and consequences. The value function (10) is standard in real options models, whereas (11) is not. However, the first part of (11) does look like the value function one obtains from a standard real options model. Suppose a (non–ambiguous) firm has an option to exchange the payoff stream $(D_0X_t)_{t\geq 0}$ for $(D_1X_t)_{t\geq 0}$, with $0 \leq D_0 < D_1$, by paying a sunk cost I. It is well known (cf. Dixit and Pindyck (1994)) that the value function in the continuation region $(0, x^*)$ can then be written as

$$V(x_t) = \frac{D_0 x}{r-\mu} + \frac{1}{\beta_1(\mu)-1} I \left(\frac{x_t}{x^*}\right)^{\beta_1(\mu)},$$

where x^* denotes the optimal investment trigger. This value function allows for a clear interpretation. If the firm never invests, then its value is the first term on the right-hand side, which is simply the expected present–value of receiving the stream $(D_0X_t)_{t\geq 0}$ forever. This value has to be corrected for the fact that at x^* the firm's value increases due to investment. At that time, it turns out that the firm's value increases by a factor $I/(\beta_1(\mu)-1)$. The factor $1/(\beta_1(\mu)-1)$ results from applying the value–matching and smooth–pasting optimality conditions at the trigger x^*. That value has to be discounted back to time t, which is achieved by the expected discount factor

$$\mathsf{E}^{\mathsf{P}}\left[e^{-r(\tau^*-t)}\middle|\mathscr{F}_t\right] = \left(\frac{x_t}{x^*}\right)^{\beta_1(\mu)},$$

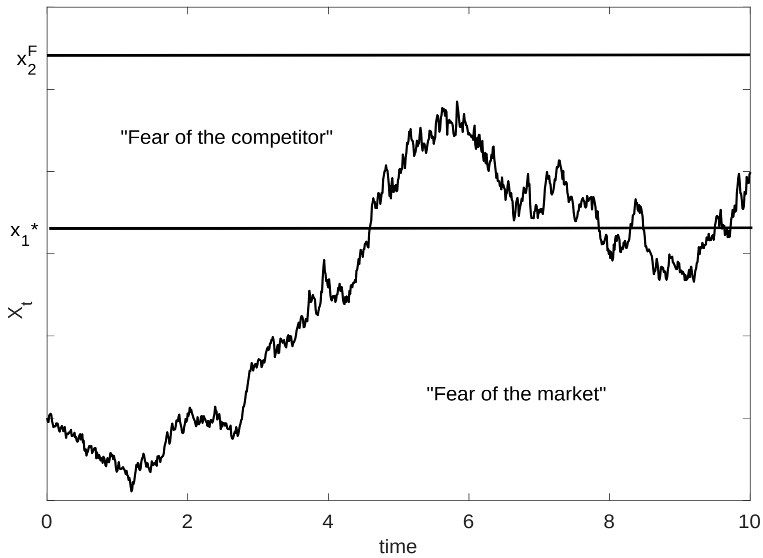


Figure 1 The critical value x_1^* differentiates between two "regimes" for firm 1.

where τ^* is the first hitting time of x^*.

In our model we have a similar interpretation. If the firm never changes the worst case from $\underline{\mu}_i$, then its present value is $D_{10}x/(r-\underline{\mu}_i)$. However, this overestimates the firm's value, because at x_i^* the worst case changes and, thus, evaluating revenues under $\underline{\mu}_i$ forever is too optimistic. At x_i^* it turns out that revenues need to be reduced by a factor $1/\beta_1(\underline{\mu}_i)$. Like in the option case, this factor is obtained by applying a smooth–pasting condition, as becomes obvious from the proof below. The expected discount factor discounts this drop in payoffs back to time t.

It is important to highlight that, while the payoff function looks like the value function of an option exercise decision, our value function does not allow for an option interpretation. The switch at x_i^* is entirely driven by the worst–case prior for the leader value function. It just so happens that smooth pasting results in similar–looking expressions, but there is no *optimality* condition being applied here. We also point out that, while it may appear that the value function for $x < x_i^*$ does not depend on $\overline{\mu}_i$, in fact it does through the value of x_i^*, which depends on both $\underline{\mu}_i$ and $\overline{\mu}_i$.

The second part of the value function (11) contains the terms

$$\frac{(x_i^*)^{\beta_2(\overline{\mu}_i)}x_t^{\beta_1(\overline{\mu}_i)}-(x_i^*)^{\beta_1(\overline{\mu}_i)}x_t^{\beta_2(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)}-(x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}} \quad \text{and} \quad \frac{(x_j^F)^{\beta_1(\overline{\mu}_i)}x_t^{\beta_2(\overline{\mu}_i)}-(x_j^F)^{\beta_2(\overline{\mu}_i)}x_t^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)}-(x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}},$$

which admit a clear interpretation: they represent the expected discount factors of the first hitting times of (i) firm j's follower threshold conditional on it being reached before x_i^* is reached, and (ii) the threshold x_i^* conditional on it being reached before firm j's follower threshold is reached, respectively.

Figure 1 depicts the implications of Theorem 1. In case the drop of the payoff from being the only one who has invested to the situation that both players have invested is sufficiently big, the value x^* distinguishes

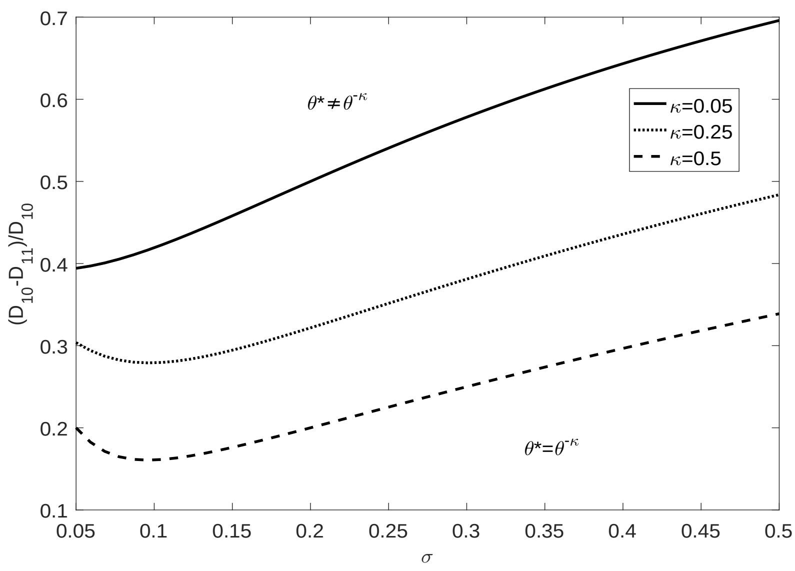


Figure 2 The boundary between the region where $\underline{\mu}$ is always the worst–case trend and the region where it is not. Other parameters are: $r = 0.1$, $\mu = 0.04$, and $I = 120$.

between the regions where each of the two "fears" dominates. For sufficiently small values of x_t, the threat of the competitor investing is relatively low, so that "fear of the market" dominates. For values of x_t close enough to x_j^F, the threat of the competitor investing dominates the payoff uncertainty, so that "fear of the competitor" dominates.

Figure 2 illustrates condition (9) as a function of both volatility (σ) and ambiguity (κ). It can be seen that the higher the level of κ, the more likely it is that for lower first–mover advantages the worst–case trend is not always given by $\underline{\mu}$. In addition, it is also clear that the relationship is not monotonic in σ for higher levels of κ. This non–monotonicity follows from the non–linear (multiplicative) way in which κ and σ interact in $\beta_1(\underline{\mu})$. This feature is, therefore, not peculiar to our particular numerical example and should be expected more generally. From Figure 2 we conclude that a switch in the worst–case leader value is most likely in low–volatility, high–ambiguity, and high first–mover advantage environments.

For the proof of Theorem 1, we need a different approach compared to the standard literature on real option games. We use backward stochastic differential equations and g-expectations as introduced by Peng (1997). The advantage of this approach lies in the fact that we know the value of our problem at the entry point of the follower. That value yields the starting point for a BSDE. A non–linear Feynman–Kac formula then reduces the problem to solving a particular non–linear partial differential equation. From this PDE we are eventually able to derive the worst–case prior and the value function.

Proof. To keep notation simple, we ignore (without loss of generality) the sunk costs I_i.

1. Denote

$$Y_t := \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_s ds + \int_{\tau_j^F}^{\infty} e^{-r(s-t)}D_{11}X_s ds \Big| \mathscr{F}_t\right].$$

Applying the time consistency property of a strongly rectangular set of density generators gives

$$\begin{aligned}
Y_t &= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_s ds + \int_{\tau_j^F}^{\infty} e^{-r(s-t)}D_{11}X_s ds \Big| \mathscr{F}_t\right]\\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\inf_{\mathsf{Q}'\in\mathscr{P}^{\Theta}} \mathsf{E}^{\mathsf{Q}'}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_s ds + \int_{\tau_j^F}^{\infty} e^{-r(s-t)}D_{11}X_s ds \Big| \mathscr{F}_{\tau_j^F}\right] \Big| \mathscr{F}_t\right]\\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_s ds + e^{-r(\tau_j^F-t)} \inf_{\mathsf{Q}'\in\mathscr{P}^{\Theta}} \mathsf{E}^{\mathsf{Q}'}\left[\int_{\tau_j^F}^{\infty} e^{-r(s-\tau_j^F)}D_{11}X_s ds \Big| \mathscr{F}_{\tau_j^F}\right] \Big| \mathscr{F}_t\right]\\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_j^F} e^{-r(s-t)}D_{10}X_s ds + e^{-r(\tau_j^F-t)}\Phi_i(x_{\tau_j^F}) \Big| \mathscr{F}_t\right],
\end{aligned}$$

where

$$\Phi_i(x_t) := \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\infty} e^{-r(s-t)}D_{11}X_s ds \Big| \mathscr{F}_t\right] = \frac{D_{11}x_t}{r-\underline{\mu}_i}. \tag{12}$$

Chen and Epstein (2002) show that Y_t solves the BSDE

$$-dY_t = g_i(Z_t)dt - Z_t dB_t, \tag{13}$$

where, in our case, the *generator*, g_i, is given by

$$g_i(z) = -\kappa_i|z| - rY_t + X_t D_{10}.$$

It is well-known (cf. Chen and Epstein (2002, Appendix B)) that the upper–rim generator satisfies

$$\theta_t^* Z_t = \max_{y\in\Theta_{i,t}} y\cdot Z_t. \tag{14}$$

This shows that $\theta_{i,t}^* \in \{-\kappa, +\kappa\}$ for all $t \geq 0$.

The terminal boundary condition of the BSDE (13) is given by

$$Y_{\tau_j^F} = \Phi_i(x_j^F),$$

In the terminology of Peng (2013), we now say that the leader value is the g_i–expectation of the random variable $e^{-r(\tau_j^F-t)}\Phi_i(x_j^F)$, and write

$$Y_t = \mathsf{E}_{g_i}\left[e^{-r(\tau_j^F-t)}\Phi_i(x_j^F)\Big|\mathscr{F}_t\right].$$

2. Denote the present value of the leader payoff by L_i, i.e. $L_i(x_t) = Y_t$. The non–linear Feynman–Kac formula (Peng (2013, Theorem 3)) implies that L_i solves the non–linear PDE

$$\mathscr{L}_X L_i(x) + g_i(\sigma x L_i'(x)) = 0,$$

where $\mathscr{L}_X$ is the characteristic operator of the SDE (1). [Note that Peng (1991) shows that the non–linear Feynman–Kac formula not only holds for deterministic times but also for first exit times like τ_j^F, even if it does not hold a.s. that $\{\tau_j^F < \infty\}$.]

Hence, L_i solves

$$\frac{1}{2}\sigma^2 x^2 L_i''(x) + \mu x L_i'(x) - \kappa_i \sigma x \left|L_i'(x)\right| - rL_i(x) + D_{10}x = 0. \tag{15}$$

Equation (15) implies that $\underline{\mu}_i$ is the worst case on the set $\{x \leq x_j^F | L_i'(x) > 0\}$ and $\overline{\mu}_i$ is the worst case on $\{x \leq x_j^F | L_i'(x) < 0\}$.

The unique viscosity solution to the PDE (15) is given by

$$L_i(x) = \frac{D_{10}x}{r-\mu} + Ax^{\beta_1(\mu)} + Bx^{\beta_2(\mu)}, \tag{16}$$

where μ equals either $\underline{\mu}_i$ or $\overline{\mu}_i$. The constants A and B are determined by some boundary conditions.

One can easily see that for x close to zero we have $L_i'(x) > 0$. Now two cases are possible: Either $L_i'(x) > 0$ for all $x \in [0, x_j^F]$ or we can find (at least) one point x_i^* at which $L_i'(x_i^*) = 0$.

3. Let us first assume that $L_i'(x) > 0$ for all $x \in [0, x_j^F)$. Then $\underline{\mu}_i$ is always the worst–case drift. Since $\beta_2(\underline{\mu}_i) < 0$, we have that $B = 0$. In order to determine the constant A, we apply a value–matching condition at x_j^F, i.e.

$$L_i(x_j^F) = \frac{D_{10}x_j^F}{r-\underline{\mu}_i} + A(x_j^F)^{\beta_1(\underline{\mu}_i)} = \frac{D_{11}x_j^F}{r-\underline{\mu}_i}.$$

This implies that

$$A = \frac{D_{10}-D_{11}}{r-\underline{\mu}_i}(x_j^F)^{1-\beta_1(\underline{\mu}_i)},$$

and therefore that

$$L_i(x_t) = \frac{D_{10}x_t}{r-\underline{\mu}_i} + \left(\frac{x_t}{x_j^F}\right)^{\beta_1(\underline{\mu}_i)} \frac{D_{11}-D_{10}}{r-\underline{\mu}_i} x_j^F. \tag{17}$$

Due to the continuity and concavity of the value function (17), the condition $L_i'(x) \geq 0$ for all $x \leq x_j^F$ is equivalent to $L_i'(x_j^F) \geq 0$. Therefore,

$$\begin{aligned} L_i'(x_j^F) = \frac{D_{10}}{r-\underline{\mu}_i} + \left(\frac{D_{11}-D_{10}}{r-\underline{\mu}_i}\right)\beta_1(\underline{\mu}_i)\left(\frac{x_j^F}{x_j^F}\right)^{\beta_1(\underline{\mu}_i)-1} &\geq 0 \\ \iff D_{11}-D_{10} &\geq -\frac{D_{10}}{\beta_1(\underline{\mu}_i)} \\ \iff \frac{D_{10}-D_{11}}{D_{10}} &\leq \frac{1}{\beta_1(\underline{\mu}_i)}. \end{aligned} \tag{18}$$

4. If inequality (18) is not satisfied, the worst–case drift changes at some point $x_i^* < x_j^F$ from $\underline{\mu}_i$ to $\overline{\mu}_i$, where x_i^* is determined by the condition $L_i'(x_i^*) = 0$. We denote by $\tilde{L}_i(x; \underline{\mu}_i)$ the solution to (16) on $[0, x_i^*]$ and by

$\hat{L}_i(x;\overline{\mu}_i)$ the solution to (16) on $[x_i^*, x_j^F]$. The unknowns in (16) are determined by twice applying a value–matching condition and once a smooth–pasting condition (see also Cheng and Riedel (2013)). Indeed, it must hold that

1. $\hat{L}_i(x_j^F;\overline{\mu}_i) = \Phi_i(x_j^F)$,
2. $\tilde{L}_i(x_i^*;\underline{\mu}_i) = \hat{L}_i(x_i^*;\overline{\mu}_i)$,
3. $\tilde{L}_i'(x_i^*;\underline{\mu}_i) = \hat{L}_i'(x_i^*;\overline{\mu}_i)$.

In case $\underline{\mu}_i$ is not always the worst–case drift, the unique viscosity solution of (16) is given by

$$L_i(x_t) = 1_{x_t<x_i^*}\tilde{L}_i(x_t;\underline{\mu}_i) + 1_{x_t\geq x_i^*}\hat{L}_i(x_t;\overline{\mu}_i),$$

where

$$\tilde{L}_i(x_t;\underline{\mu}_i) = \frac{D_{10}x_t}{r-\underline{\mu}_i} - \frac{1}{\beta_1(\underline{\mu}_i)}\frac{D_{10}x_i^*}{r-\underline{\mu}_i}\left(\frac{x_t}{x_i^*}\right)^{\beta_1(\underline{\mu}_i)},$$

and

$$\begin{aligned}\hat{L}_i(x_t;\overline{\mu}_i) =& \frac{D_{10}x_t}{r-\overline{\mu}_i} + \frac{(x_i^*)^{\beta_2(\overline{\mu}_i)}x_t^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}x_t^{\beta_2(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_{i\,i}}\right)x_j^F \\ &+ \frac{(x_j^F)^{\beta_1(\overline{\mu}_i)}x_t^{\beta_2(\overline{\mu}_i)} - (x_j^F)^{\beta_2(\overline{\mu}_i)}x_t^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^*.\end{aligned}$$

We can easily verify that $\hat{L}_i$ and $\tilde{L}_i$ satisfy the boundary conditions. Indeed,

$$\begin{aligned}\hat{L}_i(x_j^F;\overline{\mu}_i) =& \frac{D_{10}x_j^F}{r-\overline{\mu}_i} + \frac{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F \\ &+ \frac{(x_j^F)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)} - (x_j^F)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^* \\ =& \frac{D_{10}x_j^F}{r-\overline{\mu}_i} + \left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F = \frac{D_{11}x_j^F}{r-\underline{\mu}_i} = \Phi_i(x_j^F).\end{aligned}$$

and

$$\begin{aligned}\hat{L}_i(x_i^*;\overline{\mu}_i) =& \frac{D_{10}x_i^*}{r-\overline{\mu}_i} + \frac{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_i^*)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_i^*)^{\beta_2(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F \\ &+ \frac{(x_j^F)^{\beta_1(\overline{\mu}_i)}(x_i^*)^{\beta_2(\overline{\mu}_i)} - (x_j^F)^{\beta_2(\overline{\mu}_i)}(x_i^*)^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^* \\ =& \frac{D_{10}x_i^*}{r-\overline{\mu}_i} + \left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^* \\ =& \frac{D_{10}x_i^*}{r-\underline{\mu}_i} - \frac{1}{\beta_1(\underline{\mu}_i)}\frac{D_{10}x_i^*}{r-\underline{\mu}_i} = \tilde{L}_i(x_i^*;\underline{\mu}_i).\end{aligned}$$

To prove the smooth–pasting condition at x_i^* requires a bit more work. Firstly, we observe that $\tilde{L}_i'(x_i^*;\underline{\mu}_i) = 0$ by construction. The next lemma shows that there exists a unique value x_i^*, which satisfies $\hat{L}_i'(x_i^*;\overline{\mu}_i) = 0$.

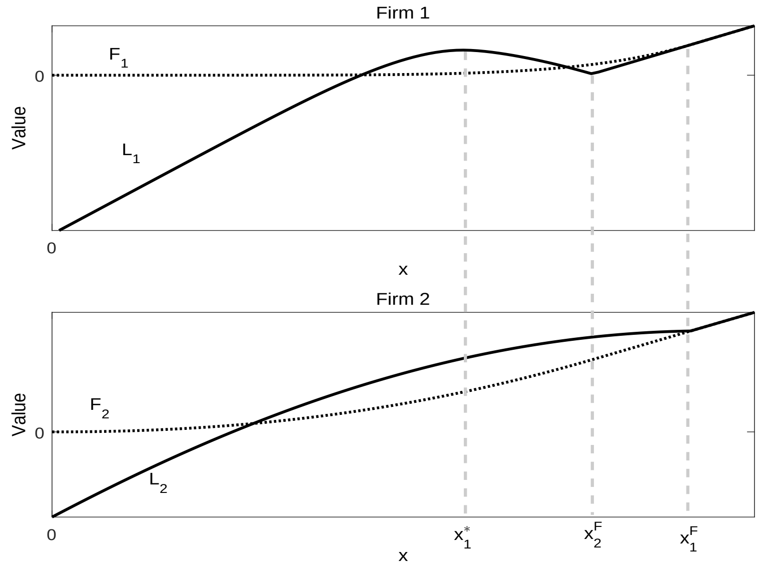


Figure 3 Typical leader and follower value functions.

Lemma 2 *If* $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu}_i)}$*, then there exists one, and only one, value* x_i^* *that solves* $\hat{L}_i'(x_i^*;\overline{\mu}_i) = 0$ *on* $(0, x_j^F]$.

The proof of this lemma is reported in Appendix A. ■

Remark 1 *The leader value function* L_i *is always concave on* $[0, x_j^F]$ *even if the worst case changes at some point. We prove this fact in Appendix B.*

Figure 3 shows a typical run of the leader and follower value functions, where we assume that κ_1 and κ_2 are such that

$$\frac{1}{\beta_1(\underline{\mu}_1)} < \frac{D_{10}-D_{11}}{D_{10}} < \frac{1}{\beta_1(\underline{\mu}_2)}.$$

We observe that the leader value of firm 1 drops below its follower value if x_t is close to x_2^F. The reason for that is that x_1^F and x_2^F are unequal (in the illustrated case we have $x_2^F < x_1^F$).

3.4. Optimal Leader Threshold

Next we want to determine the optimal time to invest as a leader. We focus (without loss of generality) on firm 1, for reasons that will become clear in the next section. If firm 1 knows that it will not be preempted by firm 2, it will determine its investment time by solving the optimal stopping problem

$$\begin{aligned} L_1^*(x_t) = \sup_{\tau_{L,1}^t \geq t} \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\Bigg[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,i}^t}^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds \\ + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - e^{-r(\tau_{L,1}^t - t)} I_1 \Big| \mathscr{F}_t \Bigg]. \end{aligned} \tag{19}$$

Again, in order to determine this stopping time, we cannot apply the standard procedure. Nevertheless, the investment trigger does not differ from the one of a non–ambiguous firm faced with drift $\underline{\mu}_1$.

Proposition 1 *The optimal time to invest for firm 1 is given by* $\tau^t_{L,1} = \inf\{s \geq t | X_s \geq x^L_1\}$, *where*

$$x^L_1 = \frac{\beta_1(\underline{\mu}_1)}{\beta_1(\underline{\mu}_1) - 1} \frac{r - \underline{\mu}_1}{D_{10} - D_{00}} I_1.$$

For the proof we refer to Appendix C.

As in the case without ambiguity, the optimal leader threshold is independent of the follower threshold. Our assumption that $D_{10} - D_{00} > D_{11} - D_{01}$ guarantees that $x^L_1 < x^F_1$ and, thus, that the firm wants to become leader before it wants to become follower. This assumption therefore ensures that this is a market characterised by preemption rather than attrition.

4. Equilibrium Analysis

The appropriate equilibrium concept for a game with ambiguity as described here is not immediately clear. We obtain our equilibrium results by using techniques developed in Riedel and Steg (2017). It should be pointed out that we cannot simply adopt their strategies to our setting. In fact, the notion of *extended mixed strategy* as introduced in Riedel and Steg (2017) presents a conceptual problem here. Such strategies consist, in essence, of a distribution over stopping times and an "investment intensity". The latter component acts as a coordination device. In our model we also need this coordination device, but we do not want ambiguity to extend to the uncertainty created by it. This presents problems if we want to define payoffs to an ambiguous firm if it plays a mixture over stopping times. For equilibrium existence, however, such mixtures are not needed, so we choose to restrict attention to what we call *extended pure strategies*, which consist of a stopping time and an element related to the coordination mechanism mentioned above. Once we have restricted ourselves to such extended pure strategies, we can refer to equilibrium results obtained by Riedel and Steg (2017) and Steg (2015). Their results mainly carry over to our framework because ambiguity is now about payoff functions only and not about strategies. In other words, there is ambiguity about *when* decisions are taken, but not about *what* decisions are taken at those times. We give a detailed description of the equilibrium concept in Appendix D.

To make comparisons with existing models easier, we will assume that $\kappa_1 = \kappa_2 \equiv \kappa$ and that $I_2 \geq I_1$. This essentially reduces our model to an extension of Steg (2015, Section 4.1) and Pawlina and Kort (2006). Note that, under these assumptions it holds that $x^F_1 \leq x^F_2$. Due to our definition of strategies, most of the results obtained by Steg (2015, Section 4.1) carry over to our setting. In particular, we have that:

Proposition 2 *If* $I_1 < I_2$, *then* $L_1 - F_1 > L_2 - F_2$ *on* $(0, x^F_2)$.

This result is similar to Steg (2015, Lemma 3.2). However, his proof relies on the linearity of the expectation operator. In the case where $\mathsf{P}^*_{L,i} = \mathsf{P}^*_{F,i} = \mathsf{P}^{-\kappa}$, Steg's result therefore applies trivially. If $\mathsf{P}^*_{L,i} \neq \mathsf{P}^{-\kappa}$, however, the leader value is a (non–linear) g–expectation. The result still holds though.

Proof. On $[x_1^F, x_2^F)$ the result is obvious. For $i = 1, 2$, define the process ξ_i^L by

$$\xi^L_{i,t} := \int_t^{\tau_j^F} e^{-r(s-t)} \left(D_{10} X_s - rI_i\right) ds + \int_{\tau_j^F}^{\infty} e^{-r(s-t)} \left(D_{11} X_s - rI_i\right) ds.$$

We then have the following inequality for any $x_t < x_1^F$:

$$\begin{aligned} L_1(x_t) - L_2(x_t) &= \inf_{\mathsf{Q}\in\mathscr{P}^\Theta} \mathsf{E}^{\mathsf{Q}}[\xi^L_{1,t}|\mathscr{F}_t] - \inf_{\mathsf{Q}\in\mathscr{P}^\Theta} \mathsf{E}^{\mathsf{Q}}[\xi^L_{2,t}|\mathscr{F}_t] \\ &= \mathsf{E}^{\mathsf{P}^*_{L,1}}[\xi^L_{1,t}|\mathscr{F}_t] - \mathsf{E}^{\mathsf{P}^*_{L,2}}[\xi^L_{2,t}|\mathscr{F}_t] \\ &\geq \mathsf{E}^{\mathsf{P}^*_{L,1}}[\xi^L_{1,t}|\mathscr{F}_t] - \mathsf{E}^{\mathsf{P}^*_{L,1}}[\xi^L_{2,t}|\mathscr{F}_t] \\ &= \mathsf{E}^{\mathsf{P}^*_{L,1}}[\xi^L_{1,t} - \xi^L_{2,t}|\mathscr{F}_t] \\ &\geq \inf_{\mathsf{Q}\in\mathscr{P}^\Theta} \mathsf{E}^{\mathsf{Q}}[\xi^L_{1,t} - \xi^L_{2,t}|\mathscr{F}_t] \\ &\overset{(*)}{=} \inf_{\mathsf{Q}\in\mathscr{P}^\Theta} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_1^F} e^{-r(s-t)} r(I_2 - I_1) ds + e^{-r\tau_1^F}\left(L_1(x_1^F) - L_2(x_1^F)\right)\Big|\mathscr{F}_t\right], \end{aligned} \tag{20}$$

where $(*)$ follows from the time consistency property of a strongly rectangular set of priors. Note that the assumption $\Theta_1 = \Theta_2$ is crucial here.

The operator (20) can be written as a g–expectation with generator

$$\tilde{g}(y,z) = -\kappa|z| - ry + r(I_2 - I_1),$$

so that we get

$$L_1(x_t) - L_2(x_t) \geq \mathsf{E}_{\tilde{g}}\left[e^{-r\tau_1^F}\left(L_1(x_1^F) - L_2(x_1^F)\right)\Big|\mathscr{F}_t\right].$$

Now define the generator

$$g(y,z) = -\kappa|z| - ry.$$

Since $\tilde{g} > g$ and $L_1(x_1^F) - L_2(x_1^F) > F_1(x_1^F) - F_2(x_1^F)$, it follows from the comparison theorem for BSDEs (Peng (2013, Theorem 2)) that

$$\mathsf{E}_{\tilde{g}}\left[e^{-r\tau_1^F}\left(L_1(x_1^F) - L_2(x_1^F)\right)\Big|\mathscr{F}_t\right] > \mathsf{E}_g\left[e^{-r\tau_1^F}\left(F_1(x_1^F) - F_2(x_1^F)\right)\Big|\mathscr{F}_t\right].$$

It is straightforward to verify that, for the generator g, the worst–case drift is always given by $\underline{\mu}$. Therefore,

$$\mathsf{E}_g\left[e^{-r\tau_1^F}\left(F_1(x_1^F) - F_2(x_1^F)\right)\Big|\mathscr{F}_t\right] = \mathsf{E}^{\mathsf{P}^{-\kappa}}\left[e^{-r\tau_1^F}\left(F_1(x_1^F) - F_2(x_1^F)\right)\Big|\mathscr{F}_t\right] = F_1(x_t) - F_2(x_t),$$

which proves the result. ■

We define the *preemption region* as the part of the state space where both firms prefer to be the leader rather than the follower, i.e.

$$\mathcal{P} = \{x \in \mathbb{R}_+ | (L_1(x) - F_1(x)) \wedge (L_2(x) - F_2(x)) > 0\}.$$

The first hitting time of $\mathcal{P}$ is denoted by τ_P. From Lemma 2 and Steg (2015, Proposition 4.1) it now immediately follows that

$$\mathcal{P} = \{x \in \mathbb{R}_+ | L_2(x) - F_2(x) > 0\} = (\underline{x}, \overline{x}),$$

for some $0 < \underline{x} \leq \overline{x} \leq x_2^F$. Note that it is possible that $\underline{x} = \overline{x}$ and, thus, that $\mathcal{P} = \emptyset$.

The following result immediately follows from the observation that firm 2's leader value (under $\mathsf{P}^*_{L,2}$) is, by definition, no larger than its leader value under $\mathsf{P}^{-\kappa}$.

Lemma 3 *If firm 2 evaluates its leader and follower values under $\mathsf{P}^*_{L,2}$ and $\mathsf{P}^*_{F,2}$, respectively, then the preemption region is always contained in the preemption region that results from their valuations under $\mathsf{P}^{-\kappa}$.*

This lemma shows that our model with ambiguity is not isomorphic to a model without ambiguity and the unique prior $\mathsf{P}^{-\kappa}$. In fact, ambiguity reduces preemptive pressure.

In the literature there are typically two types of equilibria analyzed: *preemptive equilibria* in which firms try to preempt each other at some times where it is sub–optimal to invest, and *sequential equilibria*, where one firm invests at its optimal time. In our setting, the question which of these two types of equilibrium prevails boils down to comparing the location of the optimal leader threshold x_1^L in relation to the preemption region $(\underline{x}, \bar{x})$. Roughly speaking, a preemptive equilibrium arises if $x_1^L \in (\underline{x}, \bar{x})$, whereas a sequential equilibrium arises when $x_1^L \leq \underline{x}$.

Steg (2015), however, notes that another equilibrium scenario is possible. For certain underlying parameters, it may be the case that $x_1^L > \bar{x}$. In our model, it then holds that if $\bar{x} < x_0 < x_2^F$, firm 2 will not immediately invest. That gives firm 1 the option to wait a bit longer. In fact, firm 1 solves the following optimal stopping problem to determine when to invest:

$$\sup_{\tau \leq \tau_P \wedge \tau_2^F} \inf_{\mathsf{Q} \in \mathscr{P}^\Theta} \mathsf{E}^{\mathsf{Q}} \left[\int_\tau^\infty e^{-rs}(X_s(D_{10} - D_{00}) - rI_1) ds \right]. \tag{21}$$

Since τ_1^L (the first hitting time of x_1^L) is the solution to the *unconstrained* leader problem (19), the stopping region of the *constrained* problem (21) must contain the interval $[x_1^L, x_2^F]$. This implies that the continuation region of (21) lies entirely in the region where $\underline{\mu}$ is the worst–case drift. Therefore, Steg (2015, Proposition 4.2) applies, which states that, assuming that $0 < \bar{x}(D_{10} - D_{00}) < rI_1$, the continuation region of problem (21) is $(\bar{x}, \hat{x} \wedge x_2^F)$, where $\hat{x} \in [rI_1/(D_{10} - D_{00}), x_1^L)$ is the unique solution to the equation

$$(\beta_1(\underline{\mu}) - 1)A(x)x^{\beta_1(\underline{\mu})} + (\beta_2(\underline{\mu}) - 1)B(x)x^{\beta_2(\underline{\mu})} = I_1, \tag{22}$$

where

$$\begin{bmatrix} A(x) \\ B(x) \end{bmatrix} = \frac{1}{\bar{x}^{\beta_1(\underline{\mu})} x^{\beta_2(\underline{\mu})} - x^{\beta_1(\underline{\mu})} \bar{x}^{\beta_2(\underline{\mu})}} \begin{bmatrix} x^{\beta_2(\underline{\mu})} - \bar{x}^{\beta_2(\underline{\mu})} \\ -x^{\beta_1(\underline{\mu})} \bar{x}^{\beta_2(\underline{\mu})} \end{bmatrix} \begin{bmatrix} \bar{x} \frac{D_{10}-D_{00}}{r-\underline{\mu}} - I_1 \\ x \frac{D_{10}-D_{00}}{r-\underline{\mu}} - I_1 \end{bmatrix}.$$

That is, the threshold $\hat{x}$ is determined in the standard way, by applying the value–matching and smooth–pasting optimality conditions.

If it holds that $\bar{x}(D_{10} - D_{00}) \geq rI_1$, then firm 1 will always invest immediately in the region $(\bar{x}, x_2^F)$ and we can set $\hat{x} = \bar{x}$. In either case, firm 1 invests at the first hitting time, $\hat{\tau}$, of the set

$$\mathcal{A} = (\underline{x}, \bar{x}) \cup [\hat{x}, \infty) \cup [x_2^F, \infty).$$

If $\bar{x}$ or $\hat{x}$ is hit before x_2^F, then firm 1 becomes the leader, whereas if x_2^F is hit first both firms invest at the same time. However, since $M_1(x_2^F) = L_1(x_2^F)$, we can say that firm 1 earns the leader payoff at $\hat{\tau}$, no matter which threshold is reached first.

In the case where $\underline{\mu}$ is not always the worst–case drift, the above scenario is more likely to occur than in the case where $\underline{\mu}$ is always the worst–case drift (cf. Lemma 3). Since the preemption region $(\underline{x}, \bar{x})$ under $\mathsf{P}^*_{L,2}$ and $\mathsf{P}^*_{F,2}$ is a subset of the same set under $\mathsf{P}^{-\kappa}$, and because x_1^L is the same under $\mathsf{P}^*_{L,1}$ and $\mathsf{P}^{-\kappa}$, it follows that the condition $\bar{x} < x_1^L$ is more easily met under the worst–case priors than under $\mathsf{P}^{-\kappa}$.

To summarize, depending on the starting point x_0 of the underlying stochastic process, different investment scenarios along the equilibrium path are now possible (see Appendix D for a formal statement of the subgame perfect equilibrium):

Case 1(a). $x_0 < \underline{x}$ and $\underline{x} \leq x_1^L \leq \bar{x}$: preemptive equilibrium where firm 1 becomes the leader at time τ_P and firm 2 follows at τ_2^F;

Case 1(b). $x_0 < \underline{x}$ and $x_i^L < \underline{x}$: sequential equilibrium with firm 1 investing at time τ_1^L and firm 2 following at τ_2^F;

Case 2. $\underline{x} < x_0 \leq \bar{x}$: preemptive equilibrium with at least one firm investing immediately, and where preemption results in rent equalization;

Case 3(a). $\bar{x} < x_0 < x_2^F$ and $x_1^L \leq \bar{x}$: firm 1 becomes the leader immediately and firm 2 follows at τ_2^F;

Case 3(b). $\bar{x} < x_0 < x_2^F$ and $\bar{x} < x_1^L < x_2^F$: firm 1 becomes the leader at time $\hat{\tau}$ and firm 2 follows at τ_2^F;

Case 4. $x_0 \geq x_2^F$: simultaneous investment by both firms.

5. Numerical Examples

In this section, we study numerically how the presence of ambiguity affects the equilibrium scenarios presented in the previous section. We continue to assume that $\Theta_1 = \Theta_2$, and that $I_2 > I_1$. Throughout, we will use a base–case scenario, with parameter values as in Table 1. Note that this is a so–called "new market model" ($D_{00} = D_{01}$) so that firm 1 always invests immediately in the region $(\bar{x}, x_2^F)$, i.e. Case 3(b) does not

$D_{10} = 1.8$	$D_{11} = 1$	$D_{00} = 0$
$D_{01} = 0$	$I_1 = 100$	$I_2 = 120$
$r = 0.1$	$\mu = 0.04$	$\sigma = 0.1$

Table 1 Parameter values for a base–case numerical example.

arise (cf. Steg (2015)). In the remainder, we shall denote the value function of firm i under the equilibrium scenario following from Section 4 by $x_0 \mapsto V_i(x_0)$.

The presence of ambiguity changes the drift under which the firms evaluate their payoffs. In the standard model without ambiguity a reduction in the trend μ reduces firm 1's leader threshold x_1^L and increases firm 2's preemption threshold $\underline{x}$. Typically, a reduction in μ also shifts the preemption region $(\underline{x}, \bar{x})$, as can be seen from Figure 4.

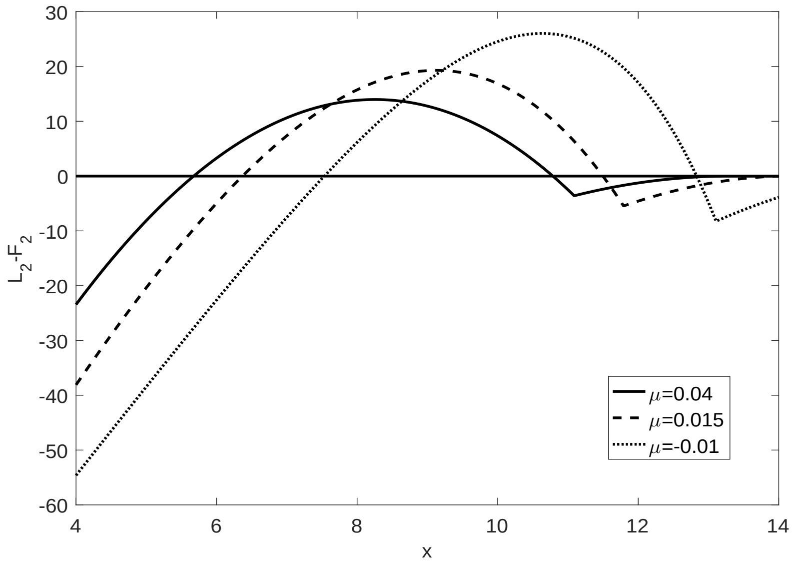


Figure 4 Preemption regions for different values of μ and no ambiguity ($\kappa = 0$). Other parameter values are taken from Table 1.

This insight also applies to cases where $\underline{\mu}$ is always the worst–case drift ($\mathsf{P}^*_{L,i} = \mathsf{P}^{-\kappa}$), i.e. when

$$\frac{D_{10} - D_{11}}{D_{10}} \leq \frac{1}{\beta_1(\underline{\mu})}.$$

For this case, the analysis of equilibria and firm values can proceed *as if* the firms use a unique prior, i.e., a GBM with trend $\underline{\mu} = \mu - \kappa\sigma$.

The situation is different when

$$\frac{D_{10} - D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu})}.$$

In this case, firm i uses the worst–case prior $\mathsf{P}^*_{L,i} \neq \mathsf{P}^{-\kappa}$ to value the leader role, which is generated by θ^*_i with

$$\theta^*_{i,t} = -\kappa \text{ on } \left\{ X_t \in [0, x^*_i] \cup [x^F_j, \infty) \right\} \quad \text{and} \quad \theta^*_{i,t} = +\kappa \text{ on } \left\{ X_t \in (x^*_i, x^F_j) \right\}.$$

Since $I_2 > I_1$, it holds that $x^F_2 > x^F_1$ and, thus, that $x^*_2 \neq x^*_1$. This implies that the two firms will use different priors to value the leader payoff on $\left\{ \theta^*_{1,t} \neq \theta^*_{2,t} \right\}$, which is a P–non–null set. This happens even though the firms are entirely symmetric, bar the sunk costs of investment.

This observation has some consequences for equilibria. First, the preemption region is different from the one under the prior $\mathsf{P}^{-\kappa}$ (cf. Lemma 3). This can be seen in Figure 5. For $\kappa = 0.25$ and $\kappa = 0.5$ the worst–case drift for the leader payoff is not always given by $\underline{\mu}$. It can be seen that the preemption region under $\mathsf{P}^*_{L,2}$ is then a subset of the preemption region that would be obtained if firm 2 were to use the prior $\mathsf{P}^{-\kappa}$ to value the leader role.

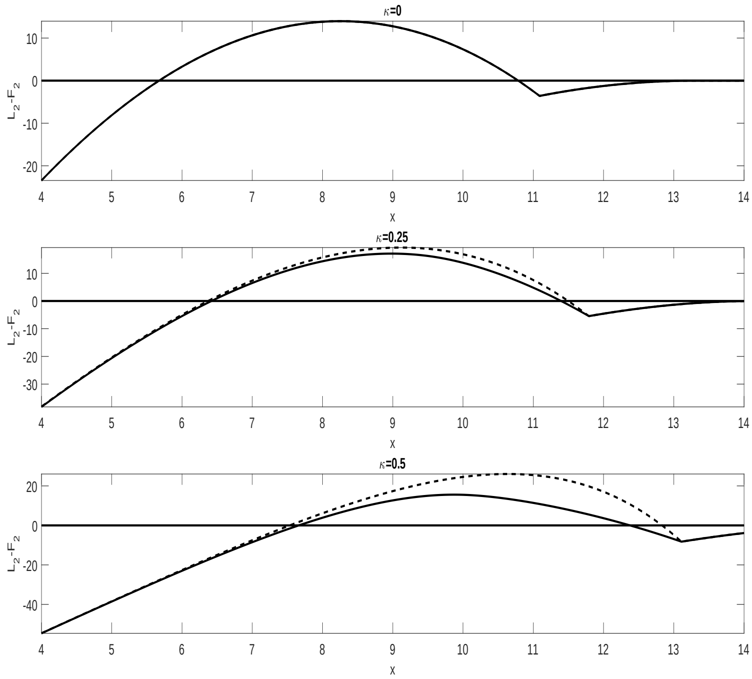


Figure 5 Preemption regions for different values of κ under the priors $\mathsf{P}^*_{L,2}$ (solid line) and $\mathsf{P}^{-\kappa}$ (dashed line). Other parameter values are taken from Table 1.

Qualitatively, equilibria may or may not change as κ increases. This can be seen from Figure 6. In Figure 6a we plot $L_2 - F_2$ for $\kappa = 0$ and $\kappa = 0.25$. In both cases it holds that $\underline{x} < x^L_1$, so that first investment takes place at the first hitting time of the set $[\underline{x}, \infty)$. To be more specific:

1. simultaneous investment takes place for $x_0 \in [x^F_2, \infty)$,

2. firm 1 immediately becomes the leader for $x_0 \in (\bar{x}, x_2^F)$ (with firm 2 following at τ_2^F),
3. firms immediately compete for the leader role if $x_0 \in (\underline{x}, \bar{x})$, resulting in rent equalization, and
4. firm 1 becomes the leader at $\tau_{\mathcal{P}}$ for $x_0 \leq \underline{x}$.

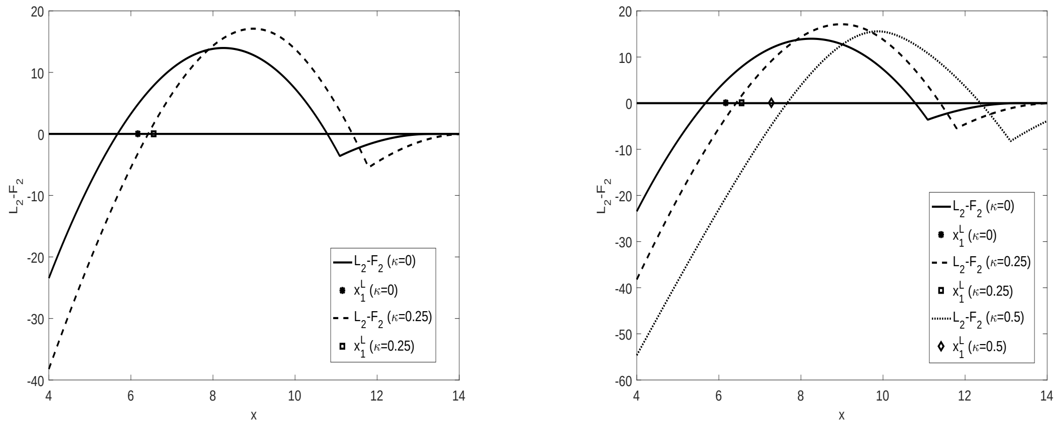
Firm values in this equilibrium are given by

$$V_1(x_t) = \begin{cases} \frac{D_{00}x_t}{r-\underline{\mu}} + \left(\frac{x_t}{\underline{x}}\right)^{\beta_1(\underline{\mu})} \left[L_1(\underline{x}) - \frac{D_{00}\underline{x}}{r-\underline{\mu}}\right] & \text{if } x_t \leq \underline{x} \\ F_1(x_t) & \text{if } \underline{x} < x_t < \bar{x} \\ L_1(x_t) & \text{if } \bar{x} \leq x_t < x_2^F \\ M_1(x_t) & \text{if } x_t \geq x_2^F, \end{cases}$$

and

$$V_2(x_t) = \begin{cases} \frac{D_{00}x_t}{r-\underline{\mu}} + \left(\frac{x_t}{\underline{x}}\right)^{\beta_1(\underline{\mu})} \left[F_2(\underline{x}) - \frac{D_{00}\underline{x}}{r-\underline{\mu}}\right] & \text{if } x_t \leq \underline{x} \\ F_2(x_t) & \text{if } \underline{x} < x_t < x_2^F \\ M_2(x_t) & \text{if } x_t \geq x_2^F. \end{cases}$$

Figure 7a shows that the increase in ambiguity lowers equilibrium value for both firms. The first (downward) jump in the value of firm 1 occurs at $\underline{x}$. At that point, firm 1 no longer expects to earn the leader value, but the follower value instead, due to preemptive pressure. Between $\bar{x}$ and x_2^F, firm 1 expects to become the leader again, which explains the second (upward) jump in its equilibrium value function.



(a) $\kappa \in \{0, 0.25\}$ (b) $\kappa \in \{0, 0.25, 0.5\}$

Figure 6 Preemption regions for different values of κ under the priors $\mathsf{P}^*_{L,2}$ (solid line) and $\mathsf{P}^{-\kappa}$ (dashed line). Other parameter values are taken from Table 1. Also indicated are the leader thresholds for firm 1, x_1^L.

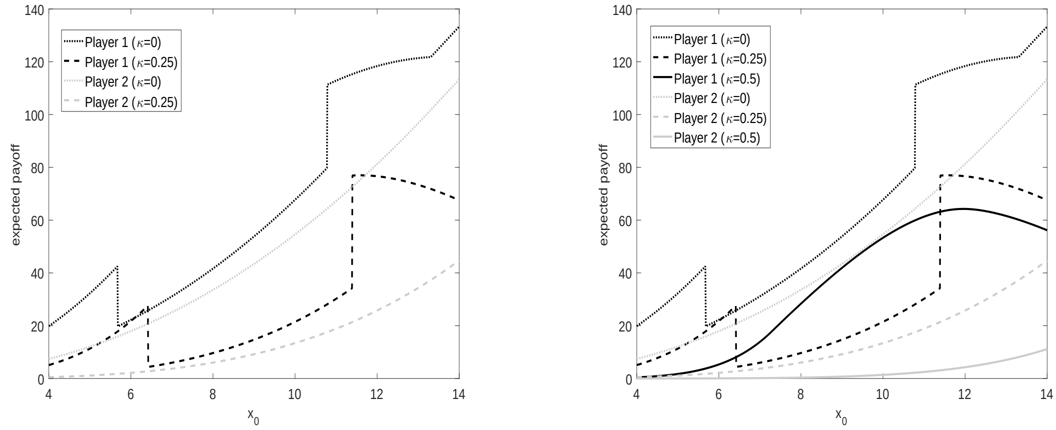
If we now add the case where $\kappa = 0.5$, we can see from Figure 6b that here it no longer holds that $\underline{x} < x_1^L$. Instead, a sequential equilibrium arises with equilibrium firm values

$$V_1(x_t) = \begin{cases} \frac{D_{00} x_t}{r - \underline{\mu}} + \left(\frac{x_t}{x_1^L} \right)^{\beta_1(\underline{\mu})} \left[L_1(x_1^L) - \frac{D_{00} x_1^L}{r - \underline{\mu}} \right] & \text{if } x_t < x_1^L \\ L_1(x_t) & \text{if } x_1^L \leq x_t < x_2^F \\ M_1(x_t) & \text{if } x_t \geq x_2^F, \end{cases}$$

and

$$V_2(x_t) = \begin{cases} \frac{D_{00} x_t}{r - \underline{\mu}} + \left(\frac{x_t}{x_1^L} \right)^{\beta_1(\underline{\mu})} \left[F_2(x_1^L) - \frac{D_{00} x_1^L}{r - \underline{\mu}} \right] & \text{if } x_t < x_1^L \\ F_2(x_t) & \text{if } x_1^L \leq x_t < x_2^F \\ M_2(x_t) & \text{if } x_t \geq x_2^F. \end{cases}$$

Figure 7b shows that, for firm 2, this change in equilibrium scenario does not change anything qualitatively: an increase in ambiguity leads to a reduction in equilibrium firm value. For firm 1, however, the story is now different. In fact, there exists a region where firm 1 has a higher value with *higher* ambiguity. This happens because in the higher ambiguity case ($\kappa = 0.5$), firm 1 always becomes the leader, whereas in the lower ambiguity case ($\kappa = 0.25$), firm 1 expects the (lower) follower value, due to preemptive pressure and rent equalization.



(a) $\kappa \in \{0, 0.25\}$ (b) $\kappa \in \{0, 0.25, 0.5\}$

Figure 7 Equilibrium firm values (black for firm 1 and gray for firm 2) for different values of κ. Other parameter values are taken from Table 1.

Another comparative statics exercise of interest is comparing the payoffs of an ambiguous firm with the payoffs of an unambiguous firm. This exercise is relevant, e.g., in cases where one firm is an incumbent who has learned the distribution of future revenues, whereas the other firm may be an entrant which has no information about its revenue distribution. A question that then arises is whether the unambiguous firm benefits from its competitor's ambiguity aversion. The intuition here is clear: yes. The ambiguity–averse competitor has a higher value of waiting and will be willing to invest later. So, even if the firms are completely symmetric except for their attitudes to ambiguity, the non-ambiguous firm is always expecting to become the leader, either under preemptive pressure, or at its optimal time. To illustrate this phenomenon, we use the same numerical example as before, with $I_1 = I_2$ and $0 = \kappa_1 < \kappa_2$. In that case the equilibrium result is as before (because $x_1^F < x_2^F$ and $x_1^L < x_2^F$) and the only determinant of which equilibrium prevails is the ordering of x_1^L and $\underline{x}$, and the ordering of x_1^F and $\overline{x}$. The results are shown in Figure 8. Note that for the case $\kappa_2 = 0$ both firms' values are the same. The value for the non–ambiguous firm is increasing in the level of ambiguity of the ambiguous firm. This is particularly stark for cases where ambiguity leads to a situation where $x_1^L < \underline{x}$ (as is the case for $\kappa_2 = 0.3$) so that firm 1 can essentially ignore the preemptive pressure exerted by firm 2.

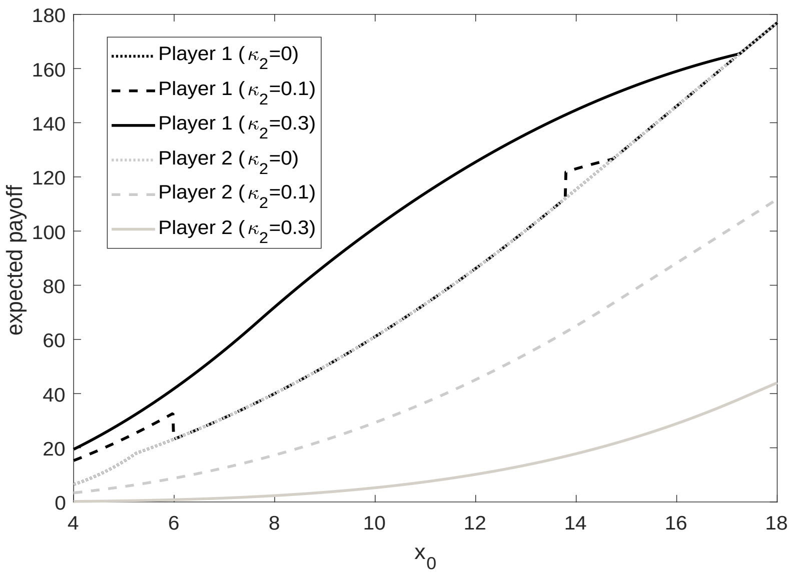


Figure 8 Equilibrium firm values (black for firm 1 and gray for firm 2) for different values of κ_2. Other parameter values are taken from Table 1, with $I_2 = I_1 = 100$ and $\kappa_1 = 0$. The non–ambiguous firm uses the reference measure (i.e. trend μ).

6. Discussion and Concluding Remarks

In this paper we have built a model of an investment game with a first–mover advantage between two firms, at least one of which is ambiguous about the probability measure over the sample paths of future revenues. Our main conclusions are that: (i) the worst–case drift for the leader value can switch between $\underline{\mu}$ and $\overline{\mu}$, (ii) despite the complicated g-expectation under which the leader value has to be computed in those cases, equilibrium existence results from the literature on non–ambiguous investment mainly carry over, and (iii) ambiguity reduces preemptive pressure, when compared to a simple reduction of the trend in a non–ambiguous model. In this section we will further discuss some of our modeling assumptions and point to some avenues for future research.

Our main assumptions fall into two categories: payoff and ambiguity assumptions. The payoff assumptions are standard in the real options literature. In particular, the multiplicative nature of the way uncertainty enters revenues is well–established. The zero (operating) cost assumption can easily be relaxed to the case of a deterministic cost flow, in which case the present value of operating costs can be subsumed in the sunk costs of investment. Note that all our $D_{k\ell}$s are non–negative. This restriction is required in our model, because for negative payoff streams, the worst–case prior is always opposite to the one with positive $D_{k\ell}$s. In other words, in a model with ambiguity, gains and losses are evaluated under different priors. We want to rule out this complication.

The assumption $D_{10} - D_{00} > D_{11} - D_{01}$ is standard and reflects the first–mover advantage that we have analysed in this paper. The assumption ensures that $x_i^L < x_i^F$ so that no firm would prefer to become the follower before it prefers to become the leader. The assumption is strong in the sense that the market can *never* exhibit a second–mover advantage. For a model with interchanging first–mover and second–mover advantages, see Steg and Thijssen (2015). Their model has a two–dimensional state–space, however, which increases the technical complexity of the model substantially.

Regarding ambiguity, the assumption of κ–ambiguity is a strong one, albeit one that is often made in the continuous–time literature. While its main technical advantage is to parameterize a strongly rectangular set of priors, it also implies that the density generator satisfies $\theta_t \in [-\kappa, \kappa]$, regardless of the time t. Therefore, there is no possibility of learning about the true measure in our model. A way to interpret this assumption is given by Chen and Epstein (2002) and Epstein and Schneider (2010) and is best explained in discrete time by means of a sequence of draws from an Ellsberg urn. In a model with learning the decision–maker repeatedly draws balls (with replacement) from the same Ellsberg urn, thereby having the opportunity to shrink the set of priors over time. A model without learning, like ours, is appropriate when the decision–maker draws from a *different* Ellsberg urn every time. This prevents the decision–maker from learning about ambiguity. Suppose that the process X represents the (market) demand flow for the firm's product. If the manager thinks that the expected growth rate might not be constant over time then our model without learning is appropriate. For example, X may represent market demand, the expected growth rate of which may vary over time due to changing consumer tastes. It might then be better for the manager to consider an interval of trends and apply a precautionary principle in valuation. If the manager beliefs that the expected growth rate of demand is constant and learnable through observations of actual demand, then a model with learning is more appropriate.

If our model were extended to include learning, we would expect the interval $[\underline{\mu},\overline{\mu}]$ to shrink over time. Essentially, such a model would be an extension to ambiguity of the model analysed in Décamps et al. (2005). Of course, it can not be predicted ex ante in which direction the interval will shrink. For example, if the true measure is $\mathsf{P}^{-\kappa}$ then, over time, the highest possible trend will reduce. This has no effect on firm values in the case where $\underline{\mu}$ always represents the worst–case drift. It will, however, increase the leader value in the case where $\underline{\mu}$ is not always the worst–case drift. This, in turn should lead to a widening of the preemption region and, thus, to earlier investment. If the true probability measure is some $\mathsf{P} \notin \{\mathsf{P}^{-\kappa},\mathsf{P}^{+\kappa}\}$, one would expect that, over time, both $\underline{\mu}$ and $\overline{\mu}$ will move closer together. This will increase both leader and follower values and will decrease the leader and follower thresholds. The effect on the preemption region and resulting equilibria is unclear and will be left for future research.

To the best of our knowledge, there have been no empirical studies into the effects of risk aversion or ambiguity on corporate investment timing decisions. In a real options model, risk aversion leads to delay of investment (see, e.g., Thijssen (2010)). In that sense, our theory leads to the same qualitative conclusion: ambiguity leads to a further delay of investment. This opens up the possibility of empirical research into ambiguity over (competitive) growth options, much in the same way as in the literature on the equity premium puzzle. As Mehra and Prescott (1985) have famously shown, in the standard expected utility theory of asset pricing empirically observed equity premia are compatible only with implausibly large coefficients of risk aversion. Trojani and Vanini (2004) have shown that a model with an ambiguity averse representative investor leads to higher risk premia and lower interest rates, thus providing a (partial) explanation of the equity premium puzzle. Our model could be used in a similar way. If it were the case that empirically observed investment timing in oligopolistic industries requires implausibly large coefficients of risk aversion, then ambiguity aversion may explain the observed behavior.

Appendix

A. Proof of Lemma 2

In this section, we show that if the worst–case for the leader value is not always given by the worst possible trend, there exists a unique value x_i^* at which the worst–case drift changes from $\underline{\mu}_i$ to $\overline{\mu}_i$. As before, we drop subscripts wherever possible.

Proof. The critical value x_i^* is found by applying the smooth–pasting condition $\hat{L}_i(x_i^*;\overline{\mu}_i) = 0$. The first derivative of $\hat{L}_i$ is given by

$$\hat{L}_i'(x;\overline{\mu}_i) = \frac{D_{10}}{r-\overline{\mu}_i} + \frac{\beta_1(\overline{\mu}_i)(x_i^*)^{\beta_2(\overline{\mu}_i)}x^{\beta_1(\overline{\mu}_i)-1} - \beta_2(\overline{\mu}_i)(x_i^*)^{\beta_1(\overline{\mu}_i)}x^{\beta_2(\overline{\mu}_i)-1}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F$$
$$+ \frac{\beta_2(\overline{\mu}_i)(x_j^F)^{\beta_1(\overline{\mu}_i)}x^{\beta_2(\overline{\mu}_i)-1} - \beta_1(\overline{\mu}_i)(x_j^F)^{\beta_2(\overline{\mu}_i)}x^{\beta_1(\overline{\mu}_i)-1}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^*.$$

In order to prove the existence of x_i^*, we will show that if $x_i^* \uparrow x_j^F$, $\hat{L}_i'(x_i^*;\overline{\mu}_i)$ becomes negative, and if $x_i^* \downarrow 0$, $\hat{L}_i'(x_i^*;\overline{\mu}_i)$ becomes positive.

We have

$$\hat{L}_i'(x_i^*;\overline{\mu}_i) = \frac{D_{10}}{r-\overline{\mu}_i} + \frac{(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i))(x_i^*)^{\beta_1(\overline{\mu}_i)+\beta_2(\overline{\mu}_i)-1}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}} \left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right) x_j^F$$
$$+ \frac{\beta_2(\overline{\mu}_i)(x_j^F)^{\beta_1(\overline{\mu}_i)}(x_i^*)^{\beta_2(\overline{\mu}_i)} - \beta_1(\overline{\mu}_i)(x_j^F)^{\beta_2(\overline{\mu}_i)}(x_i^*)^{\beta_1(\overline{\mu}_i)}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}} \left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right].$$

Clearly, $\lim_{x_i^* \downarrow x_j^F} \hat{L}_i'(x_i^*,\overline{\mu}_i)$ has the same sign as

$$\frac{D_{10}}{r-\overline{\mu}_i}\left((x_j^F)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_j^F)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}\right)$$
$$+ \left(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)\right)(x_j^F)^{\beta_1(\overline{\mu}_i)+\beta_2(\overline{\mu}_i)}\left[\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i} - \left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} + \frac{D_{10}}{r-\overline{\mu}_i}\right]. \tag{23}$$

Using the fact that $\frac{1}{\beta_1(\underline{\mu}_i)} < \frac{D_{10}-D11}{D_{10}}$ yields that (23) is smaller than

$$\left(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)\right)(x_j^F)^{\beta_1(\overline{\mu}_i)+\beta_2(\overline{\mu}_i)}\frac{1}{r-\underline{\mu}_i}\left(D_{11}-D_{10}+D_{10}-D_{11}\right) = 0. \tag{24}$$

Considering the case $x_i^* \downarrow 0$, one can easily see that $\lim_{x_i^* \downarrow 0} \hat{L}_i'(x_i^*;\overline{\mu}_i)$ has the same sign as

$$\lim_{x_i^*\downarrow 0}\left\{\frac{D_{10}}{r-\overline{\mu}_i}\left((x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}\right)\right.$$
$$+ \left(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)\right)(x_i^*)^{\beta_1(\overline{\mu}_i)+\beta_2(\overline{\mu}_i)-1}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F$$
$$\left. + \beta_2(\overline{\mu}_i)(x_j^F)^{\beta_1(\overline{\mu}_i)}(x_i^*)^{\beta_2(\overline{\mu}_i)} - \beta_1(\overline{\mu}_i)(x_j^F)^{\beta_2(\overline{\mu}_i)}(x_i^*)^{\beta_1(\overline{\mu}_i)}\left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)\right\}$$
$$= \lim_{x_i^*\downarrow 0}\left\{(x_i^*)^{\beta_2(\overline{\mu}_i)}\left(\frac{D_{10}}{r-\overline{\mu}_i}\left((x_j^F)^{\beta_2(\overline{\mu}_i)} - (x_i^*)^{\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)}\right)\right.\right.$$
$$+ \left(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)\right)(x_i^*)^{\beta_1(\overline{\mu}_i)-1}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F$$
$$\left.\left. + \left(\beta_2(\overline{\mu}_i)(x_j^F)^{\beta_1(\overline{\mu}_i)} - \beta_1(\overline{\mu}_i)(x_i^*)^{\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}\right)\left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)\right)\right\}$$
$$= \lim_{x_i^*\downarrow 0}\underbrace{(x_i^*)^{\beta_2(\overline{\mu}_i)}}_{\to+\infty}\left\{\underbrace{\frac{D_{10}}{r-\overline{\mu}_i}}_{>0}\left((x_j^F)^{\beta_2(\overline{\mu}_i)} - \underbrace{(x_i^*)^{\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)}}_{\to 0}\right)\right.$$
$$+ \underbrace{\left(\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)\right)(x_i^*)^{\beta_1(\overline{\mu}_i)-1}}_{\to 0}\left(\frac{D_{11}}{r-\underline{\mu}_i} - \frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F$$

$$+\left(\underbrace{\beta_2(\overline{\mu}_i)(x_j^F)^{\beta_1(\overline{\mu}_i)}}_{<0}-\underbrace{\beta_1(\overline{\mu}_i)(x_i^*)^{\beta_1(\overline{\mu}_i)-\beta_2(\overline{\mu}_i)}}_{\to 0}\right)\underbrace{\left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)}_{<0}\Bigg\}.$$

Therefore, we get $\hat{L}_i'(x_i^*;\overline{\mu}_i)>0$ for x_i^* close to 0. Due to continuity of L_j' on $[0,x_j^F]$, we can find in that interval a solution to $\hat{L}_i'(x_i^*;\overline{\mu}_i)=0$.

The uniqueness of x_i^* is automatically given by the uniqueness of the solution to PDE (15).

■

B. Concavity of L_i

In this section we prove that the leader value function is concave on $[0,x_j^F]$. In case the worst–case prior is always induced by the lowest possible trend, this statement is trivial. The next proof shows that concavity is not lost even if the worst–case changes at some point.

Proof. Suppose condition (9) is not satisfied (i.e. $\mathsf{P}^*_{L,i}\neq\mathsf{P}^{-\kappa_i}$). The concavity of $L_i(x)$ for $x<x_i^*$ is trivial. We therefore consider the second derivative of L_i on the interval $[x_i^*,x_j^F)$:

$$\begin{aligned}\hat{L}_i''(x;\overline{\mu}_i)=&\frac{\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)(x_i^*)^{\beta_2(\overline{\mu}_i)}x^{\beta_1(\overline{\mu}_i)-2}-\beta_2(\overline{\mu}_i)(\beta_2(\overline{\mu}_i)-1)(x_i^*)^{\beta_1(\overline{\mu}_i)}x^{\beta_2(\overline{\mu}_i)-2}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)}-(x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\\&\times\left(\frac{D_{11}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F\\&+\frac{\beta_2(\overline{\mu}_i)(\beta_2(\overline{\mu}_i)-1)(x_j^F)^{\beta_1(\overline{\mu}_i)}x^{\beta_2(\overline{\mu}_i)-2}-\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)(x_j^F)^{\beta_2(\overline{\mu}_i)}x^{\beta_1(\overline{\mu}_i)-2}}{(x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)^{\beta_1(\overline{\mu}_i)}-(x_i^*)^{\beta_1(\overline{\mu}_i)}(x_j^F)^{\beta_2(\overline{\mu}_i)}}\\&\times\left[\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right]x_i^*.\end{aligned}$$

Now, we have

$$\begin{aligned}&\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)x^{\beta_1(\overline{\mu}_i)-2}\Bigg[\left(\frac{D_{11}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F(x_i^*)^{\beta_2(\overline{\mu}_i)}\\&\qquad-\left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)x_i^*(x_j^F)^{\beta_2(\overline{\mu}_i)}\Bigg]\\&<\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)x^{\beta_1(\overline{\mu}_i)-2}x_i^*(x_j^F)^{\beta_2(\overline{\mu}_i)}\left[\left(\frac{D_{11}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)-\left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)\right]\\&=\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)x^{\beta_1(\overline{\mu}_i)-2}x_i^*(x_j^F)^{\beta_2(\overline{\mu}_i)}\left[\frac{D_{11}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\underline{\mu}_i}+\frac{1}{\beta_1(\underline{\mu}_i)}\frac{D_{10}}{r-\underline{\mu}_i}\right]\\&<\beta_1(\overline{\mu}_i)(\beta_1(\overline{\mu}_i)-1)x^{\beta_1(\overline{\mu}_i)-2}x_i^*(x_j^F)^{\beta_2(\overline{\mu}_i)}\frac{1}{r-\underline{\mu}_i}[D_{11}-D_{10}+D_{10}-D_{11}]\\&=0,\end{aligned}$$

where we used the fact that $x_i^*(x_j^F)^{\beta_2(\overline{\mu}_i)} < (x_i^*)^{\beta_2(\overline{\mu}_i)}(x_j^F)$ (because $x_i^* < x_j^F$ and $\beta_2(\overline{\mu}_i) < 0$) and $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu}_i)}$.

In a similar, way we can show that

$$\beta_2(\overline{\mu}_i)(\beta_2(\overline{\mu}_i)-1)x^{\beta_2(\overline{\mu}_i)-2}\Bigg[-\left(\frac{D_{11}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)x_j^F(x_i^*)^{\beta_1(\overline{\mu}_i)} + \left(\left(1-\frac{1}{\beta_1(\underline{\mu}_i)}\right)\frac{D_{10}}{r-\underline{\mu}_i}-\frac{D_{10}}{r-\overline{\mu}_i}\right)x_i^*(x_j^F)^{\beta_1(\overline{\mu}_i)}\Bigg] < 0,$$

which proves the concavity of L_i.

■

C. Proof of Proposition 1

The proof follows along similar lines to the proof of Theorem 1. We use the same procedure, but now we consider the value function in the continuation region, i.e. before any investment has taken place. Applying the BSDE approach with different value–matching and smooth–pasting conditions eventually yields the desired stopping time.

Proof. Denote

$$Y_t = \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\Big[\int_t^{\tau_{L,1}^t} e^{-r(s-t)}D_{00}X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)}(D_{10}X_s - rI_1)ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)}(D_{11}X_s - rI_1)ds\Big|\mathscr{F}_t\Big].$$

Using the time–consistency property of a strongly–rectangular set of density generators yields

$$\begin{aligned}
Y_t &= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\Big[\int_t^{\tau_{L,1}^t} e^{-r(s-t)}D_{00}X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)}(D_{10}X_s - rI_1)ds \\
&\qquad + \int_{\tau_2^F}^{\infty} e^{-r(s-t)}(D_{11}X_s - rI_1)ds\Big|\mathscr{F}_t\Big] \\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\left[\inf_{\mathsf{Q}'\in\mathscr{P}^{\Theta}} \mathsf{E}^{\mathsf{Q}'}\left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)}D_{00}X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)}(D_{10}X_s - rI_1)ds\right.\right. \\
&\qquad \left.\left. + \int_{\tau_2^F}^{\infty} e^{-r(s-t)}(D_{11}X_s - rI_1)ds\Big|\mathscr{F}_{\tau_{L,1}^t}\right]\Big|\mathscr{F}_t\right] \\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)}D_{00}X_s ds + e^{-r(\tau_{L,1}^t-t)}\inf_{\mathsf{Q}'\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}'}\left[\int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-\tau_{L,1}^t)}(D_{10}X_s - rI_1)ds\right.\right. \\
&\qquad \left.\left. + \int_{\tau_2^F}^{\infty} e^{-r(s-\tau_{L,1}^t)}(D_{11}X_s - rI_1)ds\Big|\mathscr{F}_{\tau_{L,1}^t}\right]\Big|\mathscr{F}_t\right] \\
&= \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_1}} \mathsf{E}^{\mathsf{Q}}\left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)}D_{00}X_s ds + e^{-r(\tau_{L,1}^t-t)}L_1(x_{\tau_{L,1}^t})\Big|\mathscr{F}_t\right].
\end{aligned}$$

Following Chen and Epstein (2002), Y_t solves the BSDE

$$-dY_t = g_1(Z_t)dt - Z_t dB_t,$$

where the generator is given by

$$g_1(z) = -\kappa_1|z| - rY_t + X_t D_{00}.$$

The boundary condition is given by

$$Y_{\tau^t_{L,1}} = L_1(x_1^L),$$

where $L_1(x_1^L)$ is given by Theorem 1 and $x_1^L = x_{\tau^t_{L,1}}$.

Denote the present value of the leader payoff by Λ, i.e.

$$\Lambda(x_t) = Y_t.$$

The non–linear Feynman–Kac formula implies that Λ solves the non–linear PDE

$$\mathscr{L}_X \Lambda(x) + g_1(\sigma x \Lambda'(x)) = 0.$$

Hence, Λ solves

$$\frac{1}{2}\sigma^2 x^2 \Lambda''(x) + \mu x \Lambda'(x) - \kappa_1 \sigma x \left|\Lambda'(x)\right| - r\Lambda(x) + D_{00}x = 0. \tag{25}$$

In the continuation region the leader function has to be increasing, hence $\Lambda' > 0$. This implies that $\underline{\mu}_1$ is the worst–case in the continuation region.

Therefore, equation (25) becomes

$$\frac{1}{2}\sigma^2 x^2 \Lambda''(x) + (\mu - \kappa_1\sigma) x \Lambda'(x) - r\Lambda(x) + D_{00}x = \frac{1}{2}\sigma^2 x^2 \Lambda''(x) + \underline{\mu}_1 x \Lambda'(x) - r\Lambda(x) + D_{00}x = 0.$$

The general increasing solution to this PDE is

$$\Lambda(x) = \frac{D_{00}x}{r - \underline{\mu}_1} + A_2 x^{\beta_1(\underline{\mu}_1)}.$$

We have to distinguish two cases here. Either the condition given in Theorem 1 holds which means that the boundary condition takes the form (10) or the boundary condition becomes (11).

We will show that for both cases, the optimal threshold to invest becomes

$$x_1^L = \frac{\beta_1(\underline{\mu}_1)}{\beta_1(\underline{\mu}_1) - 1} \frac{I_1(r - \underline{\mu}_1)}{D_{10} - D_{00}}. \tag{26}$$

If condition (9) is satisfied, the boundary condition is given by

$$L_1(x_1^L) = \frac{D_{10}x_1^L}{r - \underline{\mu}_1} + \left(\frac{x_1^L}{x_2^F}\right)^{\beta_1(\underline{\mu}_1)} \frac{D_{11} - D_{10}}{r - \underline{\mu}_1} x_2^F - I_1.$$

Otherwise, the boundary condition is given by

$$L_1(x_1^L) = \frac{D_{10}x_1^L}{r-\underline{\mu}_1} - \frac{1}{\beta_1(\underline{\mu}_1)}\frac{D_{10}x_1^*}{r-\underline{\mu}_1}\left(\frac{x_1^L}{x_1^*}\right)^{\beta_1(\underline{\mu}_1)} - I_1.$$

In addition to the value–matching condition, we apply a smooth–pasting condition. Here, smooth pasting implies that the derivatives of the value function Λ and L coincide at $x_{\tau^t_{L,1}}$, i.e.

$$\Lambda'(x_{\tau^t_{L,1}}) = L_1'(x_{\tau^t_{L,1}}). \tag{27}$$

This condition ensures differentiability at the investment threshold.

Applying condition (27) gives

$$\frac{D_{00}}{r-\underline{\mu}_1} + \beta_1(\underline{\mu}_1)A_2 {x_1^L}^{\beta_1(\underline{\mu}_1)-1} = \frac{D_{10}}{r-\underline{\mu}_1} + \beta_1(\underline{\mu}_1)A_1 {x_1^L}^{\beta_1(\underline{\mu}_1)-1},$$

where

$$A_1 = \left(\frac{1}{x_2^F}\right)^{\beta_1(\underline{\mu}_1)-1}\frac{D_{11}-D_{10}}{r-\underline{\mu}_1}$$

in the first case and

$$A_1 = -\frac{1}{\beta_1(\underline{\mu}_1)}\frac{D_{10}x_1^*}{r-\underline{\mu}_1}\left(\frac{1}{x_1^*}\right)^{\beta_1(\underline{\mu}_1)}$$

in the second.

Hence,

$$A_2 = \frac{D_{10}-D_{00}}{r-\underline{\mu}_1}\frac{1}{\beta_1(\underline{\mu}_1)}\frac{1}{{x_1^L}^{\beta_1(\underline{\mu}_1)-1}} + A_1.$$

Applying the value–matching condition finally yields

$$\begin{aligned}
\frac{D_{00}x_1^L}{r-\underline{\mu}_1} + \left(\frac{D_{10}-D_{00}}{r-\underline{\mu}_1}\frac{1}{\beta_1(\underline{\mu}_1)}\frac{1}{{x_1^L}^{\beta_1(\underline{\mu}_1)-1}} + A_1\right){x_1^L}^{\beta_1(\underline{\mu}_1)} &= \frac{D_{10}x_1^L}{r-\underline{\mu}_1} + A_1 {x_1^L}^{\beta_1(\underline{\mu}_1)} - I_1 \\
\Longleftrightarrow \frac{D_{10}-D_{00}}{r-\underline{\mu}_1}x_1^L - \frac{D_{10}-D_{00}}{r-\underline{\mu}_1}\frac{1}{\beta_1(\underline{\mu}_1)}x_1^L &= I_1 \\
\Longleftrightarrow \frac{\beta_1(\underline{\mu}_1)-1}{\beta_1(\underline{\mu}_1)}\frac{D_{10}-D_{00}}{r-\underline{\mu}_1}x_1^L &= I_1,
\end{aligned}$$

and therefore, for both cases, it holds that

$$x_1^L = \frac{\beta_1(\underline{\mu}_1)}{\beta_1(\underline{\mu}_1)-1}\frac{r-\underline{\mu}_1}{D_{10}-D_{00}}I_1.$$

■

D. Equilibrium Concept

The appropriate notion of subgame–perfect equilibrium for our game is developed in Riedel and Steg (2017). In this section we adapt their equilibrium concept to make it applicable to games with ambiguous players.

Let $\mathscr{T}$ denote the set of stopping times with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$. The set $\mathscr{T}$ will act as the set of (pure) strategies. Given the definitions of the leader, follower and simultaneous investment payoffs above, the timing game is

$$\Gamma = \left\langle (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathsf{P}), (\mathscr{P}^{\Theta_i})_{i=1,2}, \mathscr{T} \times \mathscr{T}, (L_i, F_i, M_i)_{i=1,2}, (\pi_i)_{i=1,2} \right\rangle,$$

where, for $(\tau_1, \tau_2) \in \mathscr{T} \times \mathscr{T}$,

$$\pi_i(x_0) = \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}} \left[L_i(x_0)\mathbf{1}_{\tau_i<\tau_j} + F_i(x_0)\mathbf{1}_{\tau_i>\tau_j} + M_i(x_0)\mathbf{1}_{\tau_i=\tau_j} \right].$$

The subgame starting at stopping time $\vartheta \in \mathscr{T}$ is the tuple

$$\Gamma^{\vartheta} = \left\langle (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq \vartheta}, \mathsf{P}), (\mathscr{P}^{\Theta_i})_{i=1,2}, \mathscr{T}_{\vartheta} \times \mathscr{T}_{\vartheta}, (L_i, F_i, M_i)_{i=1,2}, (\pi_i^{\vartheta})_{i=1,2} \right\rangle,$$

where $\mathscr{T}_{\vartheta}$ is the set of stopping times no smaller than ϑ a.s., i.e.

$$\mathscr{T}_{\vartheta} := \{\tau \in \mathscr{T} | \tau \geq \vartheta, \mathsf{P} - a.s.\},$$

and, for $(\tau_1, \tau_2) \in \mathscr{T}_{\vartheta} \times \mathscr{T}_{\vartheta}$,

$$\pi_i^{\vartheta}(x_{\vartheta}) = \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}} \left[L_i(x_{\vartheta})\mathbf{1}_{\tau_i<\tau_j} + F_i(x_{\vartheta})\mathbf{1}_{\tau_i>\tau_j} + M_1(x_{\vartheta})\mathbf{1}_{\tau_i=\tau_j} | \mathscr{F}_{\vartheta} \right].$$

As is argued in Riedel and Steg (2017), careful consideration has to be given to the appropriate notion of strategy. They show that the notion of *extended mixed strategy* is versatile and intuitively appealing. For the subgame Γ^{ϑ} this is a pair of processes $(G^{\vartheta}, \alpha^{\vartheta})$, both taking values in $[0,1]$, with the following properties.

1. G^{ϑ} is adapted, has right–continuous and non–decreasing sample paths, with $G^{\vartheta}(s) = 0$ for all $s < \vartheta$, P–a.s.
2. α^{ϑ} is progressively measurable with right–continuous sample paths whenever its value is in $(0,1)$, P–a.s.
3. On $\{t \geq \vartheta\}$, it holds that $\alpha^{\vartheta}(t) > 0$ implies $G^{\vartheta}(t) = 1$, P–a.s.

Note that the properties above hold for all $\mathsf{Q} \in \mathscr{P}^{\Theta_i}$, $i = 1,2$, if they hold for P, because all measures in $\mathscr{P}^{\Theta_i}$ are equivalent. We use the convention that

$$G^{\vartheta}(0-) \equiv 0, \quad G^{\vartheta}(\infty) \equiv 1, \quad \text{and} \quad \alpha^{\vartheta}(\infty) \equiv 1.$$

For our purposes extended mixed strategies are, in fact, more general than necessary. Therefore, we will restrict attention to what we will call *extended pure strategies*. For the subgame Γ^{ϑ} this is a pair of extended mixed strategies $(G_i^\vartheta, \alpha_i^\vartheta)_{i=1,2}$, where G_i^ϑ is restricted to take values in $\{0,1\}$. In other words, in an extended pure strategy a firm does not mix over stopping times, but potentially mixes over its "investment intensity" α^ϑ.

An extended pure strategy for the game Γ is then a collection $(G^\vartheta, \alpha^\vartheta)_{\vartheta \in \mathscr{T}}$ of extended pure strategies in subgames Γ^ϑ, $\vartheta \in \mathscr{T}$ satisfying the time consistency conditions that for all $\vartheta, \nu \in \mathscr{T}$ it holds that

1. $\nu \leq t \in \mathbb{R}_+$ implies $G^\vartheta(t) = G^\vartheta(\nu-) + (1 - G^\vartheta(\nu-))G^\nu(t)$, P–a.s. on $\{\vartheta \leq \nu\}$,
2. $\alpha^\vartheta(\tau) = \alpha^\nu(\tau)$, P–a.s., for all $\tau \in \mathscr{T}$.

The importance of the α component in the definition of extended pure strategy becomes obvious in the definition of payoffs. Essentially α allows both for immediate investment and coordination between firms. It leads to investment probabilities that can be thought of as the limits of conditional stage investment probabilities of discrete–time behavioral strategies with vanishing period length (cf. Fudenberg and Tirole (1985)). In the remainder, let $\hat{\tau}_i^\vartheta$ be the first time that α_i^ϑ is strictly positive, and let $\hat{\tau}^\vartheta$ be the first time that at least one α^ϑ is non–zero in the subgame Γ^ϑ, i.e.

$$\hat{\tau}_i^\vartheta = \inf\{t \geq \vartheta | \alpha_i^\vartheta(t) > 0\}, \quad \text{and} \quad \hat{\tau}^\vartheta = \inf\{t \geq \vartheta | \alpha_1^\vartheta(t) + \alpha_2^\vartheta(t) > 0\},$$

respectively. At time $\hat{\tau}^\vartheta$ the extended pure strategies induce a probability measure on the state space

$$\Lambda = \{\, \{\text{Firm 1 becomes the leader}\}, \{\text{Firm 2 becomes the leader}\}, \{\text{Both firms invest simultaneously}\}\,\},$$

for which we will use the shorthand notation

$$\Lambda = \{\, (L,1), (L,2), M \,\}.$$

Riedel and Steg (2017) show that the probability measure on Λ, induced by the pair $(\alpha_1^\vartheta, \alpha_1^\vartheta)$, is given by

$$\lambda_{L,i}^\vartheta(\hat{\tau}^\vartheta) = \begin{cases} \frac{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta(1-\alpha_{j,\hat{\tau}^\vartheta}^\vartheta)}{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta + \alpha_{j,\hat{\tau}^\vartheta}^\vartheta - \alpha_{i,\hat{\tau}^\vartheta}^\vartheta \alpha_{j,\hat{\tau}^\vartheta}^\vartheta} & \text{if } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta \text{ and } \alpha_i^\vartheta(\hat{\tau}_i^\vartheta), \alpha_j^\vartheta(\hat{\tau}_i^\vartheta) > 0 \\ 1 & \text{if } \hat{\tau}_i^\vartheta < \hat{\tau}_j^\vartheta, \text{ or } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta \text{ and } \alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0 \\ 0 & \text{if } \hat{\tau}_i^\vartheta > \hat{\tau}_j^\vartheta, \text{ or } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta \text{ and } \alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0 \\ \frac{1}{2}\Big(\liminf_{t \downarrow \hat{\tau}_i^\vartheta} \frac{\alpha_i^\vartheta(t)(1-\alpha_j^\vartheta(t))}{\alpha_i^\vartheta(t)+\alpha_j^\vartheta(t)-\alpha_i^\vartheta(t)\alpha_j^\vartheta(t)} & \text{if } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta,\ \alpha_i^\vartheta(\hat{\tau}_i^\vartheta) = \alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0, \\ +\limsup_{t \downarrow \hat{\tau}_i^\vartheta} \frac{\alpha_i^\vartheta(t)(1-\alpha_j^\vartheta(t))}{\alpha_i^\vartheta(t)+\alpha_j^\vartheta(t)-\alpha_i^\vartheta(t)\alpha_j^\vartheta(t)}\Big) & \text{and } \alpha_i^\vartheta(\hat{\tau}_i^\vartheta+), \alpha_j^\vartheta(\hat{\tau}_j^\vartheta+) > 0, \end{cases}$$

and

$$\lambda_M^\vartheta(\hat{\tau}^\vartheta) = \begin{cases} 0 & \text{if } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta,\ \alpha_i^\vartheta(\hat{\tau}_i^\vartheta) = \alpha_j^\vartheta(\hat{\tau}_i^\vartheta) = 0, \text{ and } \alpha_i^\vartheta(\hat{\tau}_i^\vartheta+), \alpha_j^\vartheta(\hat{\tau}_i^\vartheta+) > 0 \\ \frac{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta \alpha_{j,\hat{\tau}^\vartheta}^\vartheta}{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta + \alpha_{j,\hat{\tau}^\vartheta}^\vartheta - \alpha_{i,\hat{\tau}^\vartheta}^\vartheta \alpha_{j,\hat{\tau}^\vartheta}^\vartheta} & \text{otherwise.} \end{cases}$$

Note the following:

1. If $\hat{\tau}_i^\vartheta < \hat{\tau}_j^\vartheta$ there is no coordination problem: firm i becomes the leader λ–a.s. at $\hat{\tau}_i^\vartheta$;

2. If $\hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta$, but $\alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0$, there is no coordination problem: firm i becomes the leader λ–a.s. at $\hat{\tau}_i^\vartheta$;

3. In the degenerate case where $\hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta$, $\alpha_i^\vartheta(\hat{\tau}_i^\vartheta) = \alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0$, and $\alpha_i^\vartheta(\hat{\tau}_i^\vartheta+), \alpha_j^\vartheta(\hat{\tau}_j^\vartheta+) > 0$, the leader role is assigned at time $\hat{\tau}_i^\vartheta$, effectively on the basis of the flip of a fair coin;

4. Firms are not ambiguous over the measure λ.

In order to derive the payoffs to firms, let $\tau_{G,i}^\vartheta$ denote the first time that G_i^ϑ jumps to one, i.e.

$$\tau_{G,i}^\vartheta = \inf \left\{ t \geq \vartheta \,\middle|\, G_i^\vartheta(t) > 0 \right\}.$$

The payoff to firm i of a pair of extended pure strategies $((G_1, \alpha_1), (G_2, \alpha_2))$ in the subgame Γ^ϑ is given by

$$\begin{aligned}
V_i^\vartheta\left((G_i^\vartheta, \alpha_i^\vartheta), (G_j^\vartheta, \alpha_j^\vartheta)\right) := \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\tau_{G,i}^\vartheta < \min\{\tau_{G,j}^\vartheta, \hat{\tau}^\vartheta\}}\Bigg(\int_\vartheta^{\tau_{G,i}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds + \int_{\tau_{G,i}^\vartheta}^{\tau_j^F} e^{-r(s-\vartheta)} D_{10} X_s ds \\
&+ \int_{\tau_j^F}^\infty e^{-r(s-\vartheta)} D_{11} X_s ds - e^{-r(\tau_{G,i}^\vartheta - \vartheta)} I_i\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg] \\
+ \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\tau_{G,j}^\vartheta < \min\{\tau_{G,i}^\vartheta, \hat{\tau}^\vartheta\}}\Bigg(\int_\vartheta^{\tau_{G,j}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds + \int_{\tau_{G,j}^\vartheta}^{\tau_i^F} e^{-r(s-\vartheta)} D_{01} X_s ds \\
&+ \int_{\tau_i^F}^\infty e^{-r(s-\vartheta)} D_{11} X_s - e^{-r(\tau_i^F - \vartheta)} I_i\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg] \\
+ \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\tau_{G,i}^\vartheta = \tau_{G,j}^\vartheta < \hat{\tau}^\vartheta}\Bigg(\int_\vartheta^{\tau_{G,i}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds \\
&+ \int_{\tau_{G,i}^\vartheta}^\infty e^{-r(s-\vartheta)} D_{11} X_s ds\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg] \\
+ \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\hat{\tau}^\vartheta \leq \min\{\tau_{G,i}^\vartheta, \tau_{G,i}^\vartheta\}} \lambda_{L,i}^\vartheta(\hat{\tau}^\vartheta)\Bigg(\int_\vartheta^{\hat{\tau}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds \\
&+ \int_{\hat{\tau}^\vartheta}^{\tau_j^F} e^{-r(s-\vartheta)} D_{10} X_s ds + \int_{\tau_j^F}^\infty e^{-r(s-\vartheta)} D_{11} X_s ds - e^{-r(\tau_{G,i}^\vartheta - \vartheta)} I_i\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg] \\
+ \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\hat{\tau}^\vartheta \leq \min\{\tau_{G,i}^\vartheta, \tau_{G,i}^\vartheta\}} \lambda_{L,j}^\vartheta(\hat{\tau}^\vartheta)\Bigg(\int_\vartheta^{\hat{\tau}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds \\
&+ \int_{\hat{\tau}^\vartheta}^{\tau_i^F} e^{-r(s-\vartheta)} D_{01} X_s ds + \int_{\tau_i^F}^\infty e^{-r(s-\vartheta)} D_{11} X_s - e^{-r(\tau_i^F - \vartheta)} I_i\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg] \\
+ \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\Bigg[&\mathbf{1}_{\hat{\tau}^\vartheta \leq \min\{\tau_{G,i}^\vartheta, \tau_{G,i}^\vartheta\}} \lambda_M^\vartheta(\hat{\tau}^\vartheta)\Bigg(\int_\vartheta^{\hat{\tau}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds \\
&+ \int_{\hat{\tau}^\vartheta}^\infty e^{-r(s-\vartheta)} D_{11} X_s ds\Bigg)\Bigg|\mathscr{F}_\vartheta\Bigg].
\end{aligned}$$

Hence, the payoff of firm i can written as

$$
\begin{aligned}
V_i^\vartheta\left((G_i^\vartheta,\alpha_i^\vartheta),(G_j^\vartheta,\alpha_j^\vartheta)\right) := & \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\tau_{G,i}^\vartheta<\min\{\tau_{G,j}^\vartheta,\hat{\tau}^\vartheta\}} L_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right] \\
& + \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\tau_{G,j}^\vartheta<\min\{\tau_{G,i}^\vartheta,\hat{\tau}^\vartheta\}} F_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right] \\
& + \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\tau_{G,i}^\vartheta=\tau_{G,j}^\vartheta<\hat{\tau}^\vartheta} M_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right] \\
& + \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\hat{\tau}^\vartheta\le\min\{\tau_{G,i}^\vartheta,\tau_{G,i}^\vartheta\}} \lambda_{L,i}^\vartheta(\hat{\tau}^\vartheta) L_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right] \\
& + \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\hat{\tau}^\vartheta\le\min\{\tau_{G,i}^\vartheta,\tau_{G,i}^\vartheta\}} \lambda_{L,j}^\vartheta(\hat{\tau}^\vartheta) F_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right] \\
& + \inf_{\mathsf{Q}\in\mathscr{P}^{\Theta_i}} \mathsf{E}^{\mathsf{Q}}\left[\mathbf{1}_{\hat{\tau}^\vartheta\le\min\{\tau_{G,i}^\vartheta,\tau_{G,i}^\vartheta\}} \lambda_{M}^\vartheta(\hat{\tau}^\vartheta) M_i(x_\vartheta)\middle|\mathscr{F}_\vartheta\right].
\end{aligned}
$$

An equilibrium for the subgame Γ^ϑ is a pair of extended pure strategies $\left((\bar{G}_1^\vartheta,\bar{\alpha}_1^\vartheta),(\bar{G}_2^\vartheta,\bar{\alpha}_2^\vartheta)\right)$, such that for each firm $i=1,2$ and every extended pure strategy $(G_i^\vartheta,\alpha_i^\vartheta)$ it holds that

$$V_i^\vartheta(\bar{G}_i^\vartheta,\bar{\alpha}_i^\vartheta,\bar{G}_j^\vartheta,\bar{\alpha}_j^\vartheta) \ge V_i^\vartheta(G_i^\vartheta,\alpha_i^\vartheta,\bar{G}_j^\vartheta,\bar{\alpha}_j^\vartheta),$$

for $j\neq i$. A subgame perfect equilibrium is a pair of extended pure strategies $\left((\bar{G}_1,\bar{\alpha}_1),(\bar{G}_2,\bar{\alpha}_2)\right)$, such that for each $\vartheta\in\mathscr{T}$ the pair $\left((\bar{G}_1^\vartheta,\bar{\alpha}_1^\vartheta),(\bar{G}_2^\vartheta,\bar{\alpha}_2^\vartheta)\right)$ is an equilibrium in the subgame Γ^ϑ.

For the equilibrium analysis in this paper it is important to know where $L_i > F_i$. In a model with $\mathscr{P}^{\Theta_1} = \mathscr{P}^{\Theta_2}$ and $I_1 < I_2$ we have that (cf. Proposition 2)

$$L_1 - F_1 > L_2 - F_2, \quad \text{on } (0, x_2^F).$$

The following equilibrium results now follow directly from Riedel and Steg (2017) and Steg (2015). The collection

$$\left((G_1^\vartheta,\alpha_1^\vartheta),(G_2^\vartheta,\alpha_2^\vartheta)\right)_{\vartheta\in\mathscr{T}},$$

constitutes a subgame perfect equilibrium, where

1. if $x_1^L \in (\underline{x},\bar{x})$, it holds that

$$
\begin{aligned}
G_1^\vartheta &= \mathbf{1}_{\{t: x_t \ge x\}} \\
\alpha_1^\vartheta(t) &= \begin{cases} 1 & \text{if } x_t \in \{\underline{x}\}\cup[\bar{x},\infty) \\ \frac{L_2(x_t)-F_2(x_t)}{L_2(x_t)-M_2(x_t)} & \text{if } x_t\in(\underline{x},\bar{x}) \\ 0 & \text{else,} \end{cases}
\end{aligned}
$$

and

$$G_2^\vartheta = \mathbf{1}_{\{t:x_t\in(\underline{x},\bar{x})\cup[x_2^F,\infty)\}}$$

$$\alpha_2^\vartheta(t) = \begin{cases} 1 & \text{if } x_t \geq x_2^F \\ \frac{L_1(x_t)-F_1(x_t)}{L_1(x_t)-M_1(x_t)} & \text{if } x_t \in [\underline{x},\bar{x}] \\ 0 & \text{else;} \end{cases}$$

2. if $x_1^L < \underline{x}$, it holds that

$$G_1^\vartheta = \alpha_1^\vartheta = \mathbf{1}_{\{t:x_t\geq x_1^L\}},$$

and

$$G_2^\vartheta = \alpha_2^\vartheta = \mathbf{1}_{\{t:t\geq x_2^F\}};$$

3. if $x_1^L > \bar{x}$ and $\bar{x}(D_{10} - D_{00}) < rI_1$, it holds that

$$G_1^\vartheta = \mathbf{1}_{\{t:x_t\in[\underline{x},\bar{x}]\cup[\hat{x},\infty)\cup[x_2^F,\infty)\}}$$

$$\alpha_1^\vartheta(t) = \begin{cases} 1 & \text{if } x_t \in \{\underline{x},\bar{x}\} \cup [\hat{x},\infty) \cup [x_2^F,\infty) \\ \frac{L_2(x_t)-F_2(x_t)}{L_2(x_t)-M_2(x_t)} & \text{if } x_t \in (\underline{x},\bar{x}) \\ 0 & \text{else,} \end{cases}$$

and

$$G_2^\vartheta = \mathbf{1}_{\{t:x_t\in(\underline{x},\bar{x})\cup[x_2^F,\infty)\}}$$

$$\alpha_2^\vartheta(t) = \begin{cases} 1 & \text{if } x_t \geq x_2^F \\ \frac{L_1(x_t)-F_1(x_t)}{L_1(x_t)-M_1(x_t)} & \text{if } x_t \in [\underline{x},\bar{x}] \\ 0 & \text{else.} \end{cases}$$

Acknowledgments

Most of the research was conducted when Hellmann was at the Center for Mathematical Economics (IMW) at Bielefeld University, Germany. The authors gratefully acknowledge support from the Department of Economics & Related Studies at the University of York. Hellmann gratefully acknowledges financial support from the German Research Foundation (DFG) via grant RI 1128-4-4-2. Thijssen gratefully acknowledges support from the Center for Mathematical Economics (IMW) and the Center for Interdisciplinary

Research (ZiF) at Bielefeld University. Helpful comments were received from participants of the ZiF programm “Robust Finance: Strategic Power, Knightian Uncertainty, and the Foundations of Economic Policy Advice”, Mark Freeman, Philip Linsley, Simon Mollan, and seminar participants at the University of Lund. The authors would especially like to thank Frank Riedel and Jan–Henrik Steg for valuable comments and suggestions. We are grateful to the Area Editor, the Associate Editor and two anonymous referees for many helpful comments and suggestions over several rounds of revision.

References

Ball D, Watt J (2013) Further thoughts on the utility of risk matrices. *Risk Analysis* 33(11):2068–2078.

Bloom N (2009) The impact of uncertainty shocks. *Econometrica* 77(3):623–685.

Bloom N, Bond S, Reenen JV (2007) Uncertainty and investment dynamics. *Review of Economic Studies* 74(2):391–415.

Chen Z, Epstein L (2002) Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70(4):1403–1443.

Cheng X, Riedel F (2013) Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics* 7(1):29–68.

Décamps J, Mariotti T, Vileneuve S (2005) Investment timing under incomplete information. *Mathematics of Operations Research* 30(2):472–500.

Dixit AK, Pindyck RS (1994) *Investment under uncertainty* (Princeton University Press, Princeton, NJ).

Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics* 75(4):643–669.

Epstein LG, Schneider M (2003) Recursive multiple priors. *Journal of Economic Theory* 113(1):1–31.

Epstein LG, Schneider M (2010) Ambiguity and asset markets. *Annual Review of Financial Economics* 2:315–346.

Fudenberg D, Tirole J (1985) Preemption and rent equilization in the adoption of new technology. *Review of Economic Studies* 52(3):383–401.

Gilboa I (2009) *Theory of decision under uncertainty* (Cambridge University Press, New York).

Gilboa I, Schmeidler D (1989) Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18(2):141–153.

Graham J, Harvey C (2001) The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics* 60(2–3):187–243.

Gzyl H, ter Horst E, Molina G (2017) Inferring probability densities from expert opinion. *Applied Mathematical Modelling* 43:306–320.

Knight FH (1921) *Risk, Uncertainty and Profit* (Boston, MA: Houghton Mifflin Co).

Mehra R, Prescott E (1985) The equity premium: a puzzle. *Journal of Monetary Economics* 15:145–161.

Nishimura KG, Ozaki H (2007) Irreversible investment and Knightian uncertainty. *Journal of Economic Theory* 136(1):668–694.

Pawlina G, Kort PM (2006) Real options in an asymmetric duopoly: Who benefits from your competitive disadvantage? *Journal of Economics & Management Strategy* 15(1):1–35.

Peng S (1991) Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports* 37:61–74.

Peng S (1997) Backward SDE and related g–expectation. *Backward Stochastic Differential Equations*, 141–159, Pitman res. Notes Math. Ser., 364, Longman, Harlow.

Peng S (2013) Nonlinear expectation theory and stochastic calculus under knightian uncertainty. *Real Options, Ambiguity, Risk and Insurance*, volume 5 of *Studies in Probability, Optimization and Statistics*, 144–184 (IOS Press, Amsterdam: A. Bensoussan, S. Peng, and J. Sung).

Randall A (2011) *Risk and Precaution* (Cambridge University Press, Cambridge).

Riedel F, Steg JH (2017) Subgame–perfect equilibria in stochastic timing games. *Journal of Mathematical Economics* 72:36–50.

Steg JH (2015) Symmetric equilibria in stochastic timing games. *Working Paper 543, Center for Mathematical Economics, Bielefeld University* .

Steg JH, Thijssen JJ (2015) Quick or persistent? strategic investment demanding versatility. *Working Paper 541, Center for Mathematical Economics, Bielefeld University* .

Thijssen JJ (2010) Irreverisble investment and discounting: an arbitrage pricing approach. *Annals of Finance* 6:295–315.

Trojani F, Vanini P (2004) Robustness and ambiguity aversion in general equilibrium. *Review of Finance* 8:279–324.

Weeds H (2002) Strategic Delay in a Real Options Model of R&D Competition. *Review of Economic Studies* 69(3):729–747.