



WHU – OTTO BEISHEIM SCHOOL OF MANAGEMENT

PRODUCTION MANAGEMENT DEPARTMENT
BURGPLATZ 2, 56179 VALLENDAR, GERMANY

*PhD dissertation submitted in partial fulfilment of Doctor's Degree of Business
Administration (Dr. rer. pol.)*

Investment under Market and Strategic Uncertainty

by

Benoît CHEVALIER-ROIGNANT

Under the supervision of
Arnd HUCHZERMEIER, PhD and Lenos TRIGEORGIS, DBA.

July 11, 2011

A Viviane, Patrick, Aurélien et mes proches

Contents

Introduction	1
1. Motivation	1
2. A dynamic, continuous-time approach to investment	2
3. Organization of the monograph	4
4. Mathematical notations	6
5. Acknowledgement	6
Part 1. Theoretical Foundations	7
Chapter 1. Production, Capacity, and Investment	9
1. Investment frictions	9
2. Reduced-form profits	10
3. Capital adjustment decisions	12
4. References	14
Chapter 2. Stochastic Calculus	15
1. Randomness and stochastic processes	15
2. Stochastic integration and Itô-Doebelin formula	17
3. Itô processes	23
4. References	26
Chapter 3. Stochastic Control	27
1. A dynamic problem	27
2. Types of controls	28
3. Solution approaches	29
4. Other stochastic control problems	34
5. References	37
Chapter 4. Dynamic Game Theory	39
1. Game description	39
2. Information structure and strategy space	40
3. Solution concepts	41
4. Nonzero-sum differential games	42
5. Stochastic differential games	44
6. References	45
Part 2. Benchmark References	47
Chapter 5. Monopoly	49
1. Investment timing and hysteresis	49
2. Capital adjustment and marginal Tobin's q theory	51
3. References	57
Chapter 6. Social Optimality, Myopia, and Perfect Competition	59

1. Social optimum	59
2. Myopic Nash equilibrium	61
3. Perfect competition	62
4. Completion delays	64
5. References	65
Part 3. Imperfect Competition	67
Chapter 7. Strategic Investment under Uncertainty: A Synthesis	69
1. Introduction	69
2. Lumpy investment decisions: Entry, improvement, and exit	70
3. Incremental capacity expansion	77
4. Complex investment decisions under uncertainty and competition	79
5. Conclusion and outlook	81
Chapter 8. Market-Entry Sequencing under Uncertainty	85
1. Introduction	85
2. Model setup and assumptions	86
3. Market-entry sequencing	87
4. Social optimality of myopic market-entry sequencing	94
5. Conclusion	96
Appendices	96
Chapter 9. Preemptive Capacity Investment under Uncertainty	99
1. Introduction	99
2. Model primitives and assumptions	100
3. Strategy-space choice and solution approach	101
4. Markov perfect equilibrium	102
5. Conclusion	105
Summary and Outlook	107
1. Contributions to capital budgeting	107
2. Review of major results	108
3. Future Research Directions	108
Bibliography	111

Introduction

1. Motivation

A dynamic economic environment is characterized by uncertainty about microeconomic changes in the competitive landscape (strategic uncertainty) and exogenous macroeconomic developments (market uncertainty). Some European governments used to erect high (administrative) barriers shielding certain “natural monopolies” from competitive entry (e.g., telecommunications, electricity and gas utility sectors). Lately, these governments had to enforce in national laws directives of the European Union and implement deregulation schemes, opening up certain economic activities to new, potentially foreign, market participants. These moves have led to higher competitive intensity in the industries concerned. Following liberalization and deregulation of Western economies, only a limited number of industries have remained protected, whereas most companies wage fiercer competition in their respective economic sectors. At the same time, some other sectors traditionally characterized by a large number of companies have undergone significant consolidation, resulting in oligopoly situations with a reduced number of players. These two concurrent phenomena — liberalization and continuing consolidation — have increased the emphasis corporate managers should put on strategic uncertainty, the emergence of consolidated oligopolies (and the disappearance of monopolies) paving the way for higher degree of strategic interactions. Investment has been a focus of analysis in economics, finance, and operations management. They typically have long-term consequences and need to be analyzed by means of dynamic optimization techniques. Investment decisions are made at a stage where future market developments are uncertain. Owing to this, corporate managers have a higher incentive to properly assess market uncertainty when making a choice amongst strategic options. Strategy revision is both advised and necessary for an industry to sustain itself or evolve. Firms unable to adapt to actual developments, pursuing instead objectives that prove irrelevant over time, are doomed to fail.

At the heart of industry dynamics lies the investment challenge under uncertainty. Investing involves sacrificing current consumption in expectation of future benefits. The theory of *corporate finance* is meant to examine this basic tradeoff in the presence of partly unrecoverable investment costs and nonperfectly predictable future developments. The most commonly accepted theory in capital budgeting to assess, value, and compare investment alternatives is the net present value approach (NPV) that consists in discounting future expected cash flows at a discount rate accounting for risk adjustment. This view is based on the premises that the plan of managerial actions is deterministic and set once and for all. While market uncertainty is factored in through the risk-adjusted discount rate, the flexibility management has is typically ignored. In many cases, corporate managers might revise their decisions in view of the actual realization of market development (vs. its expected realization). The assumption of full commitment to a given action at certain future time is rigid and not realistic in most industries. If a manager invests today in a project, she is not necessarily committed to a given deterministic plan of actions, scaling up production in case of (unexpected) demand upsurge or scaling it down (or even closing the production plant) in downside situations. A firm’s payoff function is typically convex on the upside and concave on the downside. This calls for a valuation approach accommodating asymmetric claims. An alternative to NPV is to discount certainty-equivalent cash flows at the risk-free rate rather than the actually expected cash flows at a risk-adjusted rate. *Option-pricing theory* or *contingent claims analysis*, which is based on this alternative approach, permits the valuation of asymmetric claims. This theory

stems from the seminal work of Bachelier (1900), Samuelson (1965), Black and Scholes (1973), and Merton (1973), who popularized the use of stochastic calculus to value financial instruments. Cox and Ross (1976) and Harrison and Kreps (1979) formulated the risk-neutral valuation approach to financial assets, complementing earlier models with a more intuitive interpretation and implementation of the underlying principles. Merton (1992) and Duffie (2001) provide a broad overview of dynamic asset pricing in continuous time. It is now standard practice in finance and strategy to interpret real investment opportunities as being analogous to financial options. This view is well accepted among academics and practitioners alike and is the core of *real options analysis* (ROA). ROA is actually the extension of option-pricing theory to real investment situations (Myers, 1977). This approach effectively allows one to capture the dynamic nature of decision making because it factors in management's flexibility to revise and adapt strategy in the face of market uncertainty. For this reason, ROA is allegedly more appropriate than other capital-budgeting techniques. Dixit and Pindyck (1994) and Trigeorgis (1996) provide a good overview on ROA.

In oligopolistic industries, firms have also trouble predicting how rivals will behave forming instead beliefs about their likely behaviors. A theory that helps characterize consistent beliefs and predict which strategies opponents will most likely follow is needed to analyze such situations. *Game theory* is used to frame strategic interactions emerging in conflict situations involving parties with different objectives. Game theory attempts to mathematically capture behavior in strategic situations, or games, in which an individual's success in making choices depends on the choices of others through the influence on one another's welfare. Following Myerson (1991), game theory studies "mathematical models of conflict and cooperation between intelligent rational decision makers." A decision-maker is *rational* if she makes decisions consistently in pursuing her own objectives, generally in maximizing the expected value of her payoff valued on a certain utility scale. She is *intelligent* if she knows everything known to the modeler and can make any inference that he can make (in contrast with models of perfect competition). Noncooperative game theory assumes that firms cannot form binding contracts. Fudenberg and Tirole (1991) provide a large overview on noncooperative game theory. The theory of *industrial organization*, by employing techniques from game theory, helps give prescriptive guidance into how firms should react when faced with competition. Tirole (1988) is considered a key reference. Dynamic games used in standard industrial organization assume deterministic developments, ignoring the fact that, in reality, predictions about, e.g., market developments often fail. Fudenberg and Tirole (2002) summarizes a number of dynamic models of industrial organization. Başar and Olsder (1999) present dynamic game theory.

A proper theory of investment should consider both perspectives concurrently, leveraging on corporate finance and industrial organization. Unfortunately, real options analysis and industrial organization have developed somewhat independently: standard corporate finance elaborates solely on exogenous uncertainty, while industrial organization intends to predict how firms (should) react when faced with strategic uncertainty. One objective in this thesis is to bridge this gap. Not surprisingly, researchers with background in financial economics have already tended toward the inclusion of game-theoretic considerations into their existing valuation framework, and researchers from industrial organization and game theory have refined their models with techniques more commonly used in financial economics. There is, however, a trend toward refinement of the deterministic game-theory models and incorporation of stochastic dynamics. Both perspectives ultimately serve to derive better insights into industry dynamics under both market and strategic uncertainty. These two perspectives rely on concepts from dynamic optimization: Finance theory — building upon stochastic control — is primarily concerned with "moves by nature," while dynamic game theory focuses on dynamic optimization problems involving multiple players. Since the mid-1990s, a perceptible trend toward this definition has fostered the emergence of a new class of models, namely "option games." Smit and Trigeorgis (2004) discuss this problem with discrete-time real option techniques, whereas Grenadier (2000b) and Huisman (2001) examine a number

of continuous-time models. Chevalier-Roignant and Trigeorgis (2011) provide an introduction to both types of “option games.” A large overview of the literature is provided in Chevalier-Roignant, Flath, Huchzermeier, and Trigeorgis (2011) or Chapter 7 of the monograph.

2. A dynamic, continuous-time approach to investment

Investment analysis is essentially a dynamic problem. Many factors may influence the behavior of dynamical systems. Some are exogenous to decision-makers, others are under their controls. Under uncertainty whereby a random noise or disturbance affect the evolution of the dynamical system, the decision maker must take into account these random deviations when working out her optimal decisions. This generic framework is fairly descriptive of the issues arising for investment under uncertainty. While the general approach followed here is to keep the assumptions as weak as possible, assumptions that are more restrictive are made in those places where the tradeoff between the losses in generality and the reduction in mathematical complexity appears to be favorable. One such assumption is that of decisions are made dynamically in continuous time. Continuous-time analysis, which contributed greatly to the enhancement of mathematical economics, is based on the following *twin assumptions*:

- (i) the underlying economic phenomena (e.g., asset price movements) take place continuously;
- (ii) decision-makers can act continuously in time.

In stochastic analysis, the procedure of approximating discrete-time Markov chains that describe economic phenomena by Markov processes with continuous sample paths (Itô diffusions) is useful because the mathematical methods associated with a continuum generally lend themselves more easily to analytical treatment than those associated with discrete-time processes. Karatzas and Shreve (1991) provide a mathematically rigorous overview on the mathematics of Itô calculus for diffusion processes. The twin assumptions of continuous-time analysis lead to intertemporal investment policies — characterized by use of optimal control techniques — that are relatively simpler, richer, and more theoretically elegant than those derived from the corresponding discrete-time models. Of course, any continuous decision-making process is an abstraction from physical reality. However, if management makes decisions on a regular basis with negligible periods between decision “rounds,” then the continuous-time solution will be a reasonable approximation to the discrete-time decision-making policy. Whether or not the length of time between actions is short enough for the continuous-time solution to provide a good approximation must be decided on a case-by-case basis. If one desires to concentrate on long-term trends rather short-term, immaterial perturbations — as standard in neoclassical economics, whereby short-run business cycles and conjunctural unemployment are neglected — continuous-time analysis provides a reasonable, easy-to-handle setting.

A noted tool in control theory is the celebrated approach developed by Bellman (1957) known as *dynamic programming*. Dynamic programming provides a general methodology to solve optimization problems — possibly involving stochastic processes — based on the recursive nature of the objective function. At the core of dynamic programming lies Bellman’s principle of optimality stating that: “an optimal policy has the property that, whatever the initial action, the remaining choices constitute an optimal policy with respect to the subproblem starting at the state that results from the initial action.” In discrete time, it is possible to derive the Bellman equation linking the current values received at time t and $t + 1$. In continuous time, dynamic programming involves the derivation of a partial differential equation known as the Hamilton-Jacobi-Bellman (HJB) equation. Under certain conditions, the HJB equation admits a closed-form solution. Dynamic programming makes it possible to reduce a complex, dynamic problem into a series of simpler, static optimization problems. Mathematical finance generally provides a large and rich area of application of stochastic control techniques in continuous time. Two subjects have been particularly addressed in the literature on investment under uncertainty: the optimal timing of a lump-sum investment, modeled as a perpetual American call option (e.g., McDonald and Siegel, 1986) and the problem of capital accumulation under uncertainty as part of the literature on

marginal Tobin’s q (e.g., Abel, 1983; Abel and Eberly, 1994). To approach the investment problem, we will extend notions and techniques derived from stochastic optimal control theory in continuous time to multiplayer settings, in so-called stochastic differential games.

Game theory adopts a different perspective to optimization, whereby the focus is on the formation of beliefs about rivals’ strategies. In this context, *dynamic game theory* (see Başar and Olsder, 1999) is rich enough to help frame and solve dynamic investment problems. In continuous time, the literature on differential games (see Isaacs, 1965; Dockner, Jorgensen, Van Long, and Sorger, 2000) provides the right modeling tools to approach capital-accumulation games, whereby a state is assumed to evolve continuously in time according to a known differential equation. Such dynamic games of capital accumulation include Spence (1979), Fudenberg and Tirole (1983), and Reynolds (1987). Another subfield of dynamic industrial organization involves so-called “games of timing,” whereby firms decide on the appropriate investment timing under competition. In games of timing, two polar types of equilibria might occur: preemption or war of attrition. The theory of timing games has been considerably improved through the seminal work of Reinganum (1981a; 1981b) and Fudenberg and Tirole (1984). Reinganum assumes precommitment to a given investment path and demonstrates that the Nash equilibrium involves sequential investments by the firms or “diffusion.” Fudenberg and Tirole (1984) refine the strategy space to allow for endogenous firm roles by developing a mixed-strategy framework in continuous time and thereby describe how preemption affect investment timing under competition. In both capital-accumulation games and games of timing, the definition of the strategy space is delicate. As Kreps and Spence (1984) note: “one should not allow precommitment to enter by the back door [...] If it is possible, it should be explicitly modeled [...] as a formal choice in the game.” In such models, one should generally look for closed-loop or Markov perfect equilibria.

TABLE 1. Classification of optimization approaches ^a

<i>Time dimension</i>	<i>One player</i>	<i>Several players</i>
<i>Static</i>	Mathematical programming	Static game theory
<i>Dynamic</i>	Optimal control theory	Dynamic game theory

^a Adapted from Başar and Olsder (1999, p.2).

Table 1 classifies the different approaches to optimization. Optimal control and dynamic game theories have been developed in parallel at their inception at RAND Corporation in the 1950s. Unfortunately, these two theories have undergone separate developments since. The branch of optimal control extended its numerous notions and tools to stochastic environments, while dynamic games have been mostly developed in deterministic environments. In other words, stochastic optimal control does not address strategic interactions. Since mainstream ROA draws from stochastic control, a key feature specific to real assets is often ignored — namely, the interplay or strategic interactions taking place among (real) option holders. Many industries are characterized by imperfect competition where a oligopolistic firms have sufficient market power to influence each other’s payoff. The inadequate treatment of competitive dynamics among (real) option holders remains one of the most compelling research gaps in real options analysis (Trigeorgis 1996, p. 376). In short-term competition (static games), the payoff a firm can extract from the market is affected by rivals’ (instantaneous) actions. In long-term competition (dynamic games), present action alternatives may be affected by previous play thereby potentially creating path-dependencies. The oversight of competition may result in misestimating the value of strategic options and in suboptimal investment policies. Strategies that have long-term consequences on the industry structure may be best assessed by use of dynamic game theory. As we follow an economic analysis in continuous time, the relevant subset of game theory are stochastic differential games for which control theorists have analyzed a subclass of solutions in the set of Markov perfect equilibria. Stochastic differential games leverage on the advantage of differential games while also allowing for state equations

that involve random perturbations (a stochastic differential equation) as standard in continuous-time real options models. This lead to mathematical complexity that need be addressed in continuous-time models dealing with strategic investment under uncertainty.

3. Organization of the monograph

Preliminary discussions about the underlying mathematical tools are necessary building blocks for further developments. Part I addresses in detail the modeling challenges and provide way to overcome them. Part II discusses briefly the benchmark cases of monopoly and perfect competition. Part III offers three contributions to the theory of finance and economics dealing with investment under uncertainty and competition. Readers with expert knowledge on continuous-time real options analysis may skip Parts I and II and go to Part III directly.

In Part I, Chapter 1 elaborates on the *investment frictions* that explain why investment decisions have long-term consequences and need to be analyzed by use of dynamic optimization techniques. Besides, it sets up a number of assumptions used throughout the manuscript. Chapter 2 formally discuss the dynamic modeling of random exogenous economic phenomena by use of notions and techniques from *Itô calculus*. Chapter 3 presents dynamic optimization tools used when decision makers can act continuously in time and face uncertain developments; it effectively provides an account of *stochastic control* techniques. Dynamic game theory can be viewed as the extension of optimal control theory to settings involving several interactive players. Building upon notions and tools discussed in Chapter 3, Chapter 4 presents briefly the subset of game theory — used in the sequel — known as stochastic *differential games*.

In Part II, Chapter 5 addresses the benchmark case of *monopoly* with a focus on costly reversibility and investment lags. Chapter 6 discusses the social optimum derived by a central planner (and imposed to decentralized firms) and its connection with the investment behaviors in *perfect competition*.

Part III discusses imperfect competition involving *oligopolist* firms. Part III is composed of three papers — submitted to refereed journals — that deals with imperfect competition under uncertainty. We put a focus on the information firms possess at the time they make their investment decisions. The first paper (Chapter 7) discusses a *market-entry game* under uncertainty. In oligopolies, market entries (lumpy investments) obtain to take place sequentially under uncertainty regardless of whether or not firms can observe (and react to) their rivals' moves. Besides, unlike the case in which firms disregard competition (the myopic or open-loop case), entry does not occur in a socially optimal manner when firms condition their moves on competitors' actions. Chapter 8 discusses a *capital-accumulation game* under uncertainty with the help of techniques from the literature on stochastic differential games. In this chapter, we derive the stationary Markov perfect equilibrium for a dynamic, infinite-horizon capacity investment game in which profits are subject to industry shocks. Finally, Chapter 9 provides an overview of recent contributions to the literature on strategic investment under uncertainty and categorize the research streams. We highlight in this chapter managerial insights about the nature of the competitive advantage (first vs. second-mover advantage), the manner in which information is revealed, firm heterogeneity, capital increment size, and the number of competing firms.

Figure 1 on the following page represents the structure of the monograph, highlighting the interdependences among the chapters and sections.

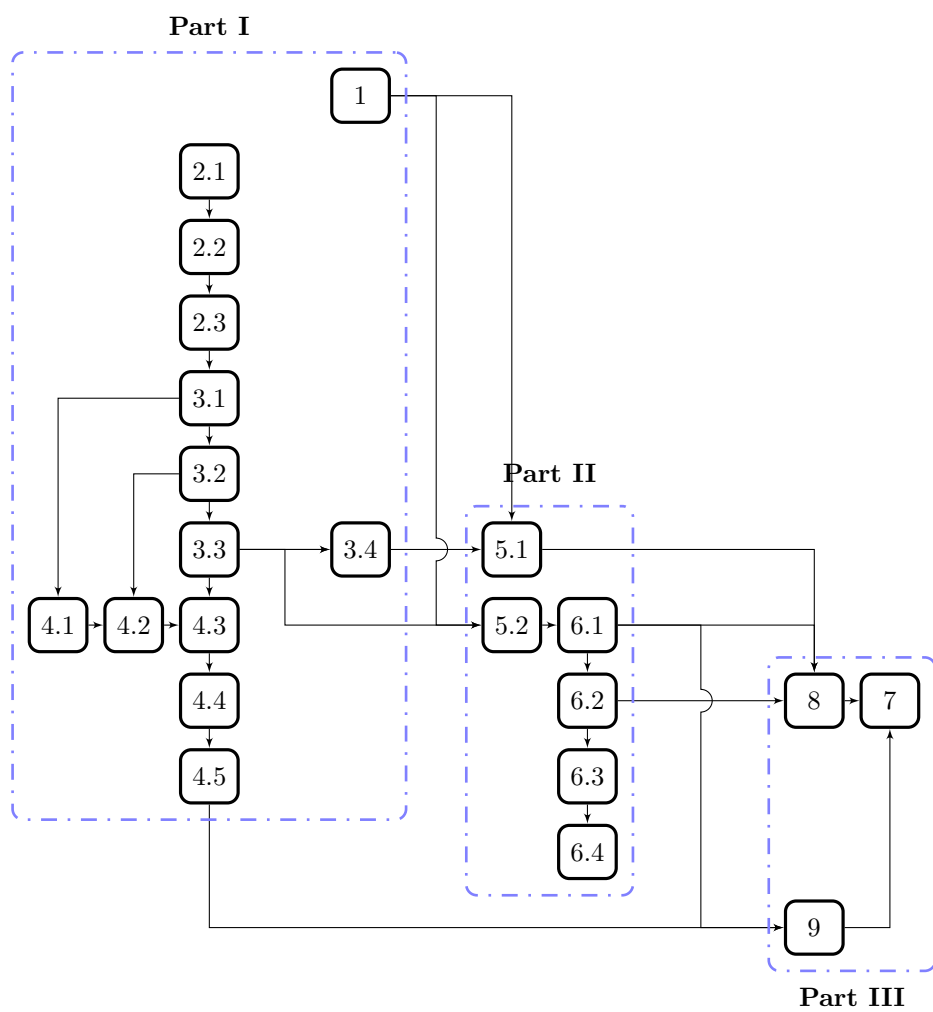


FIGURE 1. Interdependences among the chapters and sections

4. Mathematical notations

For easy reference, we provide an indicative list of symbols used throughout the manuscript. Appropriate definitions or descriptions will be given in the sequel. At places, these symbols may be used in other contexts than indicated below.

Probability and stochastic analysis. Ω represents the set of all states of the world, with outcome $\omega \in \Omega$. \mathcal{F} is a σ -algebra. ν and v are measures. \mathbb{P} is a probability measure and $\mathbb{E}[\cdot]$ is the expectation operator under \mathbb{P} . $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ is a filtration. Random variables and stochastic processes are generally noted by upper-case letters. Some capital letters are used for specific types of processes. X stands for a d -dimensional stochastic process that describes an industry shock whose realizations are not influenced by the participating parties (other than the pseudo-player, nature). Y is a n -dimensional process whose realizations describe the state of a given dynamical system at a given time; the evolution of this “state process” can be partly influenced by players’ (previous) actions. \mathcal{X} and \mathcal{Y} are the state spaces. The lower-case letters $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ are used as arbitrary realizations and initial state value of the processes X and Y respectively. B denotes the standard Brownian motion. A is a bounded-variation process. M is a square-integrable martingale. ξ is used as the cumulative investment process. $\langle \cdot, \cdot \rangle$ is the operator for quadratic covariation, while $\langle \cdot \rangle$ applies for quadratic variation. $\mu(\cdot)$ and $\sigma(\cdot)$ are the drift and diffusion terms of a stochastic process. β_1 and β_2 are the positive and the negative roots of the fundamental quadratic of Brownian motion. \mathcal{D} is the infinitesimal generator. $\tau \in \mathcal{T}$ is a stopping time and \mathcal{T} is the set of admissible stopping times.

Functional analysis. f , $f(\cdot)$ or $x \mapsto f(x)$ corresponds to a function that maps x to $f(x)$. C^k is the space of k times continuously differentiable functions. The partial derivative or gradient of a C^k -function f with respect to the argument x is noted $\nabla_x f$ or f_x . The second-order derivative or Hessian is f_{xx} . $\mathbb{1}$ is the indicator function. \circ is used for compound functions.

Description of a dynamical system or game. t and s are used as time indexes. T is used for the time horizon (finite or infinite) of a control problem. S is the salvage value function. J stands for the payoff function; V for the value function. v is used for the value of a marginal unit of capital. i and j are generalized used as indexes (e.g., matrix dimension) but more specifically serve as subscripts to refer to players in a game. “Players” with subscript $-i$ refer to all other players except player i . Γ is the symbol used for describing a game. λ is a Lagrange multiplier (shadow price). α is an arbitrary control. ψ are used for open-loop controls; ϕ for Markov controls.

Production, capacity, and investment. D stands for the inverse demand function. k represents the capital stock held by a firm (or a group of firms). The process $K = \{K_t\}_{t \geq 0}$ describes the capital stock evolution. $\kappa(\cdot)$ is the investment cost function. I is the magnitude/rate of investment. $\pi(\cdot)$ is the reduced-form profit function.

Miscellaneous. \triangleq means “is defined by.” The positive part of a number is $a^+ \triangleq \max\{a, 0\}$ with $a \in \mathbb{R}$; its negative $a^- \triangleq \max\{-a, 0\}$. \top is used for the transpose of a matrix. tr is used for the trace of a square matrix. \mathbb{R} is used for the real line.

5. Acknowledgement

No research can be the outcome of an entirely solitary process. I thank my advisors Lenos Trigeorgis and Arnd Huchzermeier for their support. I would like to pay adequate tribute to Lenos. Thanks to him, no one could have had a better introduction to real options analysis than mine. I thank my co-author Christoph Flath (Chapter 9) for the pleasant and fruitful cooperation. John Khawam and Christoph Flath provided valuable comments on Chapters 2, 3, and 8. Matt Darnell was very diligent in copy editing Chapters 7 and 9 and improving the general legibility of these chapters. Chapter 9 has benefited from comments and suggestions made by Lorenzo Peccati and two anonymous referees of the *European Journal of Operational Research*. Chapter 7 was presented at 11th Workshop on Optimal Control, Dynamic Games and Nonlinear Dynamics in Amsterdam. Both Chapters 7 and 9 were presented

at the 14th Annual International Real Options Conference in Rome. I thank conference participants for suggestions. Last but not least, I would like to thank my family, relatives, and friends whose support throughout my doctoral studies undoubtedly contributed to the completion of the dissertation. I dedicate this work to Viviane, Patrick, and Aurélien for their unbounded love and support.

Part 1

Theoretical Foundations

Production, Capacity, and Investment

Investment is a dynamic problem that should be analyzed using techniques from dynamic optimization. A number of “frictions” explain why investment decisions cannot be properly analyzed by use of simple mathematical programming instruments. In the following, we intend to give a brief account of these frictions — without being exhaustive.

To smooth out the exposition throughout the monograph, we discuss a number of assumptions that together simplify the analysis of dynamic models presented herein. We describe first the microeconomic environment of a firm and discuss next some key assumptions. The modeling of the exogenous environment via stochastic processes is addressed in the subsequent chapter.

1. Investment frictions

The reasons why investments need to be analyzed by means of dynamic optimization techniques rely on a number of frictions discussed herein.

Costly reversibility or investment sunkness. Resource specificity and unknown resource quality are the principal factors explaining investment sunkness. Capital investments are often firm-specific (e.g., marketing, advertising costs) or industry-specific (e.g., a gold-mine has no alternative use). Besides, shocks affecting firms are often (partly) industry-specific: in cases of economic downturn, there is limited opportunity for an operating firm to sell an industry-specific asset to a rival that faces the same bearish market. Potential buyers may have trouble assessing the quality of a specific “experience good,” estimating it at the average market quality. The owner of an above-average item will be reluctant to sell at a reduced price, altering the average quality and the equilibrium market price for the worse. This describes the common lemons problem (see Akerlof, 1970). Investments in new workers may also be partly irreversible because of high costs of hiring, training, and firing. Employees have gained thanks to their education and over their career path skills specific to a given function or industry. Taking account of this specificity and the lack of instantaneous adaptiveness of workers, some regulatory bodies impose sunk firing costs to avoid erratic hiring/firing waves. Sunkness may also stem from other sources. Government regulations or social conventions may also create a form of country specificity. Multinational firms sometimes face administrative hindrance in reallocating their funds away from high-cost countries. Empirical data as analyzed by Ramey and Shapiro (2001) suggests that the assumption of irreversibility of capital is realistic at least for some industries. This real-world feature can be captured in economic analysis by considering a kinked-adjustment cost function (as discussed in the following). Irreversible investment has been extensively studied as part of the real-options literature since the mid-1980s.

Indivisibility of capital or investment lumpiness. There is an ongoing debate concerning the best approach to model investment problems, the polar cases being the assumption of one-time lump-sum investment and that of incremental capacity investment. McDonald and Siegel’s (1986) deferral option model involves a single lumpy investment. Starting with Pindyck (1988), a subset of the literature on real options analysis examines investment in perfectly divisible capital stocks (e.g., capacity level or workforce) under convex capital-adjustment costs, so that firms can invest or divest by any infinitesimal increment. As Pindyck (1988) suggests, the assumption that firms can continuously and incrementally add capital is extreme as most investments are lumpy. This approach has nevertheless convenient analytical elegance, which explains its success for research purposes. The presence of lumpiness makes it difficult to choose

between a “capacity-leading” and “capacity-lagging” strategy (see Figure 1). In case of “capacity-leading” strategy, demand shortages never occur. Following a “capacity-lagging” strategy, capacity is never under-utilized. Capacity lagging has the advantages of eliminating capacity-coverage risk, being less dependent on accurate forecasting, and delaying capital expenditures. In a single-firm, single-product setting, the choice is a combination of these cases. At the macroeconomic level, all resources can be reasonably assumed to be infinitely divisible. At the microeconomic level, this assumption may not hold for all resources; it may nevertheless be justified to achieve analytical tractability.

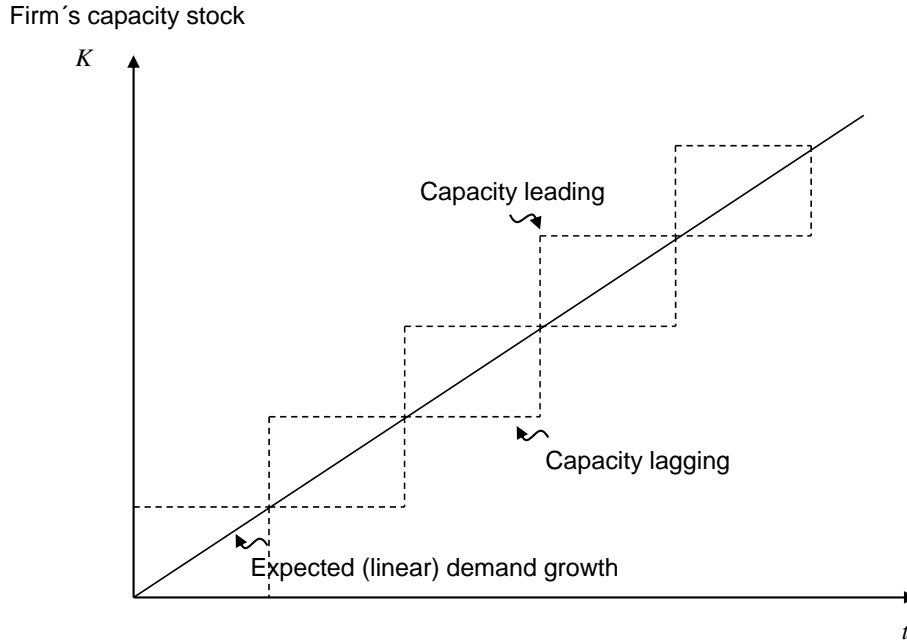


FIGURE 1. Capacity leading vs. capacity lagging

Nonconvexity of the adjustment cost function. In the presence of fixed adjustment costs or economies of scale, the adjustment cost function can be concave. At the firm level, making any capacity adjustment may involve some managerial fixed costs. Mathematically speaking, this creates a discontinuity of the adjustment cost function at 0.

Time-related constraints (e.g., discrete decision times, time-to-build delays, information lags). Herein, we conduct a continuous-time economic analysis to concentrate on a number of issues. A possible interpretation of discrete-time analysis compared to continuous-time models is that decision makers make *short-run commitments* they cannot revise over a period of fixed length considered. From this perspective, decision making at discrete intervals can be considered an investment friction. The value of the discrete-time investment policy is then benchmarked with (i) the continuous-time equivalent and (ii) a variant where firms commit indefinitely (assessing their policy based on the expected developments) and cannot revise their actions in view of actual realizations. *Completion delays* or time-to-build constitute a second category of time-related constraints. As firms may have trouble implementing their managerial decisions over a very short period of time, they have to commit to certain investment paths and assess their investment policies based on truncated expectations about the future. Besides, *information lags* may play a role. Even if firms could implement their decisions instantaneously, they could still have trouble gathering all relevant information on a short notice. Here, information lags may refer to two payoff-relevant factors: the exogenous market environment and the endogenous industry state. Economists often analyze the polar cases of negligible or infinite information lags. In case of *infinite information lags* for the market environment, a stochastic analysis loses its relevance since dispersions around the expected market development are not observed. In dynamic problems, the assumption of negligible

versus infinite information lags gives rise to two distinct state spaces discussed later: open-loop versus closed-loop strategies.

Throughout the monograph, we discuss in detail some of these frictions, with focus on costly reversibility and information lags.

2. Reduced-form profits

The market-clearing price $D : \mathcal{X} \times \mathcal{Q} \rightarrow [0, \infty)$ depends on the realization of some (exogenous) market shock $x \in \mathcal{X} = \mathbb{R}$ and the (endogenous) industry output $q \in \mathcal{Q} = \mathbb{R}_+$. Assume that the inverse demand function is twice continuously differentiable in the shock and differentiable in the industry output; formally, D belongs to the function class $\in C^{2,1}(\mathbb{R} \times \mathbb{R}_+)$. D is nonnegative, decreasing and concave in the output $q \in \mathcal{Q}$, and increasing in the shock. Two types of inverse demand functions are commonly used. The *linear demand* function takes the form

$$(2.1) \quad D(x, q) \triangleq ax - bq,$$

with $a, b \in \mathbb{R}_{++}$, while the *isoelastic demand* function involves the price-elasticity of demand $\eta (> 1)$ and reads

$$(2.2) \quad D(x, q) \triangleq xq^{-\frac{1}{\eta}}.$$

We introduce a *production technology* as a mapping $g : \mathbb{R}^d \mapsto \mathbb{R}_+$ from a given capital stock $k \in \mathbb{R}^d$ to a certain (nonnegative) output $q \in \mathbb{R}_+$. Suppose $g \in C^2(\mathbb{R}^d)$ and let g_k stand for the first-order derivative or gradient of the production function with respect to $k \in \mathbb{R}^d$; it corresponds to the column vector

$$g_k = \begin{pmatrix} g_{k_1} \\ \vdots \\ g_{k_d} \end{pmatrix},$$

with k_i denoting a generic component of $k \in \mathbb{R}^d$. Usually, the production output is assumed to increase componentwise with the capital stock ($g_{k_i} > 0$ for $i = 1, \dots, d$) but marginal returns are decreasing or g is strictly concave in each component ($g_{k_i k_i} < 0$ for $i = 1, \dots, d$). For convenience, the capital stock k is often assumed one-dimensional. Figure 2 depicts an example of a production function with declining marginal returns. In dynamic model, production technologies $g : \mathbb{R} \rightarrow \mathbb{R}$ that exhibit a constant return-to-scale are widely used, since this assumption ensures that the firm produces at full capacity $q = g(k) = k$.

We are eventually interested in the static reduced-form economic *profit* $\pi(x, k)$ received when $x \in \mathcal{X}$ is the realized shock value and $k \in \mathbb{R}^d$ is the industry capital stock. ‘‘Reduced-form’’ profits are used in economics as a modeling device that enable one to ‘‘black-box’’ static strategic interactions into a profit function with a known functional form. The aim is to simplify the analysis of dynamical systems and concentrate instead on purely dynamic aspects. Here, the reduced-form profit is a function of parameters that evolve stochastically over time. We could also consider variable production costs $c(\cdot)$, obtaining for the reduced-form profit

$$(2.3) \quad \pi(x, k) = D(x, g(k))g(k) - c(g(k)).$$

In such a case, the variable cost function typically has a number of restrictions, such as convexity, to ensure a well-behaved profit function. The profit $\pi(\cdot, \cdot)$ is increasing and twice continuously differentiable in x . It is concave in each component of $k \in \mathbb{R}^d$. Examples 1 and 2 provide examples of reduced-form profit functions. As discussed in Example 2, the reduced-form profit can be regarded as the outcome of an instantaneous optimization problem where variable inputs, such as labor or raw materials, are chosen optimally, holding the level of capital fixed.

EXAMPLE 1 (From Bertola and Caballero, 1994). *If the firm’s production and demand function have constant elasticity, the reduced-form profit function obtains as $\pi(x, k) = xk^\alpha$, with $\alpha \in (0, 1)$.*

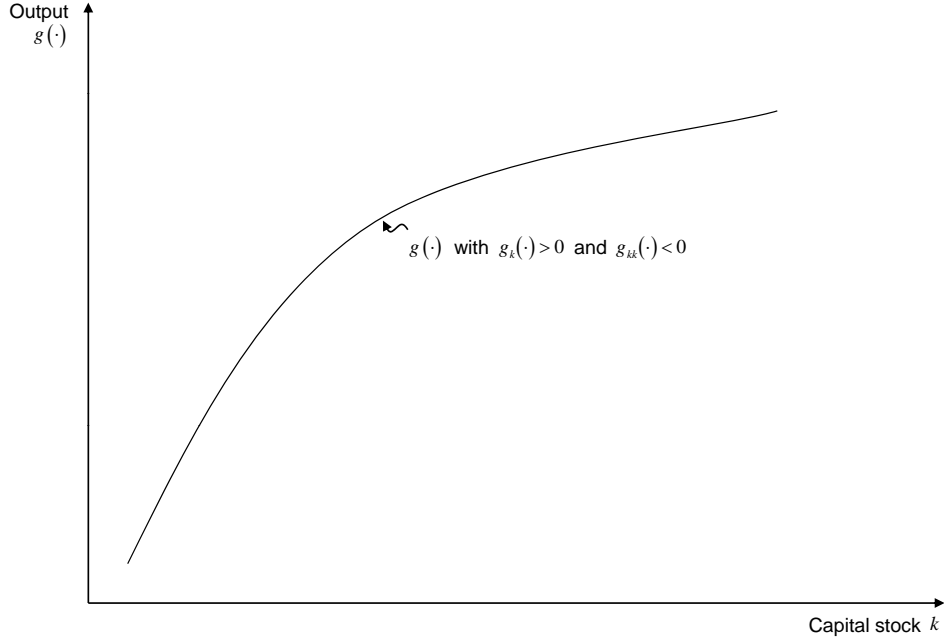


FIGURE 2. Declining marginal returns (one-dimensional case)

EXAMPLE 2 (From Abel and Eberly, 1997; McDonald and Siegel, 1985). Suppose the (exogenous) market shock realization is $x \in \mathcal{X}$. A firm holding a capital stock $k \in \mathbb{R}$ employs l units of labor for a fixed wage of w per unit. The output it produces (before optimization) is given by a Cobb-Douglas production function, $\tilde{g}(l, k) \triangleq l^\alpha k^{1-\alpha}$, with $\alpha \in (0, 1)$. In the short term, the firm cannot change its capital stock but may adjust its workforce. The static maximization program faced by the firm reads

$$\pi(x, k) = \max_{l \in \mathbb{R}_+} \{x l^\alpha k^{1-\alpha} - w l\},$$

which is concave in l . Set $\theta \equiv (1 - \alpha)^{-1} \in (0, 1)$. From the first-order condition, we have $l^* = k w^{-\theta} (1 - 1/\theta)^\theta x^\theta$. By substitution, it obtains after a few algebra steps

$$\pi(x, k) = g(k) x^\theta$$

where $g(k) = \xi k$ and $\xi = (\theta - 1)^{-1} w^{1-\theta} (1 - 1/\theta)^\theta$. In the above, ξx^θ represents the marginal revenue product of capital $\pi_k(x, k)$, i.e., the total revenue change due to a unit increase in capital.

In economics, the *Inada conditions* (see Inada, 1963) impose restrictions on the shape of a production function $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ to guarantee the stability of an economic growth path in a neoclassical economics:

- (i) no output is produced if the firm holds no capital: $g(0) = 0$;
- (ii) the function is continuously differentiable: $g \in C^2(\mathbb{R}_+)$;
- (iii) the function is strictly increasing in the shock $x \in \mathcal{X}$: $g_k(k) > 0$;
- (iv) marginal returns are decreasing: $g_{kk}(k) < 0$;
- (v) the limit of the derivative at 0 is not finite: $\lim_{k \rightarrow 0} g_k(k) = \infty$;
- (vi) the limit of the derivative towards positive infinity is 0: $\lim_{k \rightarrow \infty} \pi_k(k) = 0$.

According to Barelli and de Abreu Pessa (2003), the Inada conditions imply that the production function must be asymptotically Cobb-Douglas.

3. Capital adjustment decisions

Consider $n \in \mathbb{N}$ incumbents firms are operating at time t . Each holds a one-dimensional capital stock, with firm i having unknown capital stock $K_t^i(\omega) \in \mathbb{R}$ at time t (on path $\omega \in \Omega$).¹ Capital is homogenous among firms. The aggregate industry capital stock at time t is $K_t(\omega) = \sum_{i=1}^n K_t^i(\omega)$. The industry capital stock evolution is described by a stochastic process $K = \{K_t\}_{t \geq 0}$. This capital stock may erode over time (*depreciation*). We consider two state equations for the accumulation of capital over time.

3.1. Capital accumulation dynamics. Consider first the capital accumulation equation proposed by Solow (1956) and Swan (1994). At any instant t , firm i can adjust its capital stock by a magnitude I_t^i : if $I_t^i > 0$, the decision corresponds to the common understanding of investment; $I_t^i < 0$ is tantamount to a divestment; if $I_t^i = 0$, the firm stays put. Dixit and Pindyck (1994) argue that divestment can be interpreted as an investment as well since when a firm, e.g., closes down a plant, it incurs a certain fixed cost so as to limit operating costs and ultimately enjoy higher profits. The capital stock depreciates at a fixed proportional rate $\lambda \in \mathbb{R}_+$ per time unit. Following Leibnitz's notations, the capital stock evolves pathwise according to the differential equation

$$(3.1a) \quad dK_t^i(\omega) = (I_t^i(\omega) - \lambda K_t^i(\omega)) dt.$$

Newton notations are generally preferred in the deterministic case, the differential equation reading

$$(3.1b) \quad \dot{K}_t^i(\omega) = I_t^i(\omega) - \lambda K_t^i(\omega).$$

An alternative to Solow-Swan capital accumulation dynamics is offered by Ramsey (1928):

$$(3.2) \quad \dot{K}_t^i(\omega) = g(K_t^i(\omega)) - D_t^i(\omega) - \lambda K_t^i(\omega),$$

where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a production technology with marginal return to scale and $D_t^i(\omega)$ represents the quantity sold at time t . In this case, capital accumulates as a result of intertemporal relocation of unsold output $g(K_t^i(\omega)) - D_t^i(\omega)$, with unsold output being reintroduced in the production process.²

Note that eventhough the capital stocks are random variables, the equation describing their evolution are pathwise deterministic in the cases of equations (3.1) and (3.2)

3.2. Capital-adjustment costs. Managerial decisions may be costly to implement for two main reasons: (i) the required resources may not be readily available; and (ii) the firm may incur extra costs to overcome organizational rigidities. Two groups of adjustment cost functions are commonly considered in economics: convex and concave adjustment cost functions. We here specialize on one-dimensional capital stocks.

3.2.1. Convex adjustment costs. Abel and Eberly (1994) distinguish three types of costs possibly incurred by a firm undertaking gross investment (or divestment):

- (i) *Purchase and resale prices:* Capital can be purchased at a unit price of $\bar{\kappa} > 0$ and sold at a (unit) salvage value of $\underline{\kappa} > 0$.³ In other words, the capital-adjustment cost depends solely on

¹To reduce mathematical complexity, we here consider a one-dimensional capital stock. Considering multiple resources would require to examine whether the resources are complements or substitutes in addition to the factors analyzed herein. Eberly and Van Mieghem (1997) examine the problem faced by a monopolist with multiple resources (or a portfolio of capital stocks).

²Cellini and Lambertini (2008) prove that the type of capital accumulation dynamics considered matters in oligopoly differential games of capital accumulation. The capital accumulation rule à la Solow-Swan gives rise to a weakly time consistent open-loop equilibrium, while Ramsey's dynamics yields strong time consistency or perfection of the open-loop equilibrium.

³In case of infinite planning horizon where firms have no possibility to revise their investment decisions once taken, $\bar{\kappa}$ can be thought as comprising both investment cost and the perpetuity value of subsequent fixed cash flows; $\underline{\kappa}$ include the savings stemming from variable and fixed costs foregone to operate which are bound to be decommissioned.

the investment size I . This cost function is of the kinked piece-wise linear form

$$(3.3) \quad \kappa : I \longmapsto \bar{\kappa} I^+ + \underline{\kappa} I^-,$$

where $I^+ \triangleq \max\{0; I\}$ (investment) and $I^- \triangleq \max\{0; -I\}$ (divestment). To avoid arbitrage, the capital purchase price must exceed (or be equal to) the salvage value ($\bar{\kappa} \geq \underline{\kappa}$). If the latter relationship does not hold, the firm could massively invest in capacity and divest immediately earning infinite profit. The linear adjustment-cost function (3.3) is (weakly) convex and takes the value zero when there is no investment. $\kappa(\cdot)$ is twice differentiable everywhere except at zero. This adjustment cost function is widely used in economics and operations research (e.g., inventory and cash management). Figure 3 depicts this adjustment cost function.

- (ii) *Managerial fixed costs*: Implementing a certain decision within an organization may be costly. Such (nonnegative) costs are independent of the level of investment and are incurred every time investment or divestment takes place. Such costs are captured by the cost function

$$(3.4) \quad \kappa_1(I) \triangleq \kappa_+ \mathbb{1}_{\{I>0\}} + \kappa_- \mathbb{1}_{\{I<0\}},$$

where $\mathbb{1}$ denotes the indicator function.

- (iii) *Adjustment costs*: These costs, denoted by $\kappa_2(I, k)$, are nonnegative and minimal when the firm stays put (no investment takes place). $\kappa_2(\cdot, k)$ is assumed continuous and strictly convex in I on \mathbb{R} .⁴ It is everywhere twice differentiable (with respect to I) except possibly at $I = 0$.

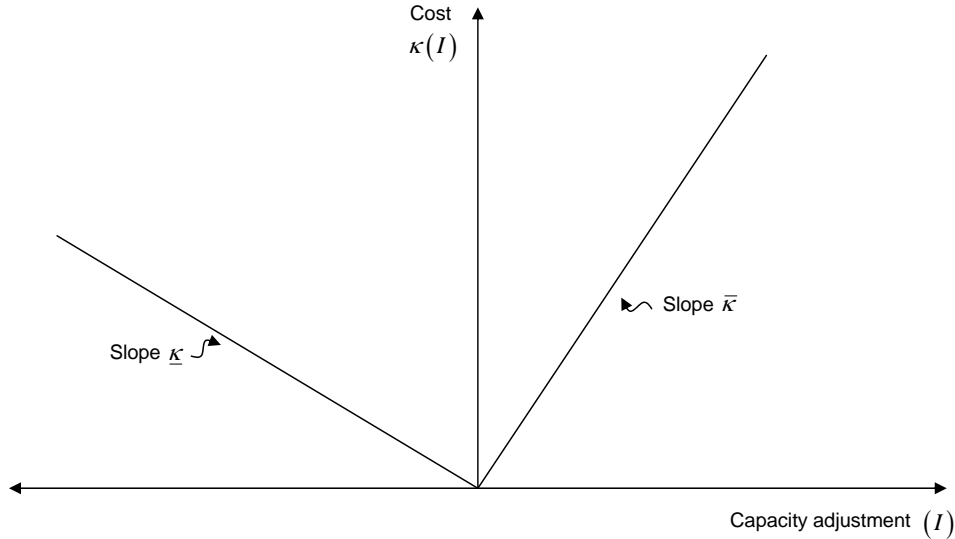


FIGURE 3. Kinked piecewise-linear capital-adjustment cost function

In some applications, we use an augmented adjustment cost function made of the three cost types discussed above, namely,

$$(3.5) \quad k(I, k) = \kappa(I) + \kappa_1(I) + \kappa_2(I, k).$$

The augmented investment cost function is continuous, strictly convex and twice differentiable with respect to I everywhere except possibly at $I = 0$.

EXAMPLE 3 (From Abel and Eberly, 1997). *Consider the (augmented) investment cost function involves two components: (i) $\kappa(I) = \kappa I$, $\kappa \in \mathbb{R}_+$ stands for the (equal) purchase and resale prices of*

⁴The partial derivative of the adjustment-cost function with respect to the investment goes to infinity as investment goes to infinity ($\lim_{I \rightarrow \infty} \frac{\partial \kappa_2}{\partial I}(I, K) = \infty$), and this partial derivative goes to negative infinity as investment goes to negative infinity ($\lim_{I \rightarrow -\infty} \frac{\partial \kappa_2}{\partial I}(I, K) = -\infty$).

capital, $\kappa = \bar{\kappa} = \bar{\kappa}$ and (ii) $\kappa_2(I) = \kappa_2 I^{\frac{2n}{2n-1}}$, with $\kappa_2 \in \mathbb{R}_{++}$ and $n \in \mathbb{N}^*$, is a convex cost of adjustment. That is,

$$k(I) = \kappa I + \kappa_2 I^{\frac{2n}{2n-1}}.$$

Abel and Eberly (1997) prove that, given this investment cost function, the optimal investment rate may be negative. When $n = 1$, $k(I) = \kappa I + \kappa_2 I^2$ corresponds to the quadratic investment cost case.

3.2.2. Concave adjustment costs. The assumption of concavity for the adjustment cost function allows to capture economies of scale. Consider for now the situation where the firm can only expand its capital stock but not reduce it. Luss (1982) gives two popular adjustment cost functions that exhibit such scale economies. These are the (strictly concave) *power cost function* of the form

$$(3.6) \quad \kappa(I) = \kappa_1 \times I^\alpha, \alpha \in (0, 1), I \in \mathbb{R}_+,$$

and the *fixed charged cost function* (in its single resource version)

$$(3.7) \quad \kappa(I) = \begin{cases} 0 & \text{if } I = 0, \\ \kappa_0 + \kappa_1 I & \text{if } I > 0, \end{cases}$$

or some combinations of both. Concave capital-adjustment cost functions are rarely used in economic analysis since the maximization program does not necessarily exhibit concavity. In the subsequent, we will consider convex adjustment cost functions and oftentimes the kinked piecewise linear cost function of equation (3.3).

4. References

Van Mieghem (2003) identifies costly reversibility, indivisibility of capital, and the presence of economies of scale (nonconvexity of the adjustment cost function) as investment frictions. We consider here time constraints and imperfect competition as additional frictions. Abel and Eberly (1994) distinguish different types of adjustment costs incurred by firms when they undertake gross investment or divestment.

Stochastic Calculus

Throughout the manuscript, we introduce market uncertainty by considering an exogenous industry shock whose realizations are distributed according to a given, a priori known probability law. One of the twin assumptions of continuous-time analysis is that the dynamical system evolves continuously in time. To introduce randomness in continuous time, we employ notions and techniques from stochastic (Itô) calculus. In Section 1, we set the premises for the definition of *stochastic processes* in continuous time. Section 2 defines stochastic integrals with an emphasis on Itô integrals and discusses the celebrated Itô-Doebelin formula. We elaborate in Section 3 on a commonly used family of stochastic processes, Itô processes, that are Markov processes with continuous sample paths. Section 4 provides references.

1. Randomness and stochastic processes

Consider a nonempty set Ω . A canonical element $\omega \in \Omega$ is an *outcome* or sample. 2^Ω is the set of all subsets in Ω . The absolute *complement* of a set A in ω is $A^c \triangleq \Omega \setminus A$ or, equivalently saying, $A^c \triangleq \{\omega \in \Omega \mid \omega \notin A\}$. The *empty set* is $\emptyset = \Omega^c$. A *tribe* \mathcal{F} — also known as σ -field or σ -algebra — satisfies the following properties: (i) $\mathcal{F} \in 2^\Omega$ with $\Omega \in \mathcal{F}$, (ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (in particular $\emptyset \in \mathcal{F}$), and (iii) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$. An element A of \mathcal{F} is a *measurable set*. In probability, a measurable set $A \in \mathcal{F}$ is called *event*. The smallest tribe containing all open sets of Ω is the *Borel set*, denoted $\mathcal{B}(\Omega)$. The pair (Ω, \mathcal{F}) with Ω finite is referred to as a *measurable space*.

Suppose two measurable spaces (Ω, \mathcal{F}) and (Ω', \mathcal{F}') . A function $X : \Omega \rightarrow \Omega'$ is *measurable* if $X^{-1}(\mathcal{F}') \subseteq \mathcal{F}$. Note that the notions of measurable set or measurable function require no measure but a tribe. In probability theory, a measurable function is called *random variable*. For each $\omega \in \Omega$, $X(\omega)$ is the *realization* of the random variable. If $(\Omega', \mathcal{F}') = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, the random variable X is said to be \mathcal{F} -measurable.

On a measurable space (Ω, \mathcal{F}) , a *measure* is a function $\nu : \mathcal{F} \rightarrow \mathbb{R}$ such that (i) the empty set has measure zero or $\nu(\emptyset) = 0$, (ii) $\nu(A) \geq 0$ for all measurable sets $A \in \mathcal{F}$, and (iii) for disjoint measurable sets A_i , $\nu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \nu(A_i)$, a property known as *countable additivity*. If $\nu(\Omega) < \infty$, the measure is finite. A measure ν satisfies a monotonicity property, namely if $A, B \in \mathcal{F}$ with $B \subseteq A$, then $\nu(B) \leq \nu(A)$. The triple $(\Omega, \mathcal{F}, \nu)$ is a *measure space*. A proposition holds *ν -almost everywhere* (a.e.) if the set of elements for which the property does not hold is a null set, i.e., if there exists a set $A \in \mathcal{F}$ with $\nu(A) = 0$ such that the proposition holds on A^c . A measure ν defined on \mathcal{F} is *absolutely continuous* with respect to ν if $\nu(A) = 0$ implies $\nu(A) = 0$. We may at places proceed to a change of measures. The *Radon-Nikodym theorem* states that, on a measurable space (Ω, \mathcal{F}) , if a measure ν is absolutely continuous with respect to another measure ν , there exists a unique nonnegative \mathcal{F} -measurable function ξ such that, for each measurable set $A \in \mathcal{F}$,

$$\nu(A) = \int_A \xi \, d\nu,$$

where the integral is understood in the Lebesgue-Stieltjes sense. The measurable function $\xi = d\nu/d\nu$ is the *Radon-Nikodym derivative* or density of ν with respect to ν . If $\nu(A) > 0$ implies $\nu(A) > 0$, then the measures ν and ν are said to be *equivalent*.

On the measurable space (Ω, \mathcal{F}) , a *probability measure* is defined as a measure $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ such that $\mathbb{P}[\Omega] = 1$. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called *probability space*. Under probability measure \mathbb{P} , an event $A \in \mathcal{F}$ “has probability” $\mathbb{P}[A]$. A measurable set $A \in \mathcal{F}$ is called \mathbb{P} -*null set* if $\mathbb{P}(A) = 0$. A probability space

is *complete* if a subset of a \mathbb{P} -null set is necessary a \mathbb{P} -null set of \mathcal{F} . If an event A is such that $\mathbb{P}[A] = 1$, then A is said to hold \mathbb{P} -almost surely (a.s.). Consider $(\Omega', \mathcal{F}') = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. On the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a \mathcal{F} -measurable random variable $X : \Omega \rightarrow \mathbb{R}^n$ implies a certain measure, \mathbb{P}_X , such that, for all $A \in \mathcal{B}(\mathbb{R}^n)$, $\mathbb{P}_X(A) = (\mathbb{P} \circ X^{-1})[A] = \mathbb{P}[\omega \in \Omega \mid X(\omega) \in A]$. A random variable is defined independently of a probability measure. \mathbb{P}_X is the *distribution* of the random variable X ; it is uniquely defined by the joint *distribution function* $F(\cdot)$ where $F(x) = \mathbb{P}[\omega \in \Omega \mid X_i(\omega) \leq x_i, 1 \leq i \leq n]$. $F(\cdot)$ is positive-valued and increasing in each variable x_i . Besides, $\lim_{x_i \rightarrow -\infty} F(x) = 0$ and $\lim_{x_i \rightarrow \infty} F(x) = 1$. Under probability measure \mathbb{P} , the *expectation* of the random variable $X : \Omega \rightarrow \mathbb{R}^n$ is defined as the Lebesgue-Stieltjes integral

$$\mathbb{E}[X] \triangleq \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}^n} x dF(x).$$

Two random variables X and Y , with distribution functions $F_X(\cdot)$ and $F_Y(\cdot)$, are *independent and identically distributed* (i.i.d.) if they are mutually independent, $F_{X,Y}(x, y) = F_X(x)F_Y(y)$, and have the same probability distribution, $F_X = F_Y$. According to the central limit theorem, the probability distribution of the sum or average of i.i.d. random variables with finite variance approaches a normal distribution. Suppose now \mathbb{P}_X is absolutely continuous with respect to the Lebesgue measure. By the Radon-Nikodym theorem, there then exists a nonnegative Lebesgue-measurable function f , corresponding to the *density* of the random variable X , such that

$$\mathbb{P}_X(A) = \int_A f(x) dx, \quad A \in \mathcal{B}(\mathbb{R}^n).$$

Equivalently, we have

$$F(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(z_1, \dots, z_n) dz_1 \cdots dz_n.$$

Consider a random variable $X : \Omega \rightarrow \mathbb{R}^n$ and a function $g : \mathbb{R}^n \rightarrow \mathbb{R}_+$. Then, $g \circ X : \Omega \rightarrow \mathbb{R}_+$ or simply $g(X)$ is also a random variable. Its expectation (if finite) is

$$\mathbb{E}[g(X)] \triangleq \int_{\mathbb{R}^n} g(x) dF(x).$$

If the random variable X admits a density function f , the expectation of $g(X)$ is

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^n} g(x) f(x) dx.$$

What if the function g takes negative values or changes signs? Here, $g : \mathbb{R}^n \rightarrow \mathbb{R}$. We may then decompose, for every $x \in \mathbb{R}^n$, the realization $g(x)$ into its positive part $g(x)^+ \triangleq \max\{0, g(x)\}$ and its negative part $g(x)^- \triangleq \max\{0, -g(x)\}$. The decomposition, $g(x) = g(x)^+ - g(x)^-$, is unique. If $\mathbb{E}[g(X)^+]$ and $\mathbb{E}[g(X)^-]$ are finite, then $g \circ X$ is *integrable* and its expectation is

$$\mathbb{E}[g(X)] \triangleq \int_{\mathbb{R}^n} g(x)^+ dF(x) - \int_{\mathbb{R}^n} g(x)^- dF(x).$$

For two random variables X and Y , the *covariance* is defined by composition as

$$\text{Cov}(X, Y) \triangleq \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^\top\right].$$

The *variance* of a random variable X is $\text{Var}(X) \triangleq \text{Cov}(X, X)$.

Thanks to the Radon-Nikodym theorem, one might proceed with a change of probability measures, say \mathbb{P} and \mathbb{Q} . $\mathbb{E}^{\mathbb{P}}$ (resp., $\mathbb{E}^{\mathbb{Q}}$) is the expectation operator under probability measure \mathbb{P} (resp., \mathbb{Q}). Suppose \mathbb{Q} is absolutely continuous with respect to \mathbb{P} . Then, the Radon-Nikodym derivative $\xi = d\mathbb{Q}/d\mathbb{P}$ exists, so that, for any random variable X , the equality $\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}^{\mathbb{P}}[\xi X]$ holds. Consider, on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a *subtribe* \mathcal{G} of \mathcal{F} , i.e., $\mathcal{G} \subseteq \mathcal{F}$. Suppose the measure \mathbb{Q} is absolutely continuous with respect to \mathbb{P} on \mathcal{G} . The *conditional expectation* of X given \mathcal{G} — denoted $\mathbb{E}[X \mid \mathcal{G}]$ — is then the Radon-Nikodym derivative ξ .

To capture exogenous uncertainty, one needs a mathematical model for the occurrence, at each moment after an initial time, of a random phenomenon. A *stochastic process* is a collection of random

variables $X = \{X_t\}_{t \geq 0}$, with t being generally interpreted as a time index. A realization of the stochastic process at time t is $X(\omega, t)$ or $X_t(\omega)$. For any sample $\omega \in \Omega$, the function $X(\omega, \cdot)$ is called the *sample path* or trajectory of the process associated with the outcome $\omega \in \Omega$. The temporal feature of a stochastic process suggests a flow of information, whereby at any instant t , one can refer to past, present, and future. A *filtration* is a family of tribes $\mathbb{F} \triangleq \{\mathcal{F}_t\}_{t \geq 0}$ such that $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$ for $t < s$. This family captures how information arrives or how uncertainty resolves as time passes, whereby information cannot decrease over time.¹ Given a stochastic process, the simplest — but not only — choice of filtration is the one generated by the process itself, namely $\mathbb{F}^X = \{\mathcal{F}_t^X\}_{t \geq 0}$, where the tribe $\mathcal{F}_t^X \triangleq \sigma(X_s; 0 \leq s \leq t)$ captures the history or former realizations of the process X up to time t . Set $\mathcal{F}_{t+} \triangleq \bigcap_{s > t} \mathcal{F}_s$ and $\mathcal{F}_{t-} \triangleq \bigcup_{s < t} \mathcal{F}_s$. A filtration \mathbb{F} is right (resp., left) continuous if $\mathcal{F}_{t+} = \mathcal{F}_t$ (resp., $\mathcal{F}_{t-} = \mathcal{F}_t$). A stochastic process is *adapted* to the filtration \mathbb{F} , or \mathbb{F} -adapted, if the realization $X_t(\omega)$ is \mathcal{F}_t -measurable for all t . The assumption of adaptedness (to the filtration) simply states that no foresight about the future economic developments is permitted; it is grossly equivalent to the non-anticipativity constraint in operations research. The existence of a stochastic process is closely related to the filtration considered.

We call $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ the *filtered probability space*. A filtered probability space satisfies the *usual conditions* if $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is complete, \mathcal{F}_0 contains the \mathbb{P} -null set in \mathcal{F} , and the filtration \mathbb{F} is right-continuous. Each X_s for $s > t$ is a random variable defined on the measurable space (Ω, \mathcal{F}_t) . \mathbb{P}_t denote the conditional probability as the restriction of \mathbb{P} to $\mathcal{F}_t \subseteq \mathcal{F}$. The conditional expectation operator defined on the probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_t)$ is $\mathbb{E}_t = \mathbb{E}[\cdot | \mathcal{F}_t]$.

The time index can be discrete, continuous, finite, or infinite. Here, we consider random phenomena in continuous time. Recall that a function f admits a *left limit* (resp., *right limit*) at t (if t belongs to the support of f and) $f(t-) \triangleq \lim_{s \uparrow t} f(s)$ (resp., $f(t+) \triangleq \lim_{s \downarrow t} f(s)$) exists. A function f is *left-continuous* (resp., *right-continuous*) at t if $f(t) = f(t-) = \lim_{s \uparrow t} f(s)$ (resp., $f(t) = f(t+) = \lim_{s \downarrow t} f(s)$) exists. The function is *continuous* at t only if $f(t) \triangleq \lim_{s \rightarrow t} f(s)$ or $f(t) = f(t-) = f(t+)$. f is discontinuous at t if $f(t) \neq f(t-)$ or $f(t) \neq f(t+)$. We denote the *jump* at t by $\Delta f(t) = f(t) - f(t-)$. A process with right-continuous sample paths for $t \geq 0$ and left limits for $t > 0$ is called a *càdlàg* or *RCLL process*, where “càdlàg” stands for the French abbreviation of “continue à droite avec limites à gauche” or “Right Continuous with Left Limits. If X is right-continuous and $\mathbb{E}[X_t^2] < \infty$ for all $t \geq 0$ a.s., the process X is *square-integrable*. A process is *continuous* or has continuous sample paths if $X(\omega, \cdot)$ is a continuous function for almost every $\omega \in \Omega$. A process $X = \{X_t\}_{t \geq 0}$ is *increasing* if $X(\omega, \cdot)$ is increasing almost everywhere.

The notion of Markov or *stopping time* plays a prominent role in stochastic analysis. It is a random variable $\tau : \Omega \rightarrow \mathbb{R}_+$, adapted to the filtration \mathbb{F} , such that the event $\{\omega \in \Omega | \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. In other words, the determination of whether “stopping” has occurred by time t can be made based on the information \mathcal{F}_t that has been made available up to time t , without anticipation of the future. In particular, suppose the process $X = \{X_t\}_{t \geq 0}$ is continuous with $X_0 = x \in \mathcal{X}$ with $\mathcal{X} \subseteq \mathbb{R}^d$ almost surely and consider a closed set $A \subset \mathcal{X}$. The stopping time $\tau_x^A(\omega) \triangleq \min\{t \geq 0 | X_t(\omega) \in A\}$ (resp., $\tau_x^A(\omega) \triangleq \min\{t \geq 0 | X_t(\omega) \notin A\}$) at which the process enters (exits) the set A is a *first-hitting time* (*first exit time*).

A stochastic process $X = \{X_t\}_{t \geq 0}$, with $\mathbb{E}[|X_t|] < \infty$ for $t \geq 0$, is called \mathbb{F} -martingale – or *martingale* as shorthand – if $\mathbb{E}[X_{t+\Delta} | \mathcal{F}_t] = X_t$ for all $\Delta > 0$ and $t \geq 0$ \mathbb{P} -a.s. In other words, when considering martingales, the best estimate or the expected value of a future random variable is its current value (as of time t). A process X is a *submartingale* if $\mathbb{E}[X_{t+\Delta} | \mathcal{F}_t] \geq X_t$ and a *supermartingale* if $\mathbb{E}[X_{t+\Delta} | \mathcal{F}_t] \leq X_t$ for all $\Delta \geq 0$ and $t \geq 0$ \mathbb{P} -a.s. Loosely speaking, martingales correspond to the “constant functions” of probability theory, while submartingales and supermartingales are the “increasing” and the “decreasing functions” respectively: on the probability space $(\Omega, \mathcal{F}_t, \mathbb{P}_t)$, the function $\Delta \mapsto \mathbb{E}[X_{t+\Delta}]$ is constant for a

¹For infinite-horizon problems, let \mathcal{F}_∞ denote the smallest tribe on Ω that contains all events in $\mathcal{F}_t, \forall t \geq 0$. Without loss of generality, we shall assume $\mathcal{F}_\infty = \mathcal{F}$.

martingale, increasing for a submartingale, and decreasing for a supermartingale. A martingale is both a supermartingale and a submartingale. Martingales are an essential building block of modern finance theory, being key to the foundation of risk-neutral pricing for complete markets that preclude arbitrage opportunities (see Harrison and Kreps, 1979).

2. Stochastic integration and Itô-Doebelin formula

Stochastic calculus is a branch of mathematics that operates on stochastic processes. Here, we present a number of concepts and techniques (e.g., stochastic integral, Itô-Doebelin formula) that are widely used in mathematical economics and finance. Our goal in this section is to give a meaningful definition of an integral in stochastic environments, as well as a number of properties and formula for applications of stochastic calculus.

The fundamental theory of calculus gave rise to the concept of integration and differential equations. We give a brief overview on Stieltjes integrals and the classical change-of-variable formula in Section 2.1. In Section 2.2, stochastic integration allows to extend such analyzes to situations involving continuous stochastic processes. We discuss a number of properties of the “quadratic variation” in Section 2.3. The change-of-variable formula for stochastic processes, namely Itô-Doebelin formula, is discussed in Section 2.4 for the case of continuous semimartingales.

2.1. Stieltjes integration. Stieltjes integration generalizes Riemann and Lebesgue integration, preserving the many advantages of these notions in a more general measure-theoretic framework. We here give a brief account of Riemann-Stieltjes integrals. Consider an interval $[0, t]$ equipped with a *partition* $\Pi^{(m)} = \{t_0^{(m)}, \dots, t_{2^m}^{(m)}\}$, where $0 = t_0^{(m)} < \dots < t_k^{(m)} < \dots < t_{2^m}^{(m)} = t$. The norm or *mesh* of the partition $\Pi^{(m)}$ is the length of the longest of these subintervals, that is $\|\Pi^{(m)}\| = \max_{1 \leq k \leq 2^m} \|t_k^{(m)} - t_{k-1}^{(m)}\|$. We shall consider a sequence $(\Pi^{(m)})_{m=1}^{\infty}$ of such partitions and assume, for simplicity, that $t_k^{(m)} = kt/2^m$. As m increases, the partition becomes “finer,” with the mesh converging to $\lim_{m \rightarrow \infty} \|\Pi^{(m)}\| = 0$.

The p th-order variation ($p > 0$) of a function g is given by

$$\langle g \rangle^{(p)}(t) \triangleq \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} \left| g(t_k^{(m)}) - g(t_{k-1}^{(m)}) \right|^p.$$

The first variation $\langle g \rangle^{(1)}$ is called *total variation*, while the $\langle g \rangle^{(2)}$ is the *quadratic variation*. A function with finite total variation is said to be of *bounded or finite variation*. Assume that g is of bounded variation on $[0, t]$ and consider a continuous function f . Let $\hat{t}_{k-1}^{(m)} \in [t_{k-1}^{(m)}, t_k^{(m)}]$ for every $k = 1, \dots, 2^m$. Then, the conditions for the existence of a *Riemann-Stieltjes integral* are satisfied, with

$$\int_0^t f(s) dg(s) \triangleq \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} f(\hat{t}_{k-1}^{(m)}) \left| g(t_k^{(m)}) - g(t_{k-1}^{(m)}) \right|.$$

The functions f and g are respectively called the *integrand* and the *integrator*. In the following example, we give a sufficient condition for this integral to exist.

EXAMPLE 4 (Differentiable integrator). Assume the integrator g is continuously differentiable on $[0, t]$ or $g \in C^1([0, t])$. By the mean value theorem, there exists $\hat{t}_{k-1}^{(m)} \in [t_{k-1}^{(m)}, t_k^{(m)}] \subset [0, t]$ such that

$$g(t_{k-1}^{(m)}) = \frac{g(t_{k-1}^{(m)}) - g(t_k^{(m)})}{t_{k-1}^{(m)} - t_k^{(m)}},$$

where $g_t(\cdot)$ is the first-order derivative of g with respect to t . The first variation is then

$$\langle g \rangle^{(1)}(t) = \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} \left| g_t(\hat{t}_{k-1}^{(m)}) \right| \left| t_{k-1}^{(m)} - t_k^{(m)} \right|,$$

where $g_t(\cdot)$ is integrable. By Riemann integration,

$$\langle g \rangle^{(1)}(t) = \int_0^t |g_t(s)| \, ds < \infty,$$

i.e., the total variation is finite. Consider now the quadratic variation. First, $\langle g \rangle^{(2)}(t) \geq 0$ for all $t \geq 0$. Besides, by the mean-value theorem and Riemann integration

$$\begin{aligned} \langle g \rangle^{(2)}(t) &= \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} |g_t(\hat{t}_{k-1}^{(m)})|^2 (t_{k-1}^{(m)} - t_k^{(m)})^2 \\ &\leq \lim_{m \rightarrow \infty} \left\| \Pi^{(m)} \right\| \times \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} |g_t(\hat{t}_{k-1}^{(m)})|^2 |t_{k-1}^{(m)} - t_k^{(m)}| \\ &\leq \lim_{m \rightarrow \infty} \left\| \Pi^{(m)} \right\| \times \int_0^t |g_t(s)|^2 \, ds \end{aligned}$$

Therefore, by the envelope theorem, the quadratic variation of a continuously differentiable function $\langle g \rangle^{(2)}(t)$ is zero for all $t \geq 0$. In the present case with a continuously differentiable integrator, the Riemann-Stieltjes integral corresponds to the Riemann integral,

$$\int_0^t f(s) g_t(s) \, ds.$$

We now extend the analysis to a certain class of càdlàg processes. A process A has “variation finie” or *bounded variation* if, for all $t > 0$ and almost every path ω ,

$$\langle A \rangle^{(1)}(\omega, t) \triangleq \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} \left| A(\omega, t_k^{(m)}) - A(\omega, t_{k-1}^{(m)}) \right| < \infty.$$

The total variation process $\langle A \rangle = \{\langle A \rangle_t^{(1)}\}_{t \geq 0}$ is increasing. A bounded-variation process can be decomposed as the difference between two continuous, increasing, adapted processes $A^+ = \{A_t^+\}_{t \geq 0}$ and $A^- = \{A_t^-\}_{t \geq 0}$ with $A_0^+ = A_0^- = 0$ a.s., that is $A_t = A_t^+ - A_t^-$ for all $t \geq 0$. From the previous discussion, a stochastic integral with a bounded variation process A as integrator can be readily defined. Consider an \mathbb{F} -adapted integrant $H = \{H_t\}_{t \geq 0}$ with continuous sample paths. Here, the stochastic integral is defined “pathwise” by a Riemann-Stieltjes integral as

$$\int_0^t H_s(\omega) \, dA_s(\omega) = \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} H_{\hat{t}_k^{(m)}}(\omega) \left| A_{t_k^{(m)}}(\omega) - A_{t_{k-1}^{(m)}}(\omega) \right|.$$

For a bounded-variation process, we have by definition $\int_0^t |dA_s(\omega)| < \infty$ for all $t \geq 0$. A formula, known as the *change-of-variables formula*, is particularly useful for Riemann-Stieltjes integral.

THEOREM 1 (Change-of-variable formula). *Consider a function $f \in C^1(\mathbb{R})$ and a continuous, bounded-variation process A . Then, (i) the process $\{f(A_t)\}_{t \geq 0}$ has bounded variation and (ii) the following change-of-variables formula holds:*

$$f(A_t) = f(A_0) + \int_0^t f_a(A_s) \, dA_s, \quad \text{or} \quad df(A_t) = f_a(A_t) \, dA_t.$$

PROOF. Since $f \in C^1(\mathbb{R})$ and A is a continuous process, $s \mapsto f_a(A_s(\omega))$ is both continuous and bounded within the interval $[0, t]$ on each path $\omega \in \Omega$. Therefore, the Stieltjes integral $\int_0^t f_a(A_s(\omega)) \, dA_s(\omega)$ is well defined pathwise. By the mean-value theorem,

$$\sum_{k=1}^{2^m} \left| f(A_{t_k^{(m)}}(\omega)) - f(A_{t_{k-1}^{(m)}}(\omega)) \right| = \sum_{k=1}^{2^m} f_a(A_{\hat{t}_k^{(m)}}(\omega)) \left| A_{t_k^{(m)}}(\omega) - A_{t_{k-1}^{(m)}}(\omega) \right|$$

The result obtains by convergence ($m \rightarrow \infty$). □

The above analyses cannot be readily extended to process with unbounded (total) variation, as commonplace in stochastic calculus. This thus calls for an appropriate theory of integration. An equivalent change-of-variables formula also exists in the stochastic case. Itô (1942) defined such a stochastic integral for the case with a standard Brownian motion as integrator. Kunita and Watanabe (1967) follow a more general approach with square-integrable martingales.

2.2. Stochastic integration. Consider a \mathbb{F} -adapted process $H = \{H_t\}_{t \geq 0}$ and an interval $[0, t]$, equipped with a sequence of partitions $\Pi^{(m)} = \{t_0^{(m)}, \dots, t_{2^m}^{(m)}\}$ with $0 = t_0^{(m)} < \dots < t_{2^m}^{(m)} = t$. If a stochastic integral exists with respect to the càdlàg process $X = \{X_t\}_{t \geq 0}$, we denote it $I_t^X(H)$.

Consider a first category of integrands. A process $H^{(m)}$ is a (strictly) *simple*, elementary, or step process if there exists a partition $\Pi^{(m)}$ such that $H_s^{(m)}(\omega) = H_{k-1}^{(m)}(\omega)$ if $s \in (t_{k-1}^{(m)}, t_k^{(m)}]$, where the random variable $H_k^{(m)}$ is $\mathcal{F}_{t_k^{(m)}}$ -measurable. In other words, the process evolves by steps, being constant over (short) time intervals. We can then decompose $H_s^{(m)}$ as

$$(2.1) \quad H_s^{(m)}(\omega) = H_0^{(m)}(\omega) \mathbb{1}_{\{0\}}(s) + \sum_{k=1}^{2^m} H_{k-1}^{(m)}(\omega) \mathbb{1}_{(t_{k-1}^{(m)}, t_k^{(m)}]}(s).$$

For a simple process $H^{(m)}$, we define an *stochastic integral* as the linear combination

$$(2.2) \quad I_t^X(H^{(m)}) \triangleq \sum_{k=1}^{2^m} H_{k-1}^{(m)}(\omega) [X_{t_k^{(m)}}(\omega) - X_{t_{k-1}^{(m)}}(\omega)].$$

The above definition involves no convergence. Besides, it is easy to show that if the integrator X is a martingale, the stochastic integral for simple integrands is also a martingale.

We want to extend the definition of the stochastic integral beyond simple processes, to a larger class of integrands. To do that, we have to be more restrictive on the admissible integrators. Among càdlàg processes, semimartingales turn out to be good integrators. A *continuous semimartingale* $X = \{X_t\}$ admits a representation

$$(2.3) \quad X_t = x + A_t + M_t, \quad \forall t \geq 0$$

where $X_0 = x$ a.s., $M = \{M_t\}_{t \geq 0}$ is a (local) martingale and $A = \{A_t\}_{t \geq 0}$ is a \mathbb{F} -adapted, bounded-variation process. The construction of stochastic integrals for more general integrand process H consists in approximating H by a sequence of simple processes $(H^{(m)})_{m=1}^\infty$, whereby, for $s \in [0, t]$, the random variable $H_s^{(m)}$ is characterized in equation (2.1). Heuristically, the *stochastic integral* $I_t^X(H)$ is the limit of the stochastic integral $I_t^X(H^{(n)})$ defined for simple processes $H^{(n)}$. Formally, if the stochastic integral $I_t^X(H)$ is well defined, then there exists a sequence of random variables $I_t^X(H^{(m)})$ defined in equation (2.2) that *converges in probability* to $I_t^X(H)$, i.e., for all $\epsilon > 0$,

$$\lim_{m \rightarrow \infty} \mathbb{P} \left[|I_t^X(H^{(m)}) - I_t^X(H)| \geq \epsilon \right] = 0.$$

The characterization of the random variables $H_{k-1}^{(m)}$ in equation (2.1) matters for the definition of the stochastic integral $I_t^X(H)$. Set

$$H_{k-1}^{(m)} = \epsilon H_{t_{k-1}^{(m)}}^{(m)} + (1 - \epsilon) H_{t_k^{(m)}}^{(m)}.$$

If $\epsilon = 1$, the stochastic integral $I_t^X(H)$ is an *Itô integral*; we denote it

$$“I_t^X(H) = \int_0^t H_{s-}(\omega) dX_s(\omega)”, \quad \text{or} \quad “I_t^X(H) = \int_0^t H_{s-} dX_s.”$$

Alternatively for $\epsilon = \frac{1}{2}$, the stochastic integral is the *Fisk-Stratonovitch integral*, denoted

$$“\int_0^t H_s \circ dX_s.”$$

When Merton first applied techniques from stochastic calculus in financial economics, he preferred the use of the Itô integral to that of the Fisk-Stratonovich integral. The Itô integral is allegedly more suitable

to economic problems, since “a Stratonovich-type formulation of the underlying price process implies that traders have a partial knowledge about future asset prices that the nonanticipating character of the Itô process does not” (Merton, 1998, footnote 6). As noted, $H_{k-1}^{(m)}$ must be $\mathcal{F}_{t_{k-1}^{(m)}}$ -measurable, which is ensured by the definition of the Itô integral. The Fisk-Stratonovitch integral finds applications in engineering but rarely in economics or operations research. Subsequently, we will not discuss this type of stochastic integral.

We now review a number of properties that the Itô integral $I_t^X(H)$ inherits from the stochastic integral $I_t^X(H^{(m)})$. First, $H \mapsto I_t^X(H)$ is a linear mapping for all $t \geq 0$: $I_t^X(\alpha H^1 + \beta H^2) = \alpha I_t^X(H^1) + \beta I_t^X(H^2)$. Besides, if X is a semimartingale, Itô integral $I^X(H)$ is a semimartingale as well. Similarly, if the integrator $X = M$ is a martingale, $I^M(H)$ is a martingale, namely $\mathbb{E}[I_{t+\Delta}^M(H) \mid \mathcal{F}_t] = I_t^M(H)$ for all $t \geq 0$ and $I_0^M(H) = 0$ a.s. The Itô integral with a martingale as integrand can thus capture the *white noise* of the phenomenon around its expected trend. For square-integrable martingales, we have

$$\mathbb{E} \left[\left(\int_0^t H_s \, dM_s \right)^2 \mid \mathcal{F}_0 \right] = \mathbb{E} \left[\int_0^t H_s^2 \, d\langle M \rangle_s \mid \mathcal{F}_0 \right],$$

a property known as *Itô isometry*. If the integrator A is a bounded-variation process, the stochastic integral $I_t^A(H)$ coincides with the Stieltjes integral defined pathwise.

So far, we have not considered which class of integrands admit a stochastic integral representation. First, the integrand H must be adapted to the filtration generated by the semimartingale X . Since we defined above the stochastic integral as the limit of sums, the integrand is required to satisfy some smoothness conditions. The space of left-continuous processes with right limits, namely “càglàd” processes, is sufficient to prove, e.g., Itô-Doeblin formula or in some applications of stochastic differential equations. In some situations a more general approach to integrands is also wishable. This generalization proceeds in a manner roughly equivalent to how Lebesgue integrals extends the Riemann integral. The class of integrands for which the stochastic integral can be extended consists of so-called locally-bounded *predictable processes* (See Protter, 2003, Chapter IV for more detail).²

2.3. Quadratic variation. We mentioned previously the importance of the notion of variation in the definition of Stieltjes integrals. In stochastic analysis, a major role is played by second-order measures of variation. The *cross-variation* or *quadratic covariation* of two semimartingales X and Y is the process $\langle X, Y \rangle = \{\langle X, Y \rangle_t\}_{t \geq 0}$ defined by

$$\langle X, Y \rangle_t \triangleq \lim_{m \rightarrow \infty} \sum_{k=1}^{2^m} [X_{t_k^{(m)}}(\omega) - X_{t_{k-1}^{(m)}}(\omega)] [Y_{t_k^{(m)}}(\omega) - Y_{t_{k-1}^{(m)}}(\omega)].$$

The following theorem gives an alternative characterization for cross-variation processes.

THEOREM 2 (Cross-variation). *For two semimartingales X and Y ,*

$$(2.4) \quad \langle X, Y \rangle_t = X_t Y_t - \int_0^t X_{s-} dY_s - \int_0^t Y_{s-} dX_s - X_0 Y_0.$$

Equation (2.4) leads to an *integration-by-parts formula* for continuous semimartingales, namely

$$\int_0^t X_s dY_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s dX_s - \langle X, Y \rangle_t.$$

²The predictable σ -algebra is the smallest tribe generated by all “càglàd” processes. All processes that are measurable with respect to this tribe are called *predictable*.

PROOF. We drop the dependence on $\omega \in \Omega$ for notational compactness. Let $S_{t_{k-1}} \triangleq [X_{t_k} - X_{t_{k-1}}] [Y_{t_k} - Y_{t_{k-1}}]$. After a few algebra steps, we obtain

$$\begin{aligned} S_{t_{k-1}}^{(m)} &= X_{t_k}^{(m)} Y_{t_k}^{(m)} - X_{t_{k-1}}^{(m)} Y_{t_{k-1}}^{(m)} - X_{t_{k-1}}^{(m)} [Y_{t_k}^{(m)} - Y_{t_{k-1}}^{(m)}] - Y_{t_{k-1}}^{(m)} [X_{t_k}^{(m)} - X_{t_{k-1}}^{(m)}] \\ \sum_{k=1}^{2^m} S_{t_{k-1}}^{(m)} &= X_{t_m}^{(m)} Y_{t_m}^{(m)} - X_{t_0}^{(m)} Y_{t_0}^{(m)} + \sum_{k=1}^{2^m} X_{t_{k-1}}^{(m)} [Y_{t_k}^{(m)} - Y_{t_{k-1}}^{(m)}] + \sum_{k=1}^{2^m} Y_{t_{k-1}}^{(m)} [X_{t_k}^{(m)} - X_{t_{k-1}}^{(m)}]. \end{aligned}$$

The left-hand term converges to $\langle X, Y \rangle_t$. Equation (2.4) then results from the definition of the Itô integral (with semimartingales as integrator). \square

The *quadratic variation* of a semimartingale — obtained from specializing (2.4) — is

$$(2.5) \quad \langle X \rangle_t^{(2)} = X_t^2 - 2 \int_0^t X_{s-} dX_s - X_0^2.$$

For the cross-variation process, the following property holds.

THEOREM 3. *For two semimartingales X and Y , the polarization formula holds:*

$$\langle X, Y \rangle_t^{(2)} = \frac{1}{4} \left[\langle X + Y \rangle_t^{(2)} - \langle X - Y \rangle_t^{(2)} \right]$$

PROOF. From equation (2.4),

$$\begin{aligned} \langle X + Y \rangle_t^{(2)} &= X_t^2 + Y_t^2 + 2X_t Y_t - 2 \int_0^t X_{s-} dX_s - 2 \int_0^t Y_{s-} dY_s \\ &\quad - 2 \int_0^t X_{s-} dY_s - 2 \int_0^t Y_{s-} dX_s - X_0^2 - Y_0^2 - 2X_0 Y_0, \\ \langle X - Y \rangle_t^{(2)} &= X_t^2 + Y_t^2 - 2X_t Y_t - 2 \int_0^t X_{s-} dX_s + 2 \int_0^t Y_{s-} dY_s \\ &\quad - 2 \int_0^t X_{s-} dY_s + 2 \int_0^t Y_{s-} dX_s - X_0^2 - Y_0^2 + 2X_0 Y_0. \end{aligned}$$

Therefore,

$$\langle X + Y \rangle_t^{(2)} - \langle X - Y \rangle_t^{(2)} = 4 \left[X_t Y_t - \int_0^t Y_{s-} dX_s + \int_0^t X_{s-} dY_s - X_0 Y_0 \right].$$

\square

The process $\langle X, Y \rangle$ is continuous and of bounded variation. The operator $\langle \cdot, \cdot \rangle$ plays the role of an inner product in stochastic calculus and is a *bilinear* form on the space of semimartingales.³ If M and N are independent or *orthogonal*, $\langle M, N \rangle = 0$.

Consider a continuous, square-integrable martingale $M = \{M_t\}_{t \geq 0}$ with starting value $M_0 = 0$ a.s. We can refine the characterization of the quadratic variation for continuous square-integrable martingales.

THEOREM 4 (Quadratic variation process for continuous square-integrable martingales). *For a square-integrable martingale M , $\langle M \rangle = \{\langle M \rangle_t\}_{t \geq 0}$ is the unique adapted, continuous, increasing process of bounded variation for which $M^2 - \langle M \rangle$ is a martingale and $\langle M \rangle_0 = 0$.*

PROOF. From equation (2.5), $\langle M \rangle_t = M_t^2 - 2 \int_0^t M_{s-} dM_s - M_0^2$. By the martingale preservation theorem, $2 \int_0^t M_{s-} dM_s$ is a martingale. Therefore, $M^2 - \langle M \rangle$ must be a martingale. \square

We have almost surely

$$\langle M \rangle_t^{(p)}(\omega) = \begin{cases} \infty & \text{if } p \in (0, 2), \\ < \infty & \text{if } p = 2, \\ 0 & \text{if } p > 2. \end{cases}$$

³Consider semimartingales, X, Y , and Z and $\alpha, \beta \in \mathbb{R}$. The operator $\langle \cdot, \cdot \rangle$ is a bilinear algebra if $\langle \alpha X + \beta Y, Z \rangle = \alpha \langle X, Z \rangle + \beta \langle Y, Z \rangle$ and $\langle X, Y \rangle = \langle Y, X \rangle$.

The sample paths of continuous square-integrable martingales are quite different from “ordinary” continuous functions. Despite its continuity, $M(\cdot, \omega)$ has unbounded total variation and therefore is not differentiable anywhere on \mathbb{R}_+ . Besides, Stieltjes integrals with a square-integrable martingale as integrator cannot be properly defined, nor the standard change-of-variable formula cannot be readily employed. The first remark motivates our earlier digression where we defined stochastic integrals for continuous semimartingales. We now turn to the second task. Note that square-integrable martingales have a bounded quadratic variation. All variations of higher order ($p > 2$) are zero. We will see that second-order terms typically matters in the appropriate change-of-variable formula.

2.4. Itô-Doebelin formula. Since standard differentiation approaches do not apply for a number of stochastic processes, an alternative *change-of-variable* technique, known as *Itô-Doebelin formula*, need be derived.⁴ For the sake of generality, we consider multivariate functions. The space of functions $f : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$ that are twice continuously differentiable in the state and continuously differentiable in time is noted $C^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$. For notational simplicity, we denote, for $f \in C^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$, the partial derivatives by

$$f_t = \frac{\partial f}{\partial t}; \quad f_x = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_d} \end{pmatrix}; \quad f_{xx} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_d} \\ \vdots & \frac{\partial^2 f}{\partial x_i \partial x_j} & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{pmatrix}.$$

where $\frac{\partial f}{\partial x_i} : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is the first-order partial derivative of f with respect to x_i , and $\frac{\partial^2 f}{\partial x_i \partial x_j} : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$ the second-order partial derivative with respect to x_i and x_j , for $i, j = 1, \dots, d$. The function f_x , as the d -dimensional vector of differential operators, is the *gradient* of f , while f_{xx} , as the square ($d \times d$) matrix of second-order partial derivatives of f , is the *Hessian*.

The use of Itô-Doebelin formula generally requires “sufficiently smooth” functions of stochastic processes and involves the quadratic (co)variation of the integrator. We state Itô-Doebelin formula for twice-continuously differentiable functions of continuous semimartingales. Itô-Doebelin formula states that a “smooth function” of a continuous semimartingale is a continuous semimartingale and provides its decomposition.

THEOREM 5 (Itô-Doebelin formula for continuous semimartingales). *Let f belong to the $C^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$ function class and $X = (X^1, \dots, X^d)$ be a continuous d -dimensional semimartingale with $X_0 = x \in \mathbb{R}^d$ a.s. Then, $\{f(X_t, t)\}_{t \geq 0}$ is a continuous semimartingale solving the stochastic differential equation*

$$\begin{aligned} f(X_t, t) = & f(x, 0) + \int_0^t f_t(X_s, s) ds + \sum_i \int_0^t f_{x_i}(X_s, s) dX_s^i \\ & + \frac{1}{2} \sum_{i,j} \int_0^t f_{x_i x_j}(X_s, s) d\langle M^i, M^j \rangle_s \end{aligned}$$

In an informal differential form, the above is expressed as:

$$df(X_t, t) = f_t(X_t, t) dt + \sum_i f_{x_i}(X_t, t) dX_t^i + \frac{1}{2} \sum_{i,j} f_{x_i x_j}(X_t, t) d\langle M^i, M^j \rangle_t.$$

PROOF. We intend to give some hints at how to derive Itô-Doebelin formula. Consider the one-dimensional case for simplicity. Taylor’s Theorem asserts that for $f \in C^2(\mathbb{R})$

$$f(y) - f(x) = f'(x)(y - x) + \frac{1}{2} f''(x)(y - x)^2 + R(x, y),$$

⁴In fact, Doebelin (1940) and Itô (1951) independently discovered this formula. Doebelin’s variant was lost for almost 60 years until the early 2000s. The term “Itô’s lemma” has become standard in financial economics. To pay tribute to both authors we employ the modern name variant.

with $|R(x, y)| \leq r(|y - x|)(y - x)^2$, where $r(\cdot)$ is an increasing function with $\lim_{u \downarrow 0} r(u) = 0$. Consider a partition $\Pi^{(m)}$ of the interval $[0, t]$. Then (dropping the dependence on ω)

$$f(X_t) - f(X_0) = \sum_{k=1}^{2^m} f'(X_{t_{k-1}^{(m)}})(X_{t_k^{(m)}} - X_{t_{k-1}^{(m)}}) + \frac{1}{2} \sum_{k=1}^{2^m} f''(X_{t_{k-1}^{(m)}})(X_{t_k^{(m)}} - X_{t_{k-1}^{(m)}})^2 + R(m),$$

where

$$R(m) \triangleq \sum_{k=1}^{2^m} R(X_{t_k^{(m)}}, X_{t_{k-1}^{(m)}}).$$

The first right-hand term converges to the Itô integral $\int_0^t f'(X_{s-}) dX_s$, while the second term converges to the Stieltjes integral $\int_0^t f''(X_{s-}) d\langle X \rangle_s$. The third term is majored by

$$R(m) \leq \left\{ \sup_k r\left(|X_{t_k^{(m)}} - X_{t_{k-1}^{(m)}}|\right) \right\} \times \sum_{k=1}^{2^m} (X_{t_k^{(m)}} - X_{t_{k-1}^{(m)}})^2.$$

The first left-hand term converges to zero with m , while the second left-hand term converges to the (finite) quadratic variation $\langle X \rangle_t$. Hence, $R(m) \rightarrow 0$. Itô-Doebelin's formula for continuous semimartingales obtains by convergence in probability. \square

A number of variations of Itô-Doebelin formula allow for more general integrators (e.g., discontinuous semimartingales) or less smooth functions. Such extensions include Itô-Tanaka formula, Itô-Tanaka-Meyer formula, Bouleau-Yor formula, and Foellmer-Protter-Shirayev formula.⁵

3. Itô processes

We now explore a certain family of continuous semimartingales, Itô processes, that are built upon the standard Brownian motion. The *standard Brownian motion* is fundamental because it serves as a building block to construct all continuous martingales and a multitude of continuous Markov processes.⁶

We first elaborate on the Markov property to stress the importance of Itô processes. A stochastic process X defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is said to have the weak *Markov property* if

$$\mathbb{P}[f(X_{t+\Delta}) | \mathcal{F}_t] = \mathbb{P}[f(X_{t+\Delta}) | X_t], \quad \forall \Delta \geq 0.$$

A process with this property is called Markov process. In other words, for a Markov process, the conditional probability distribution of future states depend only upon the present state, the past affecting the future only through the present state. When one analyzes financial markets, the Markov property is the mathematical translation of the weak form of the Efficient Market Hypothesis (EMH) as formulated by Samuelson (1965) and Fama (1965). In weak-form efficiency, future asset prices cannot be predicted from technical analysis of historical asset prices or other past data. The Markov property can be generalized to stopping times. A process X is said to have the *strong Markov property* if, for each stopping time τ (with $\tau < \infty$ a.s.),

$$\mathbb{E}[f(X_{\tau+\Delta}) | \mathcal{F}_\tau] = \mathbb{E}[f(X_{\tau+\Delta}) | X_\tau].$$

The strong Markov property implies the weak Markov property but not inversely. Lévy processes are prototypic strong Markov processes. A càdlàg process $X = \{X_t\}_{t \geq 0}$ is a *Lévy process* if (i) $X_0 = 0$ a.s., (ii) for any $0 \leq t_1 < \dots < t_n < \infty$, $X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent, and (iii) for any $s < t$, $X_t - X_s$ is equal in distribution to X_{t-s} . In other words, a Lévy process is any càdlàg process with “stationary independent increments,” whereby Lévy processes generalize the notion of independent and

⁵Harrison (1985), Chapter 4, and Karatzas and Shreve (1991), Chapter 3, discuss some of the extensions.

⁶The standard Brownian motion was first studied by Bachelier (1900) and Einstein (1905). Lévy (1965) carried out a brilliant study of its sample paths and had much influence on the subsequent research on stochastic processes. Karatzas and Shreve (1991), Chapter 3 establish that all continuous martingales can be constructed from the standard Brownian motion.

identically-distributed (i.i.d.) random variables. The most well-known Lévy processes are the standard Brownian motion and the Poisson process.

We provide first a definition of the one-dimensional standard Brownian motion and extend it to allow for multiple (finite) dimensions.⁷ In the one-dimensional case, a *standard Brownian motion* is a continuous, adapted process $B = \{B_t\}$ such that:

- (i) $B_0 = 0$ a.s.;
- (ii) $B_{t+\Delta} - B_t$ is independent of the tribe \mathcal{F}_t ;
- (iii) for $\Delta > 0$, $B_{t+\Delta} - B_t$ is normally distributed with mean zero and variance Δ .

A d -dimensional process $B = (B^1, \dots, B^d)^\top$ is a *multidimensional Brownian motion* if the processes $B^i = \{B_t^i\}$ are one-dimensional Brownian motions that are mutually independent, i.e., $\langle B^i, B^j \rangle = 0$ if $i \neq j$ a.e. The standard Brownian motion has a number of properties that make it suitable for economic analysis. It is a Lévy process, namely, it has stationary independent increments. It satisfies both the weak and the strong Markov property. Finally, it is a continuous (square-integrable) martingale. It has unbounded variation but bounded quadratic variation, given by Lévy Characterization Theorem.⁸ Stochastic integration with respect to the standard Brownian motion is permitted.

THEOREM 6 (Lévy Characterization Theorem). *The quadratic covariations of a multidimensional Brownian motion $B = (B^1, \dots, B^d)$ are $\langle B^i, B^j \rangle_t = \delta_{ij}t$, where δ_{ij} is Kronecker delta*

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

PROOF. The case $i \neq j$ obtains from the definition of a multidimensional standard Brownian motion. For the one-dimensional Brownian motion, consider the process defined by $M_t = B_t^2 - t$ for $t \geq 0$. We have $M_{t+\Delta} - M_t = B_{t+\Delta}^2 - B_t^2 - \Delta$. Since B is a martingale, $\mathbb{E}[B_{t+\Delta}B_t | \mathcal{F}_t] = B_t\mathbb{E}[B_{t+\Delta} | \mathcal{F}_t] = B_t^2$. Thus,

$$\begin{aligned} \mathbb{E}[M_{t+\Delta} - M_t | \mathcal{F}_t] &= \mathbb{E}[B_{t+\Delta}^2 - 2B_{t+\Delta}B_t + B_t^2 - \Delta | \mathcal{F}_t] \\ &= \mathbb{E}[(B_{t+\Delta} - B_t)^2 | \mathcal{F}_t] - \Delta. \end{aligned}$$

By definition of Brownian motion (variance), $\mathbb{E}[M_{t+\Delta} - M_t | \mathcal{F}_t] = 0$ for all $t \geq 0$. \square

We first look at the *one-dimensional Itô process*. Assume a process $X = \{X_t\}_{t \geq 0}$ such that the random variable

$$(3.1a) \quad X_t = x + \int_0^t \mu(X_s, s) ds + \int_0^t \sigma(X_s, s) dB_s,$$

is well defined for all $t \geq 0$. Here, $\mu : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ correspond to the drift and volatility of the process. We say that the process X solves the *stochastic differential equation (SDE)*

$$(3.1b) \quad dX_t = \mu(X_t, t) dt + \sigma(X_t, t) dB_t.$$

For simplicity, we shall use μ_t as shorthand for the random variable $\mu(X_t, t)$ and σ_t for $\sigma(X_t, t)$. The term μ_t corresponds to the expected rate of change of $X = \{X_t\}_{t \geq 0}$ over the next infinitesimal small instant. σ_t^2 is the instantaneous rate of change in the variance.

⁷The definitions of a standard Brownian motion and a Wiener process differ in some respects. The definition we provide corresponds to the definition of a Wiener process. Rigorously speaking, a Brownian motion is defined as a square-integrable \mathbb{F} -martingale such that $\mathbb{E}[(M_{t+\Delta} - M_t)^2 | \mathcal{F}_t] = \Delta$ for $\Delta > 0$. The Lévy characterization theorem — discussed next — implies that both definitions describe the same mathematical object.

⁸Lévy Characterization Theorem has a stronger implication that suggested in Theorem 6. It asserts that if M is a continuous martingale such that $M_0 = 0$ and $\langle M \rangle_t = t$ (or $\mathbb{E}[M_t^2] = t$), then M is a Brownian motion.

Consider now a d -dimensional Brownian motion $B = (B^1, \dots, B^d)^\top$. A drift $\mu = (\mu^1, \dots, \mu^n)^\top$ can be defined as a n -dimensional vector of functions $\mu^i : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$. A diffusion σ is a $n \times d$ matrix of functions $\sigma^{ij} : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$, denoted

$$\sigma = (\sigma^{ij}) = \begin{pmatrix} \sigma^{11} & \dots & \sigma^{1d} \\ \vdots & \sigma^{ij} & \vdots \\ \sigma^{n1} & \dots & \sigma^{nd} \end{pmatrix}.$$

with rows $\sigma^i = (\sigma^{i1}, \dots, \sigma^{id})$ and vectors $\sigma^j = (\sigma^{1j}, \dots, \sigma^{dj})^\top$. By extension to the one-dimensional Itô process, a n -dimensional Itô process $X = (X^1, \dots, X^n)^\top$, with starting value $x \in \mathbb{R}^n$ a.s, solves a systems of n stochastic differential equations (one for each X^i)

$$dX_t^i = \mu^i(X_t, t) dt + \sigma^i(X_t, t) dB_t.$$

In compact form, $X = (X^1, \dots, X^n)^\top$ solves the (multidimensional) stochastic differential equation

$$(3.2) \quad dX_t = \mu(X_t, t) dt + \sigma(X_t, t)^\top dB_t.$$

At places, we use matrix notations. The *trace* of a square ($d \times d$) matrix, $A = (a_{ij})_{1 \leq i, j \leq d}$ is defined as the sum of the elements on its main diagonal, namely

$$\text{tr} \langle A \rangle = a_{11} + a_{22} + \dots + a_{dd} = \sum_{i=1}^d a_{ii}$$

The trace is only defined for a square matrix. A norm on the space of matrices is $\|A\| \triangleq \sqrt{\text{tr} \langle AA^\top \rangle}$ for any matrix A .

Stochastic differential equations are generalizations of ordinary differential equations with the inclusion of a random, white noise term driven by the standard Brownian motion. There are different notions of solutions to the stochastic differential equation (3.2) depending on the roles played by the underlying filtered probability space and Brownian motion. We will primarily discuss *strong solutions* to (3.2). Before proceeding with sufficient conditions for the existence of a solution of (3.1a), we need to define a number of concepts. A function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is *Lipschitz continuous* if there exists a constant $M \in \mathbb{R}_+$ such that, for any $x, y \in \mathbb{R}^d$,

$$\|f(x) - f(y)\| \leq M \|x - y\|$$

Lipschitz continuity is thus a smoothness condition that is stronger than regular continuity because the Lipschitz condition limits the speed with which a function can change. In the one-dimensional case, Lipschitz continuity states that the slope will never be steeper than a certain number c . A function f satisfies a *growth condition* if there is a constant M such that, for any $x \in \mathbb{R}^d$ and any time t ,

$$\|f(x)\|^2 \leq M (1 + \|x\|^2).$$

This condition ensures that the function f does not “explode.” Sufficient conditions for the existence of a solution of (3.1a) are given in the following theorem.

THEOREM 7 (Sufficient conditions for existence and uniqueness of Itô process). *Consider a process $X = \{X_t\}_{t \geq 0}$ with $X_0 = x$ a.s.. X solves (3.2) if the following conditions are jointly satisfied:*

- (i) *the processes $\{\mu_t\}_{t \geq 0}$ and $\{\sigma_t\}_{t \geq 0}$ are \mathbb{F} -adapted;*
- (ii) *the processes $\{\mu_t\}_{t \geq 0}$ and $\{\sigma_t\}_{t \geq 0}$ are finite-variation processes;*
- (iii) *the functions μ and σ are Lipschitz continuous in x ;*
- (iv) *the functions μ and σ satisfy the growth condition in x .*

PROOF. See Øksendal (2007), Section 5.2. □

For strong solutions, the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and the \mathbb{F} -Brownian motion are given a priori. For weak solutions, they are parts of the solution. Note that the coefficients in equation (3.2) do not depend on the outcome ω explicitly, only implicitly through the solution $X = \{X_t\}_{t \geq 0}$.

Since Itô processes satisfy the (strong) Markov property, employing such processes to model an exogenous shock may help avoid the “curse of dimensionality.” This explains the success of Itô processes in economics and operations research. The Itô process covers a broad family of stochastic processes used in economic analysis.⁹ The *arithmetic Brownian motion* (ABM) solves

$$(3.3) \quad dX_t = \mu dt + \sigma dB_t,$$

with constant drift μ and diffusion σ . The *geometric Brownian motion* (GBM) follows

$$(3.4) \quad dX_t = \mu X_t dt + \sigma X_t dB_t;$$

The *arithmetic Ornstein-Uhlenbeck* (OU) process has the the stochastic differential equation representation

$$(3.5) \quad dX_t = \eta (X_t - \bar{X}) dt + \sigma dB_t,$$

where the constant η describes the speed or strength of reversion toward the long-term mean \bar{X} . The *geometric Ornstein-Uhlenbeck* follows

$$(3.6) \quad dX_t = \eta X_t (X_t - \bar{X}) dt + \sigma X_t dB_t.$$

An alternative mean-reverting process (used by, e.g., Sick, 1995; Tsekrekos, 2010) solves:¹⁰

$$(3.7) \quad dX_t = \eta (X_t - \bar{X}) dt + \sigma X_t dB_t.$$

Itô-Doeblin formula given in Theorem 5 for continuous semimartingales can be specialized to Itô processes.

THEOREM 8 (Itô-Doeblin formula applied to Itô process). *Consider a d -dimensional Itô process $X = \{X_t\}_{t \geq 0}$ and a function $f \in C^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$. Then, $\{f(X_t, t)\}_{t \geq 0}$ is an Itô process solving the stochastic differential equation*

$$df(X_t, t) = \left[f_t(X_t, t) + \mu_t^\top f_x(X_t, t) + \frac{1}{2} \text{tr} \langle \sigma_t \sigma_t^\top f_{xx}(X_t, t) \rangle \right] dt + f_x(X_t, t)^\top \sigma_t dB_t.$$

For a one-dimensional Itô process, the stochastic differential equation simplifies to

$$df(X_t, t) = \left[f_t(X_t, t) + \mu_t f_x(X_t, t) + \frac{1}{2} \sigma_t^2 f_{xx}(X_t, t) \right] dt + \sigma_t f_x(X_t, t) dB_t.$$

PROOF. Consider the multidimensional case. We decompose the semimartingale $X = \{X_t\}_{t \geq 0}$ as $X = x + A + M$; here, A is a bounded-variation process with $A_t = \int_0^t \mu(X_s, s) ds$ and M is a square-integrable martingale with $M_t = \int_0^t \sigma(X_s, s) dB_s$. Given that $\langle \cdot, \cdot \rangle$ is a bilinear form, we obtain from Theorem 5

$$\begin{aligned} f(X_t, t) &= f(x, 0) + \int_0^t f_t(X_s, s) ds + \int_0^t f_x(X_s, s)^\top \mu_s ds + \int_0^t f_x(X_s, s)^\top \sigma_s dB_s \\ &\quad + \frac{1}{2} \sum_{i,j \geq 1} \int_0^t f_{x_i x_j}(X_s, s) \sigma_s^i \sigma_s^j d \langle B^i, B^j \rangle_s. \end{aligned}$$

Theorem 8 now results from Lévy’s Characterization theorem 6 according to which $\langle B^i, B^j \rangle_s = \delta_{ij} s$. \square

⁹See the Appendix of Chevalier-Roignant and Trigeorgis (2011) for a detailed discussion of these processes.

¹⁰Following Tsekrekos (2010), the mean-reverting process in equation (3.7) is more plausible than the geometric OU process of equation (3.6) as a market-clearing price. Since X reverts in the long-term to the mean \bar{X} , one should require that the price of any number of goods, say n of them, should revert to $n\bar{X}$ in the same manner. In other words, it seems reasonable to require that the drift and diffusion of any mean-reverting process used as a candidate market-clearing price must be homogeneous functions of degree one of the pair (X_t, \bar{X}) .

At places, it is useful to estimate the expected change in the value of Itô process $\{f(X_t, t)\}$ over an infinitesimal time interval. The *infinitesimal generator* \mathcal{D} is the appropriate operator to achieve this goal.¹¹

$$(3.8) \quad \mathcal{D}f(X_t, t) \triangleq \lim_{\Delta \downarrow 0} \frac{\mathbb{E}_t[f(X_{t+\Delta}, t+\Delta)] - f(X_t, t)}{\Delta}.$$

The following theorem permits to characterize the infinitesimal generator as a *second-order differential operator* if $f \in C^{2,1}(\mathbb{R}^d, \mathbb{R})$.

THEOREM 9 (Infinitesimal generator). *Consider an Itô process associated with the drift vector $\mu = (\mu^1, \dots, \mu^d)^\top$ and diffusion matrix $\sigma = (\sigma_{ij})$. If $f \in C^{2,1}(\mathbb{R}^d, \mathbb{R})$,*

$$\mathcal{D}f(X_t, t) = f_t(X_t, t) + \mu_t^\top f_x(X_t, t) + \frac{1}{2} \text{tr} \langle \sigma_t \sigma_t^\top f_{xx}(X_t, t) \rangle.$$

PROOF. Consider $f \in C^{2,1}(\mathbb{R}^d, \mathbb{R})$. Let \mathcal{L} denote the second-order operator given by

$$\mathcal{L}f(X_t, t) = f_t(X_t, t) + \mu_t^\top f_x(X_t, t) + \frac{1}{2} \text{tr} \langle \sigma_t \sigma_t^\top f_{xx}(X_t, t) \rangle.$$

From Itô-Doeblin formula (Theorem 8),

$$f(X_{t+\Delta}, t+\Delta) = f(X_t, t) + \int_t^{t+\Delta} \mathcal{L}f(X_s, s) ds + \int_t^{t+\Delta} f_x(X_s, s)^\top \sigma_s dB_s$$

The third right-hand term is a martingale with starting value 0. Hence,

$$(3.9) \quad \mathbb{E}_t[f(X_{t+\Delta}, t+\Delta)] = f(X_t, t) + \mathbb{E}_t \left[\int_t^{t+\Delta} \mathcal{L}f(X_s, s) ds \right]$$

The theorem obtains by the mean-value theorem. □

Using the infinitesimal generator notation, Itô-Doeblin formula in case of Itô processes becomes

$$df(X_t, t) = \mathcal{D}f(X_t, t) dt + f_x(X_t, t)^\top \sigma_t^\top dB_t.$$

The expected value of a sufficiently smooth function of an Itô process at a stopping time is given by Dynkin's formula.

THEOREM 10 (Dynkin's formula). *Consider an admissible stopping time τ (with $\mathbb{E}_0[\tau] < \infty$) and a function $f \in C^2(\mathbb{R}^d)$. If X is a time-homogenous Itô process or diffusion, then*

$$\mathbb{E}_0[f(X_\tau)] = f(x) + \mathbb{E}_0 \left[\int_0^\tau \mathcal{D}f(X_s) ds \right]$$

PROOF. Readily obtains from Theorem 9 extended to stopping times. □

4. References

Karatzas and Shreve (1991) provides a rigorous, detailed discussion of Brownian motion and generally continuous semimartingales. Karatzas and Shreve (1991), Chapter 3, discuss generalization of Itô-Doeblin's formula to less "smooth" functions (with continuous martingales as integrator). Protter (2004), Chapter 2, define stochastic integrals with respect to (non necessarily continuous) semimartingales.

¹¹Two definitions of the infinitesimal generator exist in the literature. One applies to time-homogenous Itô processes or diffusion, while the second allows the function to depend on time as well. To avoid a multitude of notations, we take the second approach, being understood that the first-order derivative with respect to time drops in the operator for time-homogeneous settings.

Stochastic Control

As noted in the introduction, continuous-time analysis relies on the twin assumptions that economic phenomena take place in continuous time and that decision makers can act at any instant within a time interval. In this chapter, we discuss mathematical concepts useful for the modeling of random changes in continuous time. Now, we want to proceed with determining how to model and solve the problem of determining how agents decide optimally in a continuous-time setting. Control theory is a branch of applied mathematics concerned with determining, over a certain (finite or infinite) period of time, optimal decisions and state values for a given dynamic system (under an appropriate optimality notion). Optimal control theory was initially developed to support needs for military applications during the “cold war;” Not surprisingly, two independent approaches were followed, one in the Soviet Union (Pontryagin) and the other in the US (Bellman). Since its inception, control theory has offered a rich toolkit to economic researchers because it provides way to solve single-agent dynamic optimization problems. Optimal control in continuous time and stochastic environment is often called *stochastic control*.

The chapter is organized as follows. In Section 1, we describe the structure of a typical stochastic control problem. Section 2 discusses different types of control functions used in dynamic problems and develops the notion of time consistency. Section 3 describes distinct approaches to solve stochastic control problems. Section 4 presents a number of dynamic problems that have a different structure but can be solved with use of instruments borrowed from control theory as well. Section 5 provides references. In Chapter 4, we will extend the analysis to allow for multiple players, as part of continuous-time *dynamic game theory*.

1. A dynamic problem

We examine first a problem over a known planning period $[0, T]$ with finite horizon $T > 0$. Suppose a standard one-dimensional Brownian motion $B = \{B_t\}_{t \geq 0}$ and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Consider a d -dimensional *state space* $\mathcal{Y} \subseteq \mathbb{R}^d$, where $Y_t \in \mathcal{Y}$ describes the state of the dynamic system at time t . $Y = \{Y_t\}_{t \geq 0}$ is the *state process*; $Y(\omega, \cdot)$, is called *state trajectory* or *path*. \mathcal{Y} is the space consisting of all permissible state trajectories, with $Y \in \mathcal{Y}$. The state of the system is almost surely known at the outset with $Y_0 = y \in \mathcal{Y}$ and evolves going onwards according to a certain known law of motion — e.g., a difference or differential equation — possibly depending on previous state realizations and past “actions” meant to alter the dynamical system. At each time, the decision maker has a set $A \subseteq \mathbb{R}^n$ of *actions* she can chose from.¹ Over the period $[0, T]$, the decision maker earns a certain profit characterized by a *running profit* function $\pi : \mathcal{Y} \times [0, T] \times A \rightarrow \mathbb{R}$. At the end of the planning horizon (if finite), the decision maker receives a certain salvage value, as given by a terminal reward, scrap value, or *salvage value* function $S : \mathcal{Y} \times \mathbb{R}_+ \rightarrow \mathbb{R}$. At time t , the decision maker knows which scenario has occurred among the set of all possible scenarios \mathcal{G}_t ; the characterization of the tribes $\mathcal{G}_t \subseteq \mathcal{F}_t^{Y, \alpha}$ — where $\mathcal{F}_t^{Y, \alpha} \triangleq \sigma(Y_s, a_s; 0 \leq s \leq t)$ — determines the *information structure* of the problem.

A control or *strategy* is a complete contingent plan of actions, conditional on available information, that indicates which action to take at each and every state and time. Formally, a control α is a random function that maps $\Omega \times [0, T]$ into A . To stress the influence of the strategy choice on the state evolution, we shall denote $Y^\alpha = \{Y_t^\alpha\}_{t \geq 0}$. The distinction between static actions and controls as a contingent plan

¹The set might be the same over all stages or depending on the state trajectory in more advanced settings.

of actions is essential to the understanding of dynamic problems. The set of “admissible” controls or strategies over $[0, T]$ is $\mathcal{A}_{[0, T]}$ or \mathcal{A} as shorthand. In the present case, we consider an equation of motion of the form

$$(1.1) \quad dY_t^\alpha = \mu(Y_t^\alpha, t, a_t) dt + \sigma(Y_t^\alpha, t, a_t) dB_t; \quad Y_0^\alpha = y,$$

where $\mu : \mathcal{Y} \times [0, T] \times A \rightarrow \mathbb{R}^d$ is a *controlled drift* function and $\sigma : \mathcal{Y} \times [0, T] \times A \rightarrow \mathbb{R}^d$ a *controlled diffusion* function. For existence and uniqueness, we assume that the drift and the diffusion processes are \mathbb{F} -adapted and have finite variations and that μ and σ are Lipschitz continuous and satisfy the growth condition (see Theorem 7, page 25). The decision maker chooses a control α^* to maximize its payoff. Given a (known) initial state $y \in \mathcal{Y}$, the *payoff* of a given admissible control α is defined as

$$(1.2) \quad J(y, 0, \alpha) = \mathbb{E}_0 \left[\int_0^T \pi(Y_t^\alpha, t, a_t) dt + S(Y_T^\alpha, T) \right].$$

We denote such a control problem by $\Gamma(y, 0)$. The objective function in equation (1.2) with $\pi(\cdot, \cdot) \neq 0$ and $S(\cdot, T) \neq 0$ is said to be of the *Bolza form*. If $S(\cdot, T) = 0$ and $\pi(\cdot, \cdot) \neq 0$, it has the *Lagrange form*, whereas if $\pi(\cdot, \cdot) = 0$ and $S(\cdot, T) \neq 0$, it is of the *Mayer form*. It is well known and easy to show that these three problems are mathematically equivalent. The *value* received in optimum for initial state $y \in \mathcal{Y}$ is

$$(1.3) \quad V(y) = V(y, 0) \triangleq \sup_{\alpha \in \mathcal{A}} J(y, 0, \alpha) = J(y, 0, \alpha^*).$$

Such an α^* is called *optimal control*. The path Y^* generated by the optimal control α^* is the *optimal state trajectory*. (α^*, Y^*) is the *optimal pair*.

2. Types of controls

So far, we provided a loose description of the information structure and the strategy space. In applied control literature, a control function $\alpha : \Omega \times [0, T] \rightarrow A$ is often called *decision rule*. Several types of *admissible* controls or strategies are considered. We rank them below in terms of increasing computational complexity:

- *Deterministic or open-loop controls* involve strategies of the form $\alpha : [0, T] \rightarrow A$, where $\alpha(\omega, t) = \alpha(t)$ for all $t \in [0, T]$;
- *Markov controls* are functions α of the form $\alpha : \mathcal{Y} \times [0, T] \rightarrow A$. The decision maker’s choice at time t depends only on the current value of the state. If $\alpha : \mathcal{Y} \rightarrow A$, the control is *stationary Markov*;
- The decision maker may have information about the true state value that is partial, delayed, or “noisy” (affected by some random disturbances). In such a case, the control strategy must be adapted to a “biased” filtration of histories. This situation relates to the *filtering problem*;
- *Closed-loop controls* involve strategies $\alpha : \mathcal{F}_t^{Y, \alpha} \rightarrow A$ that take into consideration the entire observed state trajectory, i.e., $\mathcal{F}_t^{Y, \alpha} = \sigma(Y_s, a_s; 0 \leq s \leq t)$. In other words, a closed-loop strategy is adapted to the filtration.²

Closed-loop strategies are more general but are less tractable. In *closed-loop* strategies, the decision makers can condition their actions (when called upon to act) on the observed sequence of actions and state realizations. In contrast, *open-loop* strategies do not depend on past observations and/or previous play and instead set ex ante a certain investment path to be pursued regardless of the actual developments. Given an initial state $y \in \mathcal{Y}$, a Markov control $\alpha \in \mathcal{A}$ is *admissible* if there exists a unique regulated or *controlled Itô process* Y^α that solves the stochastic differential equation (1.1). In other words, the state trajectory Y^α resulting from a particular control must be well defined (strong solution). Markov

²The optimal control literature is not consistent in defining *feedback strategies*. Sometimes, the notion is used to describe closed-loop strategies as defined above or Markov strategies. To avoid confusion, we decide not to use this notion further.

strategies are decision rules in which the choice of a player's current action is conditioned on current time and vector space $y \in \mathcal{Y}$. By introducing the Markovian assumption, one assumes that the history of state realizations and actions is not relevant for the choice of action at time t but only the present consequences as reflected in the current value of the state. The use of Markovian strategies is a natural choice in the set-up of state-space problems or games where the history until time t is summarized in the value of the state vector $y \in \mathcal{Y}$.

Even though an optimal strategy constitutes an optimal solution in its control class, some solutions might be preferred due to certain appealing properties. A *refinement scheme* helps select among strategy candidates that are optimal in a certain control class. A noted refinement scheme imposes *time consistency* of the preferred optimal solution. Two forms may be considered: the weak and strong form. Consider the control problem $\Gamma_{[0,T]}$ over the time interval $[0, T]$. $\alpha_{[0,T]}^* \in \mathcal{A}_{[0,T]}$ is an optimal solution within the class $\mathcal{A}_{[0,T]}$ of controls. Let $\alpha_{[s,t]} \in \mathcal{A}_{[s,t]}$ be the strategy played over the time interval $[s, t] \subseteq [0, T]$. $\Gamma_{[s,t]}^\beta$ is a truncation of $\Gamma_{[0,T]}$ to the interval $[s, t]$, whereby the decision maker follows strategy β outside the time interval $[s, t]$. A strategy $\alpha^* \in \mathcal{A}$ is *weakly time consistent* if its reduction to the time interval $[t, T]$, $\alpha_{[t,T]}^*$, solves the continuation control problem $\Gamma_{[t,T]}^{\alpha^*}$ generated by itself, i.e., by $\alpha_{[0,t]}$, for all $t \in (0, T]$. A strategy $\alpha^* \in \mathcal{A}$ is *strongly time consistent* if its reduction to the time interval $[t, T]$, $\alpha_{[t,T]}^*$, constitutes an optimal solution of the continuation control problem $\Gamma_{[t,T]}^\beta$ generated by any other strategy $\beta_{[0,t]}$, for all $t \in (0, T]$. Both of these refinements assert that for an optimal solution to be time consistent, the decision maker should have no reason, at any future stage, to deviate from the adopted strategy. The strong form requires in addition that, even in case the actions taken or the state trajectory are different from what was initially expected, the players should have no reason to deviate from their prescribed actions going forwards. Trivially, the strong form implies the weaker one. In *deterministic* control problems, an optimal open-loop control is weakly but not necessarily strongly time consistent. In contrast, an optimal Markov control satisfies both time consistency forms. Here, it makes no material difference which type of control is selected. The notion of time consistency has application to both optimal (one-player) control problems as well as their extensions to multi-player settings as part of dynamic game theory. In *stochastic* control problems, however, a stand-alone decision maker should have the possibility to react to deviations from an expected market development. In dynamic *games*, if a rival may take actions over the course of the game that differ from what was initially expected, a player should be in a position to change her course of actions. In stochastic environments, a policy that is strongly time consistent should be preferred as at a future time a state realization might deviate from the state value that the decision maker planned to attain, owing to some random perturbations. When dealing with single-agent dynamic optimization problems under uncertainty, mathematical economists thus prefer Markov controls among possible admissible controls. In stochastic environments, strong time consistency of the solution is ensured if one employs *dynamic programming*. According to the principle of optimality, "an optimal policy has the property that, whatever the initial state and control are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" (see Bellman, 1957). In other words, the agents are supposed to act optimally on *every* trajectory of the state. In dynamic games, strong time consistency is achieved if the selected Nash equilibrium is also *perfect*: players are required to act optimally as part of a Nash equilibrium both on and off the equilibrium path. If the information structure allows economic agents to condition their actions on previous realizations, as in closed-loop strategies, then the resulting strongly time consistent policy may yield stronger predictions and reduce the set of optimal strategy candidates. In most cases, strongly time consistent solutions are best because agents are typically assumed to face negligible information lags. However, weakly time consistent policies are fine and easier to determine when information lags are infinite or agents precommit (thanks, e.g., to a binding contract) to a certain path of actions.

3. Solution approaches

Two approaches are usually followed to solve optimal control problems: dynamic programming and the maximum principle. We first provide heuristics for the Hamilton-Jacobi-Bellman equation, a partial differential equation used in dynamic programming. For the deterministic (resp., stochastic) case, dynamic programming leads to a first-order (resp., second-order) partial differential equation (PDE) together with a final boundary condition for the value function. We then discuss the stochastic maximum principle, which provides a set of necessary conditions. Alternatively, the deterministic maximum principle (resp., stochastic maximum principle) characterizes the solution by a maximum condition and two (resp., three) ordinary differential equations (resp., stochastic differential equations) for the state and the adjoint processes, together with boundary conditions for the initial value of the state process and the terminal value of the adjoint process.

3.1. Dynamic programming. A standard approach in control theory exploits the *recursive, embedded* nature of the problem. The subproblem $\Gamma(Y_t, t)$ is embedded or concatenated in the whole initial problem $\Gamma(y, 0)$. We denote restriction of a given admissible Markov strategy $\alpha_{[0, T]} \in \mathcal{A}_{[0, T]}$ to the subproblem $\Gamma(Y_t, t)$ by $\alpha_{[t, T]} \in \mathcal{A}_{[t, T]}$.

At the end period T , the subproblem $\Gamma(Y_T, T)$ is trivial, as given by the reward function:

$$(3.1a) \quad V(Y_T, T) = S(Y_T, T).$$

At a prior instant t , consider a short time interval of length Δ . If the agent takes an action $a \in A$ at time t , it is presumably committed to it until time $t + \Delta$. After this time, the agent faces no commitment constraint and can take action at any instant. For infinitesimal time interval, the continuation payoff is

$$\begin{aligned} J(Y_t, \alpha_{[t, T]}) &= \lim_{\Delta \downarrow 0} \left\{ \pi(Y_t, t, \alpha(t)) \Delta + \mathbb{E}_t \left[\int_{t+\Delta}^T \pi(Y_s, s, \alpha(s)) ds + S(Y_T, T) \right] \right\} \\ &= \lim_{\Delta \downarrow 0} \left\{ \pi(Y_t, t, \alpha(t)) \Delta + \mathbb{E}_t [J(Y_{t+\Delta}, t + \Delta, \alpha_{[t+\Delta, T]})] \right\} \end{aligned}$$

by the law of iterated conditional expectations. By taking the supremum,

$$\begin{aligned} V(Y_t, t) &= \lim_{\Delta \downarrow 0} \max_{\substack{a \in A \\ \alpha_{[t+\Delta, T]}}} \left\{ \pi(Y_t, t, a) \Delta + \mathbb{E}_t [J(Y_{t+\Delta}, t + \Delta, \alpha_{[t+\Delta, T]})] \right\} \\ &= \lim_{\Delta \downarrow 0} \max_{a \in A} \left\{ \pi(Y_t, t, a) \Delta + \max_{\alpha_{[t+\Delta, T]}} \mathbb{E}_t [J(Y_{t+\Delta}, t + \Delta, \alpha_{[t+\Delta, T]})] \right\} \\ &= \lim_{\Delta \downarrow 0} \max_{a \in A} \left\{ \pi(Y_t, t, a) \Delta + \mathbb{E}_t [V(Y_{t+\Delta}, t + \Delta, \alpha_{t+\Delta})] \right\} \end{aligned}$$

Hence, dividing the above equation by $\Delta (> 0)$, the optimality condition becomes:

$$\lim_{\Delta \downarrow 0} \max_{a \in A} \left\{ \pi(Y_t, t, a) + \frac{\mathbb{E}_t [V(Y_{t+\Delta}, t + \Delta)] - V(Y_t, t)}{\Delta} \right\} = 0.$$

For tractability, one generally conjectures that $V \in C^{2,1}(\mathcal{Y} \times [0, T])$. Using the infinitesimal generator notation, one obtains:

$$(3.1b) \quad \max_{a \in A} \{ \pi(Y_t, t, a) + \mathcal{D}V(Y_t, t) \} = 0$$

with boundary (terminal) condition (3.1a). Under certain circumstances, the maximum might be attained, in which case we note

$$\pi(Y_t, t, \alpha^*(t)) + \mathcal{D}V(Y_t, t) = 0, \quad t \in [0, T].$$

This partial differential equation (PDE) is the continuous-time variant of the Bellman equation. It is called *Hamilton-Jacobi-Bellman* or HJB equation. The existence of a unique solution is often difficult to prove because the Markov control depends on the value dynamics without usually being an explicit function. The existence of well-behaved solutions in a continuous-time setting is delicate and we shall focus mainly on *sufficient* conditions. We can conjecture an optimal Markov control policy

$\phi : \mathcal{Y} \times [0, T] \rightarrow A$ that solves equations (3.1a)–(3.1b) and attempt to verify its optimality. A “verification argument” implies that if one finds a (twice continuously differentiable) solution to the HJB equation (4.5), then that solution must equal the value function of the control problem, whereby the control that solves the HJB equation. Under strategy ϕ , the *optimum state trajectory*, Y^* , solves the stochastic differential equation

$$dY_t = \mu(Y_t, t, \phi(Y_t, t)) dt + \sigma(Y_t, t, \phi(Y_t, t)) dB_t; \quad Y_0 = y \quad \text{a.s.}$$

The above analysis also works under certain condition to infinite-horizon problems, whereby the boundary condition is replaced by a *transversality condition*, such as:

$$(3.1c) \quad \lim_{T \rightarrow \infty} \mathbb{E}V(X_T, T) = 0.$$

Note that the above transversality condition is satisfied when one introduces a (positive) discount rate.

3.2. Maximum principle. The stochastic maximum principle builds upon Pontryagin’s dynamic optimization approach. Originally, the approach – which involves a dual process called *adjoint process* – was developed for *deterministic* environments. The maximum principle provides a set of *necessary* conditions for the optimal state trajectory and the corresponding adjoint equation. The original formulation of the maximum principle involved open-loop controls but, as noted, in deterministic, one-agent setting, the values obtained under the different class of controls do not differ. The maximum can be either proved directly by use of variational arguments or indirectly as an extension of the (first-order) HJB equation. Since the inception of Pontryagin’s maximum principle, research has been undertaken to allow for Markov controls and find an equivalent maximum principle applying in stochastic environments. We, here, build upon the (second-order) HJB equation derived previously.

3.2.1. Deterministic maximum principle. Consider first the deterministic (one-player) case whereby the optimal open-loop control policy is noted $\psi : [0, T] \rightarrow A$. Here, the payoff function under finite control horizon reads

$$(3.2) \quad J(y, t, \alpha) = \int_0^T \pi(Y_t, t, \alpha(t)) dt + S(Y_T, T)$$

with $Y_0 = y \in \mathcal{Y} \subseteq \mathbb{R}^n$ and state equation

$$(3.3) \quad \dot{y}(t) = \mu(y, t, \alpha(t)) \quad t \in [0, T].$$

For the sake of argumentation, we now digress and consider that the decision maker intends to maximize at each instant $t \in [0, T]$ her utility flow $\pi(Y_t, t, \alpha(t))$ in a static manner. The state equation (1.1) imposes then a set of (n) constraints on the optimization program. In static optimization, a Lagrange multiplier corresponds to the increase in the value of the objective function due to the relaxation of a given constraint. With each state $(y, t) \in \mathcal{Y} \times [0, T]$, we associate an *adjoint* or *costate* variable, $\lambda_1(y, t)$, of the same dimension as $y \in \mathcal{Y}$ that corresponds to a vector of n Lagrange multipliers. In other words, the costate variable is a dual representing the marginal cost of violating the constraints imposed by the state equation. A dynamic equivalent of the Lagrangian of a static optimization problem is the current-value *Hamiltonian function* defined as a real-valued function H :

$$(3.4) \quad H(t, y, a, \lambda_1) \triangleq \pi(t, y, a) + \lambda_1(y, t)^\top \mu(t, y, a).$$

By differentiation the Hamiltonian function (3.4) it obtains a new statement of the state equation, namely

$$(3.5) \quad \dot{y}(t) = H_{\lambda_1}(t, y, a, \lambda_1).$$

In the deterministic case, the HJB equation is a first-order partial differential equation:

$$V_t(y, t) = \max_{a \in A} \{ \pi(y, t, a) + V_y(y, t)^\top \mu(y, t, a) \}.$$

Considering the Hamiltonian in equation (3.4), we may state the HJB equation as

$$(3.6) \quad V_t(y, t) = \max_{a \in A} H(y, t, a, V_y(y, t)).$$

Therefore, the optimal open-loop control, $\psi(\cdot)$, satisfies at each time $t \in [0, T]$ the *maximum condition*:

$$(3.7) \quad \psi(t) \in \arg \max_{a \in A} H(t, y, a, \lambda_1(t)).$$

Equation (3.6) expresses that on the optimal state trajectory the equality $\lambda_1(y, t) \triangleq V_y(y, t)$ always holds. The costate vector $\lambda_1(y, t)$ should then be interpreted as the first-order derivative or gradient of the value function $V(y, t)$, namely

$$(3.8) \quad \lambda_1(y, t) = \begin{pmatrix} \frac{\partial V}{\partial y_1}(y, t) \\ \vdots \\ \frac{\partial V}{\partial y_n}(y, t) \end{pmatrix}.$$

The costate variable $\lambda(y, t)$ has thus an interesting economic interpretation. It corresponds to a *shadow price* or marginal value of the resource, namely the rate of change in value with respect to a change in the resource (state variable). In other words, in equilibrium it represents the highest hypothetical price a decision maker would be willing to pay for an additional, infinitesimal unit of the state variable (at time t in state y).

The merit of the maximum principle lies in the economic interpretation of the maximum condition. For a dynamical system, the control a taken at a given instant $t \in [0, T]$ should not be aimed at maximizing the instantaneous profit $\pi(y, t, \cdot)$. If, at each instant $t \in [0, T]$, the decision maker maximizes the direct instantaneous contribution $\pi(Y_t, t, \cdot)$ over A , she would not necessarily maximize the overall intertemporal payoff $J(y, \cdot)$ over \mathcal{A} because the static program $\pi(Y_t, t, \cdot)$ fails to reflect, at time $t \in [0, T]$, the intertemporal effect of the control choice $\alpha(t)$ on the state equation (1.1). By using the shadow price of the state we can explicitly consider the dynamic tradeoff in a new static optimization program, the one given by the Hamiltonian function of equation (3.4). The Hamiltonian $H(y, t, a)$ offers a simple way of altering the instantaneous objective function $\pi(Y_t, t, a)$ to take into account the future consequences of the choice of the control. The amount $H(y, t, a)$ signifies the sum of the *direct* contribution $\pi(y, t, a)$ to the overall profit and the accumulation rate $\mu(t, y, a)$ of the marginal value of the resource, $\lambda_1(y, t)$. In other words, the second term, $\lambda_1(y, t)^\top \mu(y, t, a)$, represents the *indirect* instantaneous monetary contribution to the payoff, as each additional unit of state yields a value increase of $\lambda_1(y, t)$ dollars and $\mu(y, t, a)$ units are added in total. $H(y, t, a)$ is thus the *total* instantaneous contribution, where the second term properly values the indirect effect. According to the maximum condition (3.7), the decision maker desires to make this sum as large as possible and is faced with the static maximization problem $H(Y_t, t, \cdot, \lambda_1(Y_t, t))$ at each instant $t \in [0, T]$ she is allowed to act. The maximum principle thus decouples the dynamic optimization problem (3.2) in the interval $[0, T]$ into a set of static maximization problems indexed by $t \in [0, T]$, namely the Hamiltonian maximization programs (3.6).

The maximum principle also characterizes the dynamics of the dual variable via the *adjoint equation*. By differentiating totally the expression in equation (3.8), we obtain on the optimal path

$$\begin{aligned} \dot{\lambda}(t) &= \frac{dV_y}{dt}(y(t), t) = V_{yy}(y(t), t)^\top \dot{y}(t) + V_{yt}(y(t), t) \\ &= V_{yy}(y(t), t)^\top \mu(y(t), t, \psi(t)) + V_{yt}(y(t), t). \end{aligned}$$

For V sufficiently smooth, we have $V_{yt} = V_{ty}$. The above yields

$$(3.9) \quad \dot{\lambda}(t) = -H_y(t, y, \psi(t), \lambda(t)).$$

The adjoint equation (3.9) also has a useful economic interpretation. A marginal unit of capital yields a direct, instantaneous marginal return $\pi_y(y, t, a)$. Besides, it generates in the next period an extra change in state $\mu_y(y, t, a)$, each unit being currently valued at λ_1 . The change in shadow price, $\dot{\lambda}_1(t)$, is like a capital gain expressed in present value terms. On the optimal path, the total marginal return as the sum of these components should be zero. In other words, on the optimal state trajectory, the shadow price

takes on values that do not permit a pure or excess return from holding the stock. The adjoint equation (3.9) expresses an intertemporal no-arbitrage condition. By integrating the adjoint equation, it obtains

$$\lambda(t) = - \int_t^T H_y(Y_s, s, a_s, \lambda_s) ds + S_y(Y_T, T).$$

The last term expresses that the price $\lambda_1(T)$ of a unit of capital at time T is its marginal salvage value $S_y(Y_T, T)$. The above simply means that, along the optimal state trajectory, the (instantaneous) depreciation rate of the shadow price is the sum of its contribution to paid-out profit, $\pi_y(t, Y_t^*, \alpha_t^*)$, and its contribution to the enhancement of the marginal value of the resource, $\mu_y(t, Y_t^*, \alpha_t^*)$. The maximum principle relates to the so-called *duality principle*: “maximizing the total value amounts to minimizing the total depreciation of the marginal value.” The adjoint function $\lambda_1(\cdot)$ evolves backwards in time, solving the *adjoint equation* – an ordinary differential equation – provided a given terminal condition, namely

$$(3.10) \quad \lambda_1(T) = S_y(T, Y_T^*).$$

The system consisting of the the state equation (3.5), the adjoint equation (3.9), and the maximum condition (3.7) is referred to as an extended *Hamiltonian system*. The deterministic *maximum principle* asserts that the Hamiltonian system provides a set of necessary conditions that characterize the optimal triple (Y^*, ψ, λ^*) . If certain convexity conditions hold, the Hamiltonian system characterizes an optimal control. To obtain a sufficiency theorem, we generally make additional convexity assumptions.

3.2.2. Stochastic maximum principle. In the above statement of the deterministic maximum principle, the presence of an adjoint process that evolves backward in time makes the extension of the maximum principle to stochastic environments nontrivial because of the nonanticipativity constraint (\mathbb{F} -adaptedness). Besides, we need to refine the space of admissible controls to Markov controls to circumvent the problem of time inconsistency involved with open-loop controls. The (recent) advent of *backward stochastic differential equations* (BSDE) – SDE with terminal boundaries – provided a natural framework for handling this problem. The key computational issue is that the adjoint equations are solved backwards but their solutions are required to be nonanticipative. Recall that the objective function for the stochastic case reads

$$J(y, 0, \alpha) = \mathbb{E} \left[\int_0^T \pi(Y_t^\alpha, t, \alpha(t)) dt + S(Y_T, T) \mid y \right],$$

subject to state equation (1.1). We introduce new adjoint variables together with their adjoint equations that are involved in the formulation of the stochastic maximum principle. We now generalize the notion of the Hamiltonian to a stochastic environment. A time t and state y , the current-state generalized *Hamiltonian* is defined as:

$$(3.11) \quad \begin{aligned} \mathcal{H}(y, t, a, \lambda_1, \lambda_2) &\triangleq \pi(y, t, a) + \lambda_1(y, t)^\top \mu(y, t, a) + \frac{1}{2} \text{tr} \left\langle \sigma(y, t, a)^\top \lambda_2(y, t) \sigma(y, t, a) \right\rangle, \\ &= H(y, t, a, \lambda_1) + \frac{1}{2} \text{tr} \left\langle \sigma(y, t, a)^\top \lambda_2(y, t) \sigma(y, t, a) \right\rangle \end{aligned}$$

The first term on the right-hand side of equation (3.11) was already consider for the deterministic Hamiltonian function (3.4). As noted it signifies the sum of the expected direct contribution to the overall profit and the expected accumulation rate of the marginal value of the resource. The second term is new. It reflects the riskiness of the venture, i.e., it represents the reduction/addition in expected marginal value of the resource due to the decision maker’s uncertainty about the anticipated change in resource. This third term is called risk-adjustment. If V is convex (in case, e.g., the decision maker is risk-averse), the risk-adjustment effect reduces the value of current-value generalized Hamiltonian. The opposite holds if V is concave.

Assume the Markov control policy ϕ is optimal. Using the notations of equation (3.11), the second-order HJB equation (3.1) becomes

$$(3.12) \quad V_t(y, t) = \max_{a \in A} \mathcal{H}(y, t, a, V_y(y, t), V_{yy}(y, t)).$$

The *maximum condition* is thus

$$(3.13) \quad \phi(y, t) \in \arg \max_{a \in A} \mathcal{H}(y, t, a, \lambda_1, \lambda_2).$$

Besides, from equation (3.12), the first-order adjoint variable should be interpreted as the gradient of the value function $V(y, t)$, $\lambda(y, t) = V_y(y, t)$. By differentiating the current-state Hamiltonian in equation (3.11) with respect to λ_1 , it obtains

$$(3.14) \quad \mathcal{H}_{\lambda_1}(y, t, a, \lambda_1, \lambda_2) = \mu(y, t, a).$$

The adjoint variable λ_1 corresponds in the deterministic case with the so-called shadow price. Nonetheless, in a stochastic environment, the marginal value alone may not be able to fully characterize the true contribution/influence of the control on the objective function. One has to introduce another variable to reflect the uncertainty or risk factor in the system. It similarly obtains from (3.11)

$$(3.15) \quad \mathcal{H}_{\lambda_2}(y, t, a, \lambda_1, \lambda_2) = \frac{1}{2} \sigma(y, t)^\top \sigma(y, t).$$

It follows from (3.8) that

$$(3.16) \quad \lambda_2(y, t) = \frac{\partial \lambda_1}{\partial y}(y, t) = \sigma(y, t, a) V_{yy}(y, t).$$

In other words, λ_2 is the instantaneous standard deviation of the depreciation rate of the random shadow price.

We want to derive a system of stochastic differential equations describing the optimal state trajectory $Y^* = \{Y_t^*\}_{t \geq 0}$ and adjoint equations λ_1 and λ_2 on the optimal path. From equation (3.14), the optimal state trajectory $\{Y_t^*\}_{t \geq 0}$ solves the forward stochastic differential equation

$$(3.17) \quad dY_t = H_{\lambda_1}(Y_t, t, \phi(Y_t, t), \lambda_1, \lambda_2) dt + \sigma(Y_t^*, t, \phi(Y_t, t)) dB_t,$$

subject to $Y_0 = y \in \mathcal{Y}$ a.s. From HJB equation (3.17) and Itô-Doebelin formula,

$$(3.18) \quad dV_y(Y_t, t) = \mathcal{D}V_y(Y_t, t) dt + V_{yy}(Y_t, t) + \sigma(Y_t, t, a) dB_t,$$

where \mathcal{D} is the infinitesimal generator. Equivalently stated, the *first-order adjoint equation* is a backward stochastic differential equation given by

$$(3.19a) \quad d\lambda_1(Y_t, t) = \mathcal{D}\lambda_1(Y_t, t) dt + \lambda_y(Y_t, t) + \sigma(Y_t, t, a) dB_t,$$

with transversality condition

$$(3.19b) \quad \lambda_1(T) = -S_y(Y_T, T).$$

To account for the influence of the control on the diffusion coefficient, we consider an additional *second-order adjoint equation*, given by

$$(3.20a) \quad d\lambda_2(t) = - \left\{ \lambda_2(t)^\top \mu_y(y, t, a) + \lambda_2(y, t)^\top \mu_y(y, t, a) + \sum_{j=1}^n \sigma_y^j(y, t, a)^\top \lambda_2^j(y, t) \sigma_y^j(y, t, a) \right. \\ \left. + \sum_{j=1}^n \sigma_y^j(y, t, a)^\top \Lambda_2^j(y, t) + \Lambda_2^j(y, t) \sigma_y^j(y, t, a) + H_{yy}(y, t, a, \lambda_1, \lambda_2) dt + \sum_{j=1}^n \Lambda_2^j(y, t) dB_t^j \right\}$$

with terminal condition

$$(3.20b) \quad \lambda_2(T) = -S_{yy}(Y_T, T).$$

In the above the unknown is a pair of processes (λ_2, Λ_2) that must be \mathbb{F} -adapted.

A stochastic Hamiltonian system consists, along with the maximum condition (3.13), of a n -dimensional forward SDE (3.17) for the original state equation and two adjoint n -dimensional backward SDEs —

one for each adjoint variable. The *stochastic maximum principle* asserts that under certain conditions, there exists for every admissible Markov control costate trajectories λ_1 and λ_2 such that the maximum condition (3.13), (backward stochastic differential) adjoint equations (3.19a) and (3.20a), and the transversality conditions (3.19b) and (3.20b) are satisfied almost everywhere.

PROPOSITION 1 (Stochastic maximum principle). *Suppose the Markov control ϕ^* solves the dynamic optimization problem $\Gamma(y, 0)$. Then, there exist a state trajectory $Y^* = \{Y_t^*\}_{t \leq 0}$ and adjoint processes λ_1^* and λ_2^* such that for each $t \in [0, T]$:*

- (i) *the maximum condition (3.13) is satisfied on the optimal state trajectory;*
- (ii) *Y^* solves equation (3.17) with initial boundary condition $Y_0^* = y$ a.s.;*
- (iii) *λ_1 solves equation (3.19a) with terminal boundary condition (3.19b);*
- (iv) *λ_2 solves equation (3.20a) with terminal boundary condition (3.20b).*

4. Other stochastic control problems

Stochastic control theory is richer than sketched previously. A number of situations encountered in mathematical economics correspond to refinements of the control problem defined previously.

4.1. Cauchy problem and Feynman-Kac representation. Consider first the polar case in which no control is exercised and the horizon is finite and of known duration. As before, we assume, on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, a d -dimensional Itô process X that solves the stochastic differential equation (3.1) and such that $X_0 = x \in \mathcal{X}$ a.s. A function $r : \mathcal{X} \times [0, T] \rightarrow \mathbb{R}_+$ determines for each state the *discount rate*. The value function — which is of the Bolza type — reads

$$(4.1) \quad V(x, t) = \mathbb{E} \left[\int_t^T D^{x,t}(s) \pi(X_s, s) ds + D^{x,t}(T) S(X_T) \mid x, t \right],$$

where D stands for the stochastic discount factor:

$$(4.2) \quad D^{x,t}(s) \triangleq \exp \left(- \int_t^s r(X_s, s) ds \right), \quad s \in [t, T].$$

According to the *Feynman-Kac theorem* (see proof in Karatzas and Shreve, 1991, 5.7.C), as long as the solution of (4.1) is sufficiently smooth, there exists a correspondence between that value and the solution of a certain Cauchy problem. Here, the *Cauchy problem* consists in finding a function V , in the class $C^{2,1}(\mathcal{X}, [0, T])$, that solves the partial differential equation

$$(4.3a) \quad r(x, t) V(x, t) = \pi(x, t) + \mathcal{D}V(x, t),$$

everywhere on $\mathcal{X} \times [0, T)$ and also satisfies the terminal condition

$$(4.3b) \quad V(X_T, T) = S(X_T).$$

More precisely, the Feynman-Kac theorem gives a probabilistic solution, namely (4.1), to the functional analysis problem of equations (4.3a)–(4.3b). It also asserts that the value function (4.1) admits a functional representation (4.3a)–(4.3b).

The Feynman-Kac solution to the Cauchy problem is oftentimes used in mathematical economics and continuous-time finance (see, e.g., Duffie, 2001). In economic terms, equation (4.3a) states an equilibrium condition to be satisfied in each state $x \in \mathcal{X}$ and at each time $t \in [0, T)$: the required return on value $r(x, t) V(x, t)$ must equate the instantaneous reward flow $\pi(x, t)$ plus any additional capital gain flow $\mathcal{D}V(x, t)$. This equilibrium condition is relatively restrictive: it asserts that this equality holds not only on the optimal state path but also on all other feasible state trajectories. The Black-Scholes-Merton partial differential equation can be derived and interpreted in this light; this p.d.e. corresponds to equation (4.3a) in the Cauchy problem, with no dividend payment or $\pi = 0$ (Mayer form). Thereby, the salvage value function (4.3b) is given by the payoff function of a European call option with strike price I , $S(x) = (x - I)^+$. In some other cases, the Feynman-Kac partial differential equation (4.3a) can be

used as the basis for a (implicit or explicit) finite-difference numerical scheme that describes an asset price dynamics (see e.g., Brennan and Schwartz, 1977).

4.2. Autonomous problem. We would also like to discuss problems with *infinite-horizon* ($T = \infty$) where $S = 0$ and the functions π , μ , σ , and α do not depend on time. Such problems — called *autonomous* — are often easier to solve analytically than their counterparts with finite horizon and time-dependent processes and controls. We consider a (constant) discount rate $r (> 0)$ and a state equation

$$dY_t = \mu(Y_t, a) dt + \sigma(Y_t, a) dB_t,$$

with $Y_0 = y \in \mathcal{Y}$ a.s. and $\mathcal{Y} = \mathbb{R}^n$. We use “tilde” for time-homogenous functions. The profit function is $\pi(y, t, a) = e^{-rt} \tilde{\pi}(y, a)$. We now consider the payoff function

$$J(Y_t, t, \alpha_{[t, \infty)}) = \mathbb{E} \left[\int_t^\infty e^{-rs} \tilde{\pi}(Y_s, a_s) ds \mid Y_t \right]$$

subject to state equation (1.1) and $Y_0 = y$ a.s. The value to the decision maker — if it exists — is $V(y) = \max_{\alpha \in \mathcal{A}} J(y, 0, \alpha)$. As noted the presence of a discount rate ensures that the transversality condition (3.1c) is satisfied. In this case, it is natural to consider stationary Markov controls. Define the time-homogenous payoff and value functions by

$$\tilde{J}(y, \alpha) \triangleq \mathbb{E} \left[\int_0^\infty e^{-rt} \tilde{\pi}(Y_t, \alpha(t)) dt \mid y \right] \quad \text{and} \quad \tilde{V}(y) \triangleq \max_{\alpha \in \mathcal{A}} \tilde{J}(y, \alpha).$$

We have $J(Y_t, t, \alpha) = e^{-rt} \tilde{J}(Y_t, \alpha)$. By taking the supremum, it follows $V(y, t) = e^{-rt} \tilde{V}(y)$, the latter function admitting as first- and second-order derivatives $V_t(y, t) = -re^{-rt} \tilde{V}(y)$, $V_y(y, t) = e^{-rt} \tilde{V}_y(y)$, and $V_{yy}(y, t) = e^{-rt} \tilde{V}_{yy}(y)$. The infinitesimal generator yields now

$$\begin{aligned} \mathcal{D}V(y, t) &= V_t(y, t) + V_y(y, t)^\top \mu(y, a) + \frac{1}{2} \text{tr} \left\langle \sigma(y, a)^\top V_{yy}(y, t) \sigma(y, a) \right\rangle \\ &= -re^{-rt} \tilde{V}(y) + e^{-rt} \tilde{V}_y(y)^\top \mu(y, a) + \frac{1}{2} e^{-rt} \text{tr} \left\langle \sigma(y, a)^\top \tilde{V}_{yy}(y) \sigma(y, a) \right\rangle. \end{aligned}$$

The HJB equation in (3.1b) is then

$$\max_{a \in A} \left\{ e^{-rt} \tilde{\pi}(y, a) - re^{-rt} \tilde{V}(y) + e^{-rt} \tilde{V}_y(y)^\top \mu(y, a) + \frac{1}{2} e^{-rt} \text{tr} \left\langle \sigma(y, a)^\top \tilde{V}_{yy}(y) \sigma(y, a) \right\rangle \right\} = 0.$$

Multiplying the above expression by e^{rt} yields a time-homogenous variant of the HJB equation:

$$r \tilde{V}(y) = \max_{a \in A} \left\{ \tilde{\pi}(y, a) + \tilde{V}_y(y)^\top \mu(y, a) + \frac{1}{2} \text{tr} \left\langle \sigma(y, a)^\top \tilde{V}_{yy}(y) \sigma(y, a) \right\rangle \right\}.$$

This second-order partial differential equation has wide application in economics and continuous-time finance.

4.3. Random control horizon and optimal stopping. The problem may involve a *random control horizon*. Consider a given open set $A \subseteq \mathcal{Y}$. $Y = \{Y_t\}_{t \geq 0}$ solves again the stochastic differential equation (1.1) with $Y_0 = y \notin A$. Pathwise, the time at which the state enters the set A for the first time, noted $\tau(\omega, \alpha) \triangleq \inf \{t \geq 0 \mid Y_t^\alpha(\omega) \in A\}$, depends on the control α chosen by the decision maker via the influence of the control on the state trajectory $Y^\alpha(\omega, \cdot)$. In this case, the objective function is

$$J(y, 0, \alpha) = \mathbb{E} \left[\int_0^{\tau(\alpha)} \pi(Y_t^\alpha, t, \alpha) dt + S(\tau(\alpha), Y_{\tau(\alpha)}) \mid \mathcal{F}_0 \right].$$

In the above, the control influences the stopping time $\tau(\alpha)$ only indirectly via the controlled state evolution. Alternatively, the controller may decide directly on the control horizon — via the choice of a stopping time τ in an admissible set \mathcal{T} — as part of an *controlled optimal stopping problem*; here, an optimal pair of control and stopping time (α^*, τ^*) verifies

$$V(y, 0) = \max_{\substack{\alpha \in \mathcal{A} \\ \tau \in \mathcal{T}}} \mathbb{E} \left[\int_0^\tau \pi(t, Y_t^\alpha, \alpha(t)) dt + S(Y_\tau, \tau) \mid y \right].$$

In most cases studied in economics, the state space is one-dimensional ($\mathcal{Y} \subseteq \mathbb{R}$) and the payoff function is

$$J(y, 0, \alpha, \mathcal{F}) = \mathbb{E} \left[\int_0^\tau D^y(t) \pi(Y_t^\alpha, t, \alpha(t)) dt + D^y(\tau) S(Y_\tau, \tau) \mid y \right],$$

where the discount factor function $D^y(\cdot)$ is a time-homogenous variant of (4.2). If the value function is sufficiently smooth, a heuristic argument implies the HJB equation

$$(4.4) \quad 0 = \max_{a \in A} \{S(y, t) - V(y, t); -r(y, t) V(y, t) + \pi(y, t, a) + \mathcal{D}V(y, t)\}.$$

We face an *optimal stopping problem* if the decision maker exercises no control (α) on the state. In this case, we denote the (one-dimensional) exogenous shock process by X . The HJB equation here simplifies to

$$(4.5) \quad 0 = \max \{S(x, t) - V(x, t); -r(x, t) V(x, t) + \pi(x, t) + \mathcal{D}V(x, t)\}, \quad (x, t) \in \mathcal{X} \times [0, T].$$

where the second-order infinitesimal generator is

$$\mathcal{D}V(x, t) = V_t(x, t) + \mu(x, t) V_x(x, t) + \frac{1}{2} \sigma(x, t)^2 V_{xx}(x, t).$$

One generally considers an autonomous version of the optimal stopping problem. Here, in state $x \in \mathcal{X}$, the HJB equation reads

$$(4.6) \quad 0 = \max \left\{ S(x) - V(x); -r(x) V(x) + \pi(x) + \mu(x) V_x(x) + \frac{1}{2} \sigma(x)^2 V_{xx}(x) \right\}.$$

The optimal stopping problem involves solving the ODE (4.6) or PDE (4.5) subject to certain boundary conditions.

If the decision-maker may impose additional boundary conditions, optimal stopping involves solving a “free boundary problem” associated with such ODEs and PDEs. When, for an autonomous problem, we concentrate on stationary Markov controls, it is reasonable to partition the state space into an optimal *stopping region* $\mathcal{Z} = \{x \in \mathcal{X} \mid S(x) = V(x)\}$ and a *continuation region* $\mathcal{X} \setminus \mathcal{Z}$. The boundary of the set \mathcal{Z} is $\partial\mathcal{Z}$. HJB equation (4.6) and the terminal condition then correspond to the *Dirichlet problem*:

$$(4.7a) \quad r(x) V(x) = \pi(x) + \mu(x) V_x(x) + \frac{1}{2} \sigma(x)^2 V_{xx}(x), \quad x \notin \mathcal{Z}$$

$$(4.7b) \quad V(x) = S(x), \quad x \in \mathcal{Z}.$$

Boundary condition (3.7b) is often called *value-matching* condition. The random variable $\tau^* : \omega \mapsto \tau(\omega) = \inf\{t \geq 0 \mid X_t(\omega) \in \mathcal{Z}\}$ is the optimal first-entry time in \mathcal{Z} . If the initial state $x \in \mathcal{Z}$, the firm stops right at the outset almost surely. An optimality condition often encountered in optimal stopping problems impose a constraint on the first-order derivative, V_x , of the value function (see, e.g., Dumas, 1991). More precisely, this condition — known as *smooth-pasting*, high contact, or smooth-fit principle — states that V_x is continuous at the boundary z of the continuation set $\mathcal{X} = \mathbb{R}$. The smooth-pasting condition can be heuristically obtained from the envelope theorem. Let $V(x, z)$ be the solution of the optimal-stopping problem. $\tau^*(\omega) \triangleq \inf\{t \geq 0 \mid X_t(\omega) \geq z\}$ is the first time on path $\omega \in \Omega$ that the process X exceeds the upper threshold z . Following the value matching condition, for $x \in \mathcal{Z}$ and especially for $x = z$, the firm is indifferent between exercising no control earning $V(x, z)$ and stopping receiving $S(x)$. By total differentiation with respect to z , we have $V_x(z, z) + V_z(z, z) = S_x(z)$. Since z is optimally chosen, $V_z(z, z) = 0$. Hence, the smooth-pasting condition:

$$(4.7c) \quad V_x(z, z) = S_x(z).$$

4.4. Other control problems. A case of interest is the (stochastic) *linear quadratic* case, where the state equation is linear in the state and control and the objective function is quadratic in both the state and the control. Closed-form solutions are often obtained for such control problems.

Consider now a control in the form of a càdlàg process ξ with bounded variation and such that $\mathbb{E}[|\xi|_T] < \infty$. The process ξ is made of a continuous part ξ^c and a pure-jump part ξ^j . On a given path $\omega \in \Omega$, a jump occurring at time s is noted $\Delta\xi_s(\omega) \triangleq \xi_s(\omega) - \xi_{s-}(\omega)$, with the pure-jump part of the control being $\xi_t^j(\omega) \triangleq \sum_{0 \leq s \leq t} \Delta\xi_s(\omega)$. The continuous part is uniquely decomposed into an absolutely continuous part ξ^{ac} and a singularly continuous part ξ^{sc} , $\xi^c = \xi^{ac} + \xi^{sc}$. *Singular control* deals with problems involving an additive control term in the stochastic integral that characterizes the state evolution:

$$Y_t = y + \int_0^t \mu(s, Y_s) ds + \int_0^t \sigma(s, Y_s) dB_s + \xi(t), \quad t \in [0, T].$$

As noted, the control ξ is not necessarily absolutely continuous with respect to the Lebesgue measure but it has bounded variations. This leads to some technical issues that will be discussed in Chapter 5. For the general singular control problem, the objective function is

$$J(y, 0, \xi) = \mathbb{E} \left[\int_0^T \left\{ \pi(Y_t, t) + f^{ac}(t) \|\dot{\xi}^{ac}(t)\| \right\} dt + \int_0^T f^{sc}(t) |d\xi^{sc}(t)| + \sum_{\Delta\xi_t \neq 0} l(\Delta\xi_t, t) + S(Y_T, T) \mid \mathcal{F}_0 \right],$$

where $\|\cdot\|$ and $|\cdot|$ are the measures generated by the total variations of ξ^{ac} and ξ^{sc} . If ξ is a pure jump process, i.e., if $\xi^{ac} = \xi^{sc} = 0$, we face an *impulse control* problem. Here, equivalently, a decision maker may change, at distinct stopping times $\{\tau_i\}$, the state value by a discrete amount from an initial value $Y(\omega, \tau_i^-)$ to a new value $Y(\omega, \tau_i)$ with associated cost $l(\tau_i(\omega), \Delta\xi_{\tau_i}(\omega))$. Stochastic singular control and optimal stopping are related in the one-dimensional case, as proved in Karatzas and Shreve (1984) and discussed further in Chapter 5. In *ergodic control* problems, the decision maker desires to optimize a long-run average performance criterion:

$$\lim_{T \rightarrow \infty} \max_{\alpha \in \mathcal{A}} \frac{1}{T} \int_0^T \mathbb{E}[\pi(Y_t, t, \alpha_t) \mid y] dt.$$

5. References

Harrison (1985), Dixit (2002), and Stokey (2008) summarize useful properties on several continuous-time processes and discuss how to apply them in economic contexts involving stochastic control models. Merton (1992) and Duffie (2001) apply notions of stochastic control to mathematical finance. Øksendal (2007) elaborates on the different types of controls. Başar and Olsder (1999) define the notion of weak and strong time consistency. Yong and Zhou (1999) discuss the relationship between the maximum principle and dynamic programming in stochastic environments and derive the stochastic maximum principle in settings where the control affects the diffusion term in the state equation. Borkar (2005) offers an overview on controlled diffusion processes and discusses some further control problems. Peskir and Shiryaev (2006) focus on optimal stopping problems.

Dynamic Game Theory

Game theory is the branch of applied mathematics aiming at describing and predicting behaviors in environments in which actors are interdependent and have potentially conflicting objectives. Thus described, game theory may help derive insights in all social sciences and beyond, with applications in economics, political sciences, sociology, psychology, law, biology. We herein provide a brief account of game theory focusing on those notions and solution concepts that are most relevant to our research. In particular, we discuss differential games which are state-space games cast in continuous time and which share common characteristics with control problems.

The chapter is organized as follows. In Section 1, we describe what constitutes a normal-form game with a focus on defining (stochastic) differential games. Section 2 discusses the a critical model assumption relating to the choice of information structure or strategy space. In Section 3, we briefly present two solution concepts used in dynamic game theory: Nash equilibrium and perfect Markov equilibrium. Sections 4 and 5 discuss, respectively, (nonzero-sum) differential and stochastic differential games. Section 6 gives references to advanced literature on dynamic game theory.

1. Game description

A game in *normal form* has three basic elements:

- (i) a set of players $N = \{1, \dots, n\}$ with an arbitrary player noted $i \in N$;
- (ii) a strategy set \mathcal{A}_i for each player $i \in N$;
- (iii) a payoff function J^i for each player $i \in N$.

A game consisting of these three elements is represented by $\Gamma = \{N, (\mathcal{A}_i)_{i \in N}, (J^i)_{i \in N}\}$. A *player* is an individual, firm, entity, or actor that makes decisions. An additional *pseudo-player*, such as nature, may be considered to account for or explain outside exogenous factors. This is the case in stochastic environments where “nature” selects the state of the world at random irrespective of the players’ actions. In most cases, the probability distribution of nature’s moves is known to players. A *strategy* is a contingent plan of actions for all possible time and history/state. Let $\alpha_i \in \mathcal{A}_i$ denote firm i ’s pure strategy and \mathcal{A}_i player i ’s strategy set. By convention, α_{-i} represents the strategies played by all other players except i . A *strategy profile* α encompasses the strategies pursued by all the players, namely $\alpha = (\alpha^1, \dots, \alpha^n)^T$. $\mathcal{A} = \otimes_{i=1}^n \mathcal{A}_i$ is the strategy space. Each strategy combination, $\alpha \in \mathcal{A}$, results in a specific expected payoff for each player, $J^i(\alpha)$. Here, all players know the structure of the strategic form and know that their opponents know it, and know that their opponents know they know it, and so on ad infinitum. In other words, the structure of the game is *common knowledge*.

Dynamic games model strategic situations involving a certain notion of time. They have been extensively used in industrial organization but are loosely defined mathematical objects when one considers the choice of the strategy space. In *dynamic game theory* (see, e.g., Başar and Olsder, 1999), the formal description of the game — understood as a dynamical system — and the definition of the appropriate strategy space play a key role. To avoid the curse of dimensionality potentially occurring when the entire game history affects the current situation and action choices, a modeler may adopt more restrictive assumptions about the game evolution. In a *state-space game*, a (n -dimensional) variable describes at any particular instant t the state $Y_t \in \mathcal{Y} \subseteq \mathbb{R}^n$ of the game together with a state equation that

describes the evolution of the dynamical system over time. In other words, the payoff-relevant influence of past events and actions are assumed adequately summarized in the state variable. A *differential game* is a state-space game considered in continuous time, whereby the states evolve following a set of differential equations. Differential games originated in the work of Isaacs (1965) as an extension of optimal control theory considered in continuous time, whereby analytical tools, such as the maximum principle or dynamic programming, are also utilized. Differential games constitute a class of decision problems wherein the evolution of the state is described by a differential equation and the players may act continuously throughout a time interval. Formally, a differential game — expressed in normal form $\Gamma = \left\{ N, (\mathcal{A}_i)_{i \in N}, (J^i)_{i \in N} \right\}$ — involves

- (i) A set of players $N = \{1, \dots, n\}$;
- (ii) A time interval $[0, T]$ of specified duration (not necessarily finite);
- (iii) A strategy set \mathcal{A}_i for each player i , where a strategy $\alpha_i \in \mathcal{A}_i$ is understood as a contingent plan of actions where $\alpha_i(t)$ is the action taken at time $t \in [0, T]$;
- (iv) A trajectory space of the game \mathcal{Y} , with an element $\{Y_t\}_{t \geq 0} \in \mathcal{Y}$ being a permissible trajectory of the game. For each time $t \in [0, T]$, the state Y_t is located somewhere in the *state space* $\mathcal{Y} \subseteq \mathbb{R}^n$;
- (v) A differential equation, known as the *state equation*, describing the state trajectory of the game

$$(1.1) \quad \dot{y}(t) = \mu(y, t, \alpha(t)) = (\mu^1(y, t, \alpha(t)), \dots, \mu^n(y, t, \alpha(t)))^\top$$

where the action profile at time $t \in [0, T]$ is $\alpha(t) = (\alpha_1(t), \dots, \alpha_n(t))^\top$ and $Y_0 = y = (y^1, \dots, y^n)^\top \in \mathcal{Y}$;

- (vi) Functions $\pi^i : \mathcal{Y} \times A \rightarrow \mathbb{R}$ and $S^i : \mathcal{Y} \rightarrow \mathbb{R}$ for each player i such that the payoff function J^i

$$J^i(y, \alpha) \triangleq \int_0^T \pi^i(Y_t, t, \alpha(t)) dt + S^i(Y_T)$$

is well defined.

Başar and Olsder (1999) provide a wide overview over dynamic game theory with focus on differential games, while Dockner, Jorgensen, Van Long, and Sorger (2000) present a number of applications of differential games in economics and management science. When the state evolution is affected by a noise factor and the state equation is a stochastic differential equation of the form

$$(1.2) \quad dY_t = \mu(y, t, \alpha(t)) dt + \sigma(y, t, \alpha(t)) dB_t,$$

one deals with *stochastic differential games*. In a game with *simultaneous play*, all players choose their strategy at the outset so that no player enjoys an informational advantage over its player. If one player chooses her strategy before the other and commits to it, the information structure and strategy space are affected. Such games involving hierarchical play are called *Stackelberg games*. Differential games can also accommodate hierarchical play if the leader selects her strategy at the outset forming consistent beliefs about how the follower will react over time to its commitment.

2. Information structure and strategy space

As noted, in dynamic settings, one needs to specify the information structure and strategy space of the decision makers as part of the modeling stage. In a differential game, at each time $t \in [0, T]$, a tribe \mathcal{G}_t^i determines the information gathered by player i over the time interval $[0, t]$. Specification of the tribe \mathcal{G}_t^i for each $t \in [0, T]$ characterizes the *information pattern* of player i . The collection of the information patterns of all players determines the *information structure* of the game. In economic terms, the choice of the information structure determines which information is effectively collected by players over time and whether the players may react to new information over the game play.¹ In other words, one specifies

¹One often distinguishes various types of information structure, such as perfect, incomplete, or imperfect information. Here, we assume common knowledge about the strategy space and payoff functions but consider informational issues relating to the conditioning of the control function.

upon what information a player conditions her strategy. The largest reasonable choice of information pattern for each player i is the tribe $\mathcal{F}_t^{Y,\alpha} \triangleq \sigma(Y_s, a_s; 0 \leq s \leq t)$ generated by the whole history of state realizations and action choices. At places, the tribe $\mathcal{F}_t^{Y,\alpha}$ may be untractable to allow the modeler to characterize a well-behaved economic phenomena. In that case, we need to be more restrictive on describing the tribe \mathcal{G}_t^i for all $t \in [0, T]$ and player $i \in N$.

In dynamic games, the assumption concerning the information structure is intimately related to the hypothesized *strategy space*. This model assumption is the counterpart in case of dynamic games of the choice of control class in one-player control problems.

A strategy $\alpha^i : \Omega \times [0, T] \rightarrow A$ is *permissible* if it precludes foresight, i.e., if the function $\alpha^i(\cdot, t)$ is measurable with respect to $\mathcal{F}_t^{Y,\alpha}$ for all $t \in [0, T]$. In other words, if a player follows such a *closed-loop strategy*, she can condition, at each time $t \in [0, T]$, her actions on the whole past play of the game as summarized in the tribe $\mathcal{F}_t^{Y,\alpha}$. Employing closed-loop strategies is the less restrictive strategy space choice and is best in most dynamic games when information lags are negligible. Defining strategies in closed-loop form is often involved and possibly leading to an untractable model. This explains the success of alternative approaches.

We may assume, as the other extreme, that the players use a minimum of information and base their strategy only on the time that has elapsed so far, namely that players employ *open-loop strategies*. Ultimately, what particular action a player takes depends on calendar time only, the decision maker disregarding all information other than time. In other words, an open-loop strategy is a precommitment to a certain investment path, pursued regardless of the investment choices made by rivals. Open-loop strategies are static in a sense since they do not allow genuine strategic interactions over the game play. A natural question is then whether a device may endogenously justify the credibility of these commitments. In settings where a reduced number of players interact, modeling the strategic conflict via open-loop strategies is ill-advised and may even lead to wrong predictions. There are, however, some arguments in favor of open-loop strategies. An obvious reason is that players may lack any information other than time or face infinite information lags. If the planning horizon is short, an open-loop strategy could represent a rigid strategy for short-term operational and tactical planning. It could also be used as an approximation in a game with a large number of small players where strategic interactions have no material influence on each other's welfare. Following Fudenberg and Levine (1988), as the number of players grows, the outcome of the closed-loop equilibria — which is a priori derived based on less restrictive assumptions — converges to open-loop equilibria. In industrial organization, the open-loop equilibrium thus provides a good prediction for oligopolies with a large number of firms; This latter situation resembles perfect competition in microeconomics.

Inbetween the two polar cases of closed-loop and open-loop strategies, a player may condition her action choice on the current value of the state $Y_t \in \mathcal{Y}$ and the current time $t \in [0, T]$. Such controls are called *Markov strategies*. If they depend on the state value but not on time, they are called *stationary Markov strategies*. Markov strategies help strike the balance between realistic assumptions and mathematical tractability. In state-space games — of which differential games are a subset — the previous action history is adequately summarized in the current state vector. In other words, the game history is not relevant for the choice of action at time t but only the consequences of past actions as reflected in the current state value. In such state-space games, the Markov assumption — which is not stronger than the closed-loop assumption — is a natural choice of strategy class.

3. Solution concepts

In problems involving a single decision maker, *optimality* has no ambiguous meaning. In contrast, the notion is not uniquely defined in multi-player decision settings. In games, a *solution concept* defines a certain notion of optimality according to which players behave and prescribes criteria an equilibrium strategy profile α^* should fulfill. If such a strategy profile exists, player i receives *value* $V^i = J(\alpha^*)$.

A solution concept is thus a methodology for predicting players' behavior intended to determine the strategies that maximize each player's payoff. In game theory, players' rationality is typically considered common knowledge. *Rationalizability*, introduced by Bernheim (1984) and Pearce (1984), is considered a fundamental solution concept. Based on the weak assumption that the game structure is common knowledge and that players are rational, it predicts that a player will not play a strategy that is not a best response to some beliefs about her opponents' strategies. By iteration, we can narrow down the set of strategies that could be reasonably played (so-called rationalizable strategies). By design, rationalizability makes weak predictions owing to a generally large set of rationalizable strategies. To obtain stronger predictions, one may impose assumptions beyond common knowledge of rationality. Nash's (1951) equilibrium concept is usually preferred because the latter equilibrium solution has the advantage of existing for a larger class of games. A strategy profile $\alpha = (\alpha_i^*, \alpha_{-i}^*)$ is *Nash equilibrium* if, for any player $i \in N$,

$$J^i(\alpha_i^*, \alpha_{-i}^*) \geq J^i(\alpha_i, \alpha_{-i}^*), \quad \forall \alpha_i \in \mathcal{A}_i.$$

We can interpret the Nash equilibrium in terms of best-reply correspondences or reaction functions. Let $R(\alpha_{-i})$ denote firm i 's *best responses* to her rivals' strategies α_{-i} :

$$R(\alpha_{-i}) = \{\alpha_i \in \mathcal{A}_i \mid J^i(\alpha_i, \alpha_{-i}) \geq J^i(\gamma_i, \alpha_{-i}), \forall \gamma_i \in \mathcal{A}_i\}$$

If $R(\alpha_{-i})$ is a singleton, $R(\alpha_{-i})$ is a *reaction function*. We can now re-interpret the Nash equilibrium notion. In Nash equilibrium, each player formulates her best response to the other players' optimal decisions; that is, $\alpha_i^* \in R(\alpha_{-i}^*)$. Nash (1951) proved that every finite normal-form game has (at least) one Nash equilibrium in mixed strategies. This result is proved by applying Kakutani's fixed-point theorem to the players' best response sets. When multiple Nash equilibria exist, one may use *refining schemes* to narrow down the set of equilibria to those equilibria that fulfill additional requirements, such as perfection. Subgame perfection was introduced by Selten (1975) in the framework of an extensive form (with subgames). A Nash equilibrium α^* is *subgame perfect* if, for *every* history, the restriction of the strategy profile to the continuation game constitutes a Nash equilibrium of the continuation game. Subgame perfection generalizes our notion of strongly time consistency to multiplayer settings. The notion of Nash equilibrium is widely used in dynamics games whereby one may distinguish different strategy space.

When we restrict the strategy spaces to open-loop strategies, we have an obvious application of this equilibrium concept. The profile $\psi = (\psi^1, \dots, \psi^n)$ of open-loop strategies $\psi^i : [0, T] \rightarrow A$, $i \in N$, is an *open-loop equilibrium* if a Nash equilibrium profile $\alpha^* = (\alpha^{i*}, \alpha^{-i*})$ exists and is given by the open-loop strategy $\alpha_i^*(t) = \psi^i(t)$. The Nash equilibrium in a game modeled with Markov strategies is called Markov Nash Equilibrium.² The profile $\phi = (\phi^1, \dots, \phi^n)$ of Markov strategies $\phi^i : \mathcal{Y} \times [0, T] \rightarrow A$, $i \in N$, is called a *Markov Nash equilibrium* if a Nash equilibrium $\alpha = (\alpha^{i*}, \alpha^{-i*})$ exists and is given by the Markov strategy $\alpha^{i*}(t) = \phi^i(Y_t, t)$. By determining a nondegenerate Markov Nash equilibrium one makes the assumption that the state variable can be observed and that the players can condition their actions on these observations.

Dynamic games of complete information are often solved by means of the subgame-perfect Nash Equilibrium solution concept that is strongly time consistent and precludes empty threats. The Markov Perfect Equilibrium concept captures the same idea of subgame perfection but does not explicitly refer to the extensive form of the state-space game. Denote by $\Gamma(y, t)$ the normal-form game $\Gamma = \{N, (\mathcal{A}_i)_{i \in N}, (J_i)_{i \in N}\}$ in state $y \in \mathcal{Y}$ at time t . A Markov Nash Equilibrium $\phi = (\phi^1, \dots, \phi^n)$ is *Markov perfect* if the strategy profile ϕ is a Markov Nash equilibrium of $\Gamma(y, t)$ for all $(y, t) \in \mathcal{Y} \times [0, T]$. In other words, for every $(y, t) \in \mathcal{Y} \times [0, T]$, the game $\Gamma(y, t)$ admits a Markov Nash equilibrium $\alpha_{[t, T]} = (\alpha_{[t, T]}^1, \dots, \alpha_{[t, T]}^n)$ such that $\alpha^i(\omega, s) = \phi^i(Y_s, s)$, for all $i \in N$ and $(Y_s, s) \in \mathcal{Y} \times [0, T]$. The

²A Markov Nash Equilibrium is often referred to as a feedback or closed-loop equilibrium (see Fudenberg and Tirole, 2002).

restriction of Markov perfection prescribes a credible rule to determine the actions of players i not only along the equilibrium path $Y^* = \{Y_t^*\}_{t \geq 0}$ but also for all other trajectories in \mathcal{Y} .

Again, the choice whether to look for an open-loop Nash or a Markov perfect equilibrium in a differential game is part of the modeling stage. Ultimately, one should consider the strategy space and solution concept that best describe the situation at hand.

4. Nonzero-sum differential games

Isaacs (1965) discusses differential games (i.e., with deterministic state equation) in the context of two-player, zero-sum games. In zero-sum games, the fixed-point strategies for which the game has a value are obtained by the minmax theorem. If zero-sum games are good to describe situations arising in military conflicts, they have limited applicability in economics. Starr and Ho (1969) extend the theory of differential games to nonzero-sum games with n players, discussing the difficult question of determining the right solution concept that emerges in such a situation (e.g., minmax, Nash equilibrium, or Pareto optimum). In the general nonzero-sum, n -player differential game, player i 's payoff for strategy profile α at time t and in state $y \in \mathcal{Y}$ is:

$$J^i(y, t, \alpha) = \int_t^T \pi^i(Y_s^\alpha, s, \alpha(s)) ds + S^i(Y_T^\alpha, T), \quad t \in [0, T]$$

subject to the deterministic state equation (1.1) on page 40. In Nash equilibrium, firm i 's value in state $(y, t) \in \mathcal{Y} \times [0, T]$ is

$$V^i(y, t) \triangleq \max_{\alpha_{[t, T]}^i \in \mathcal{A}_{[t, T]}^i} J^i(y, t, \alpha_{[t, T]}^i, \alpha_{[t, T]}^{-i})$$

As usual, we denote the first-order derivatives of the value function by $V_t^i = \partial V^i / \partial t$ and $V_y^i = \partial V^i / \partial y$.

Consider first that the profile of Markov strategies $\phi = (\phi^i, \phi^{-i})$ form a *Markov Nash equilibrium*. In Nash equilibrium, a player assumes that rivals will follow their Nash equilibrium strategies. For notational simplicity, we denote

$$\bar{\mu}^i(y, t, a^i) = \mu^i(y, t, (a^i, \phi^{-i}(y, t))),$$

so that the state equation (1.1) on page 40 now reads $\dot{y}(t) = \bar{\mu}^i(y, t, a^i)$. Likewise, we rewrite the instantaneous utility flow:

$$\bar{\pi}^i(y, t, a^i) = \pi^i(y, t, (a^i, \phi^{-i}(y, t))).$$

Firm i 's payoff function when rivals follow their MNE strategies ϕ^{-i} is then

$$\bar{J}^i(y, t, \alpha^i) = \int_t^T \bar{\pi}^i(Y_s, s, \alpha^i(s)) ds + S^i(Y_T, T), \quad t \in [0, T].$$

By convention, $\bar{J}^i(y, \alpha^i) = \bar{J}^i(y, 0, \alpha^i)$. By definition, the profile $\phi = (\phi^i, \phi^{-i})$ of strategies $\phi^i : \mathcal{Y} \times [0, T] \rightarrow A^i$ is a Markov Nash equilibrium if, for each player $i \in N$, there exists a strategy α^i that maximizes $\bar{J}^i(y, \cdot)$ provided that (i) the state equation (1.1) is satisfied with initial condition $Y_0 = y \in \mathcal{Y}$ and (ii) the optimal action at time $t \in [0, T]$ and state $Y_t(\omega) \in \mathcal{Y}$ is given by $\alpha^i(t) = \phi^i(Y_t(\omega), t)$. The above optimization program appears to be similar to a one-player problem under the assumption that firm i 's rivals follow strategies ϕ^{-i} . In the deterministic case, the HJB equation reads

$$V_t(y, t) = \max_{a^i \in A^i} \left\{ \bar{\pi}^i(y, t, a^i) + V_y^i(y, t)^\top \bar{\mu}^i(y, t, a^i) \right\}.$$

We introduce a set of co-state variables $\lambda_1^i = (\lambda_1^{i1}, \dots, \lambda_1^{in})^\top$ and define player i 's Hamiltonian function H^i :

$$H^i(y, t, a, \lambda_1^i) = \bar{\pi}^i(y, t, a) + \lambda_1^{i\top} \bar{\mu}(y, t, a) = \bar{\pi}^i(y, t, a) + \sum_{j=1}^n \lambda_1^{ij} \mu(y, t, a).$$

As before, assume rivals follow their Nash equilibrium strategies. We then have the following variant of the Hamiltonian function

$$\bar{H}^i(y, t, a^i, \lambda_1^i) = \bar{\pi}^i(y, t, a^i) + \lambda_1^{i\top} \bar{\mu}(y, t, a^i).$$

The value functions $V^i \in C^{2,1}(\mathcal{Y}, [0, T])$, $i \in N$ are then solutions of the *generalized HJB equation*

$$(4.1a) \quad V_t^i(y, t) = \max_{a^i \in A^i} \bar{H}^i(y, t, a^i, V_y^i(y, t)), \quad i \in N,$$

subject to the terminal boundary condition

$$(4.1b) \quad V^i(y, T) = S^i(y, T).$$

To characterize the Markov perfect equilibrium, one must be able at each $(y, t) \in \mathcal{Y} \times [0, T]$ to find a Markov Nash equilibrium $\phi(y, t) = (\phi^1(y, t), \dots, \phi^n(y, t))$ that solves a static, continuous, nonzero-sum, n -player game with payoff functions $H = (H^1, \dots, H^n)^\top$. Finding such a solution is not always feasible but is possible for an important class of games. A differential game is said to be *normal* if (i) it is possible to find such a Markov Nash equilibrium ϕ for $H = (H^i, H^{-i})$ for all (y, t) and (ii) when equations (1.1) and (4.1a), whereby the controls are given by the Markov rules $\alpha^i(t) = \phi^i(Y_t^*, t)$ for all $(y, t) \in \mathcal{Y} \times [0, T]$, are satisfied. We can also adopt the maximum principle approach for situations involving n -players. Following this approach, the maximum condition,

$$\phi^i(y, t) \in \arg \max_{a^i \in A^i} \bar{H}^i(y, t, a^i, \lambda_1^i),$$

is satisfied for all point $(y, t) \in \mathcal{Y} \times [0, T]$. A Markov Nash equilibrium must also satisfy, for $i \in N$, the adjoint equation

$$(4.2a) \quad \dot{\lambda}_1^i(t) = - \left\{ \frac{\partial H^i}{\partial y}(y, t, (a, \phi^{-i}(y, t)), \lambda_1^i) + \sum_{j=1, j \neq i}^n \frac{\partial H^i}{\partial a^j} \frac{\partial \phi^j}{\partial y}(y, t) \right\}, \quad t \in [0, T]$$

$$(4.2b) \quad \lambda_1^i(T) = S_y^i(Y_T, T),$$

where $\frac{\partial H^i}{\partial y}$ and $\frac{\partial \phi^j}{\partial y}$ are, respectively, the gradients of the Hamiltonian and player j 's policy with respect to the (n -dimensional) state. In the optimal control setting ($n = 1$), the second term for the adjoint equation (4.2a) is absent. The adjoint equation is then an ordinary differential equation which is less computationally involved. In games ($n \geq 2$), the adjoint equation admits an additional strategic effect term that captures the dependence of the remaining $(n - 1)$ players' strategies on the current value of the state y . In other words, players care about how their opponents will react to changes in the state variables. This gives rise to a partial differential equation for the adjoint equation, in contrast to an ordinary differential equation in case $n = 1$. The presence of the cross-influence term in the adjoint equation transforms the adjoint equation from an ordinary differential equation (ODE) into a system of partial differential equations (PDEs), which are more involved. As evidenced in equation (4.2a), each player takes account in her objective function H^i of the influence of her actions on the state, both directly (via $\partial H^i / \partial y$) and indirectly through the influence of the state on the strategies of the player's opponents (via $\partial H^i / \partial a^j \times \partial \phi^j / \partial y$). In a deterministic one-player setting, the restriction to Markov controls is not needed since the optimal open-loop and Markov controls give rise to different representations corresponding to the same control path.

Among "normal games," *linear-quadratic games* are often used in differential games. Such games are characterized by (i) a state equation that is linear in the state and the control and (ii) an objective function that is quadratic in the state and the control. For the special case of linear-quadratic game, Nash equilibrium, minimax, and the Pareto-optimal strategies are obtained by solving sets of ordinary matrix differential equations, known as Riccati equations.

5. Stochastic differential games

We presented above how to solve nonzero-sum, n -person differential games. Following the discussion in Dockner et al. (2000), Chapter 8, we present techniques to solve stochastic differential games, whereby the state evolution is affected by noisy disturbances. The state equation is now a SDE of the Itô diffusion type:

$$(5.1) \quad dY_t = \mu(Y_t, t, a_t) dt + \sigma(Y_t, t, a_t) dB_t \quad \text{and } Y_0 = y \in \mathbb{R}^n,$$

where μ and σ satisfy the usual conditions, and B is a standard Brownian motion. If players follow the strategy profile (α^i, α^{-i}) , player i 's payoff (at the outset) yields

$$J^i(y, \alpha) = \mathbb{E} \left[\int_0^T \pi^i(Y_t^\alpha, t, \alpha) + S(Y_T^\alpha, T) \mid y \right],$$

subject to state equation (5.1). We first intend to give sufficient conditions for a Markov Nash equilibrium to exist. For notational simplicity, let $\bar{\sigma}(y, t, a^i) = \sigma^i(y, t, (a^i, \phi^{-i}(y, t)))$. Given her belief about her rivals' equilibrium play, player i faces a one-person dynamic optimization problem, $\bar{\Gamma}^i(y, 0)$, which consists in maximizing the payoff

$$\bar{J}^i(y, \alpha^i) = \max_{\alpha^i \in \mathcal{A}^i} \mathbb{E} \left[\int_0^T \bar{\pi}^i(Y_t, t, \alpha^i(t)) dt + S^i(Y_T, T) \mid y \right],$$

subject to the state equation (5.1). The profile $\phi = (\phi^i, \phi^{-i})$ of Markov strategies is a *Markov Nash equilibrium* of $\Gamma(y, 0)$ if, for each player i , the strategy α^i given by $\alpha^i(\omega, t) = \phi^i(Y_t(\omega), t)$ solves the dynamic problem $\bar{\Gamma}^i(y, 0)$. Consider a vector of Lagrangian multipliers: $\lambda^i \triangleq (\lambda_1^i, \lambda_2^i)$. Player i 's current-value stochastic Hamiltonian function is

$$\begin{aligned} \mathcal{H}^i(y, t, a, \lambda^i) &\triangleq \pi^i(y, t, a) + \lambda_1^{i\top} \mu(y, t, a) + \frac{1}{2} \text{tr} \langle \sigma(y, t, a) \lambda_2^{i\top} \sigma(y, t, a) \rangle \\ \bar{\mathcal{H}}^i(y, t, a^i, \lambda^i) &\triangleq \bar{\pi}^i(y, t, a^i) + \lambda_1^{i\top} \bar{\mu}(y, t, a^i) + \frac{1}{2} \text{tr} \langle \bar{\sigma}(y, t, a^i) \lambda_2^{i\top} \bar{\sigma}(y, t, a^i) \rangle. \end{aligned}$$

Following the dynamic programming approach, player i 's value function $V^i \in C^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$ solves for each $(y, t) \in \mathcal{Y} \times [0, T]$ the HJB equation

$$(5.2a) \quad V_t^i(y, t) = \max_{a^i \in A^i} \bar{\mathcal{H}}^i(y, t, a^i, (V_y^i(y, t), \bar{\sigma}(y, t, a^i)^\top V_{yy}^i)),$$

with boundary condition

$$(5.2b) \quad V^i(Y_T, T) = S^i(Y_T, T).$$

In the present stochastic case, we obtain a system of second-order partial differential equations. To obtain a Markov Nash equilibrium, one must be able at each (y, t) to find a Nash equilibrium for the static optimization problem $\mathcal{H} = (\mathcal{H}^i, \mathcal{H}^{-i})$. Owing to the second-order term in the HJB equation, it is generally difficult to find closed-form solutions for the value functions V^i , unless certain ‘‘smoothness’’ restrictions are imposed.

We now characterize the Markov perfect equilibrium further. The first-order adjoint equation reads

$$(5.3a) \quad \dot{\lambda}_1^i(t) = - \left\{ \frac{\partial \bar{\mathcal{H}}^i}{\partial y}(y, t, a^i, \lambda^i) - \sum_{j=1, j \neq i}^n \frac{\partial \mathcal{H}^i}{\partial a^j} \frac{\partial \phi^j}{\partial y}(y, t) \right\}$$

with terminal condition

$$(5.3b) \quad \lambda_1^i(T) = S_y^i(Y_T, T).$$

Along the equilibrium path, the maximum condition

$$\phi^i(y, t) \in \arg \max_{a^i \in \mathcal{A}^i} \bar{\mathcal{H}}^i(y, t, a^i, \lambda^i)$$

is satisfied for all $(y, t) \in \mathcal{Y} \times [0, T]$ and every player i . The terms $\partial \mathcal{H}^i / \partial a^j \times \partial \phi^j / \partial y$ capture as before the strategic effect. The difference with the previous differential game lies in the higher order term in HJB equation (5.2a).

6. References

Isaacs (1965) discusses differential games (i.e., with deterministic state equation) in the context of two-player, zero-sum games. Starr and Ho (1969) extend the theory of differential games to nonzero-sum games with n players. Başar and Olsder (1999) and Dockner et al. (2000) are standard references on dynamic game theory and differential games.

Part 2

Benchmark References

Monopoly

Investment is a major research topic in economics and finance. To simplify the analysis, investment has been often considered either costlessly reversible or fully sunk (irreversible). In between these polar cases lies costly reversibility, a subject which has been thoroughly investigated in the literature on real options analysis. Herein, we briefly present two themes: a monopolist firm may have to decide *when* to make a one-time lump-sum investment or *how much* to invest incrementally at any instant. Section 1 discusses some simple cases of lump-sum investments for a monopolist firm, while Section 2 focuses on the process of capital adjustments over time. Section 3 gives references to dedicated academic works.

1. Investment timing and hysteresis

McDonald and Siegel (1986) sets the benchmark case for the analysis of the timing option. Consider that a firm may enter a market whose underlying value $X = \{X_t\}_{t \geq 0}$ follows the GBM of equation (3.4). If it invests at, say, time $\tau(\omega)$ it incurs an entry cost I . The payoff to the firm if it invests at time τ is

$$J(x, \tau) = \mathbb{E} \left[e^{-r\tau} (X_\tau - I)^+ \right],$$

where r is a constant discount rate such that $r > \mu + \sigma^2/2$. The decision maker will maximize $J(x, \cdot)$ over the set of admissible stopping times \mathcal{T} . Here, it is reasonable to assume that the firm will invest if the underlying project value is sufficiently large to justify the expense of market entry, i.e., higher than an optimally selected threshold $\bar{X} \in \mathcal{X}$. In Chapter 3, we discussed optimal stopping problems and derived a relevant HJB equation (see equation 4.5). According to this equation, the value function solves a partial differential equation,

$$(1.1) \quad rV(x) = \mu xV_x(x) + \frac{1}{2}\sigma^2 V_{xx}(x),$$

for $x < \bar{X}$, together with a terminal boundary

$$(1.2) \quad V(x) = (x - I)^+$$

if $x \geq \bar{X}$. Besides, V should be at least continuously differentiable, obtaining the smooth-pasting condition

$$(1.3) \quad V_x(x) = 1.$$

In the continuation region (∞, \bar{X}) , the second-order ordinary differential equation (1.1) has a general solution of the form

$$V(x) = A x^{\beta_1} + B x^{\beta_2}.$$

β_1 and β_2 are, respectively, the positive and the negative root of

$$(1.4) \quad \mathcal{Q}(\beta) = r - \left(\mu - \frac{1}{2}\sigma^2 \right) \beta - \frac{1}{2}\sigma^2 \beta^2.$$

$\mathcal{Q}(\cdot)$ is known in the real options literature as the “fundamental quadratic” (see Dixit and Pindyck, 1994, Chapter 4). It is reasonable to assume that the value function V has a finite right limit at the origin: $V(x) \rightarrow 0$ as $x \rightarrow 0$. Since x^{β_2} blows up as $x \downarrow 0$, this suggests that $B = 0$ and $V(x) = A x^{\beta_1}$ if $x < \bar{X}$,

the latter expression capturing the option to wait. Hence, it obtains from the value-matching and the smooth-pasting conditions in equations (1.2) and (1.3), respectively that

$$\bar{X} = \frac{\beta_1}{\beta_1 - 1} I \quad \text{and} \quad A = (\bar{X} - I) \bar{X}^{-\beta_1}.$$

Consequently,

$$V(x) = (\bar{X} - I) \left(\frac{x \wedge \bar{X}}{\bar{X}} \right)^{\beta_1}, \quad \forall x \in \mathbb{R}_+.$$

Dixit (1989) analyzes both entry and exit decisions under uncertainty. Suppose a firm has to decide whether to enter a market or stay out having leeway in selecting the optimal entry and exit times. If the firm decides to enter at time t , it incurs an investment cost $\bar{\kappa}$ but can produce at a unit cost c a product sold at market price X_t . The price process $X = \{X_t\}_{t \geq 0}$ follows a GBM. Once it has entered, the firm is allowed to leave the market at a cost $\underline{\kappa}$, with $|\underline{\kappa}| < \bar{\kappa}$. We follow the risk-neutral approach and use a discount rate $r (> \mu)$. We have two state variables: the price $x \in \mathcal{X}$ and a variable $l \in \{0, 1\}$, which indicates whether the firm is active ($l = 1$) or not ($l = 0$). $V(x, l)$ denotes the value associated with states $x \in \mathcal{X}$ and l . Here, it seems natural to conjecture that, in state $l = 0$, the firm will enter the market if the price is sufficiently large, i.e., if it goes above a certain threshold \bar{X} . Similarly, if the firm is active ($l = 1$), it is sensible to surmise that exit takes place when the price falls below a particular boundary \underline{X} .

Consider first the case where $l = 0$. Then, the HJB equation is

$$(1.5) \quad 0 = \max \{-r V(x, 0) + \mathcal{D}V(x, 0); V(x, 1) - V(x, 0) - \bar{\kappa}\},$$

where

$$\mathcal{D}V(x, l) = \mu x V_x(x, l) + \frac{1}{2} \sigma^2 x^2 V_{xx}(x, l), \quad l \in \{0, 1\}.$$

Here, according to the HJB equation (1.5), the value function must satisfy a second-order ordinary differential equation in the continuation region. The second term indicates the value change when the firm actually enters. In the region $(-\infty, \bar{X})$, the ODE has a solution of the form

$$V(x, 0) = A_0 x^{\beta_1} + B_0 x^{\beta_2}.$$

where β_1 and β_2 are the roots of the fundamental quadratic. To ensure that the value function is finite at $x \searrow 0$, we need $B_0 = 0$.

Consider now the case where $l = 1$. The HJB equation is

$$(1.6) \quad 0 = \max \{(x - c) - r V(x, 1) + \mathcal{D}V(x, 1); V(x, 0) - V(x, 1) - \underline{\kappa}\}.$$

In the continuation region, the homogenous part is the same as before. We now look for a particular solution, say of the form $V(x, 1) = ax + b$. By substitution in the PDE, we have

$$0 = (x - c) - r(ax + b) + \mu x a,$$

whence $a = 1/(r - \mu)$ and $b = -c/r$. Our candidate solution is thus

$$V(x, 1) = A_1 x^{\beta_1} + B_1 x^{\beta_2} + \frac{x}{r - \mu} - \frac{c}{r}.$$

We want to ensure that as the price x grows arbitrarily large, the value function does not go to infinity as well. To achieve that, we must impose $A_1 = 0$.

We are left with

$$V(x, 0) = A_0 x^{\beta_1}, \quad \text{and} \quad V(x, 1) = \frac{x}{r - \mu} - \frac{c}{r} + B_1 x^{\beta_2}.$$

$A_0 x^{\beta_1}$ is the value of the entry option and $B_1 x^{\beta_2}$ that of the exit option. From equations (1.5) and (1.6) as well as smooth-pasting arguments, we can further determine A_0 , B_1 , and the cutoff values:

$$\begin{aligned} V(\bar{X}, 0) &= V(\bar{X}, 1) - \bar{\kappa} & \text{and} & \quad V_x(\bar{X}, 0) = V_x(\bar{X}, 1), \\ V(\underline{X}, 1) &= V(\underline{X}, 0) - \underline{\kappa} & \text{and} & \quad V_x(\underline{X}, 0) = V_x(\underline{X}, 1). \end{aligned}$$

Unfortunately, this system cannot be solved in closed form. A solution nevertheless exists. Note that the Marshallian cutoffs for entry and exit, respectively, $\underline{M} \triangleq c - r\underline{\kappa}$ and $\overline{M} \triangleq c + r\overline{\kappa}$. Consider the function $G(\cdot) \triangleq V(\cdot, 1) - V(\cdot, 0)$; $G(x)$ can be written as

$$G(x) = -A_0 x^{\beta_1} + B_1 x^{\beta_2} + \frac{x}{r - \mu} - \frac{c}{r}.$$

From our previous analysis, we obtain the second-order ordinary differential equation

$$(1.7) \quad rG(x) = (x - c) + \mu x G_x(x) + \frac{1}{2} \sigma^2 G_{xx}(x),$$

which has boundary conditions, $G(\underline{X}) = -\underline{\kappa}$, $G(\overline{X}) = \overline{\kappa}$, $G_x(\underline{X}) = G_x(\overline{X}) = 0$, with $G_{xx}(\underline{X}), G_{xx}(\overline{X}) > 0$. By substitution in equation (1.7), we obtain $\overline{X} > \overline{M}$ and $\underline{X} < \underline{M}$ or $(\underline{M}, \overline{M}) \subset (\underline{X}, \overline{X})$ if $\sigma > 0$. The introduction of uncertainty has the effect of widening the price range of inaction of the firm. When uncertainty fades away ($\sigma \rightarrow 0$), the triggers under uncertainty converge to the Marshallian cutoffs: $\underline{X} \rightarrow c - r\underline{X}$ and $\overline{X} \rightarrow c + r\overline{X}$.

2. Capital adjustment and marginal Tobin's q theory

Abel and Eberly (1994) tackle the problem of determining the optimal investment behavior for a monopolist facing uncertain market developments and costly reversibility. We want to give an account of their model. In the following, we proceed as follows. Section 2.1 sets the stochastic control problem faced by the monopolist firm. Section 2.2 provides necessary conditions for the optimal investment policy and describes the dynamics of the value function. Finally, we discuss several examples from the literature on investment under uncertainty.

In microeconomic models dealing with investment in capital stocks, the adjustment cost function is often assumed to be strictly convex. Consider the ‘‘augmented version’’ of the capital-adjustment cost function in equation (2.1); it is given by

$$(2.1) \quad k(I, K) = \kappa(I) + \kappa_1(I) + \kappa_2(I, K),$$

whereby ‘‘ $k(0, K) > 0$ ’’ captures fixed adjustment costs. $k(\cdot, K)$ is everywhere differentiable with respect to investment magnitude I , except possibly at $I = 0$, with the first-order derivative on $(-\infty, 0)$ and $(0, \infty)$ denoted by $k_I(\cdot, K)$. The left-limit of $k_I(\cdot, K)$ at $I = 0$ is denoted by $\underline{k}_I(K) \triangleq \lim_{I \uparrow 0} k_I(I, K)$, while the right-limit is $\overline{k}_I(K) = \lim_{I \downarrow 0} k_I(I, K)$. If $\underline{k}_I(K) = \overline{k}_I(K)$, $k(\cdot, K)$ is differentiable at $I = 0$. If the adjustment costs are partly sunk, the capital stock $K_t(\omega)$ installed at time t may depend on the whole past history, $\{X_s(\omega); 0 \leq s \leq t\}$ of the exogenous shock process X .

2.1. Value dynamics under optimal policy. Suppose the industry aggregate shock $\{X_t\}_{t \leq 0}$ follows a time-homogenous Itô process as per equation (3.1). An investment strategy is a capital-accumulation process ξ . Note the negative (resp., positive) part of ξ by ξ^- (resp., ξ^+). The processes ξ^- and ξ^+ are right-continuous with left limits, nondecreasing, and adapted to the filtration \mathbb{F} . Besides, $\xi^-(0-) = 0$ and $\xi^+(0-) = 0$ almost surely. The sample paths of the process $\xi = \xi^+ - \xi^-$ have bounded variations as ξ is the difference between two nondecreasing processes. If the firm pursues the investment strategy ξ , its capital stock at time t and on path $\omega \in \Omega$ is

$$(2.2) \quad K_t(\omega) = k + \xi_t^+(\omega) - \xi_t^-(\omega).$$

An instantaneous change in capital stock is noted $d\xi_t(\omega) = \xi_t(\omega) - \xi_{t-}(\omega)$. $d\xi_t(\omega)$ is a Lebesgue-Stieltjes differential and cannot properly be interpreted in terms of investment rate since it is not absolutely continuous with respect to the Lebesgue measure. If it were possible to determine an ‘‘investment rate,’’ we could define a control policy prescribing to the firm, at each instant $\in \mathbb{R}_+$, whether to adjust its capital stock or not (action $a = 1$ or 0 respectively). and if the firm adjusts its capital stock ($a = 1$) and incurs fixed adjustment costs how much to invest or divest (investment rate I). Here, a and I would be the control variables to be decided upon at each instant. Abel and Eberly (1994) follow this heuristic

approach. The capital stock of the monopolist, $\{K_t\}_{t \leq 0}$, then evolves pathwise according to (differential) equation, namely

$$\dot{K}_t = I_t^i(\omega) - \lambda K_t^i(\omega),$$

with λ being the depreciation rate. A control is an \mathbb{F} -adapted process $\alpha = \{a_t; I_t\}_{t \geq 0}$.

The firm's instantaneous profit at time t on path $\omega \in \Omega$ is driven by its capital stock level $K_t(\omega)$ and the industry shock $X_t(\omega)$. At each instant t , the firm receives an unknown revenue flow $\pi(X_t(\omega), K_t(\omega))$ depending on the shock realization $X_t(\omega)$ and the current capital stock, $K_t(\omega)$. The profit function π is strictly increasing and twice continuously differentiable in x ; it is continuously differentiable, strictly decreasing, and concave in K . The marginal profit is increasing in the shock $\pi_{xx} > 0$; this ensures that the optimal strategy can be expressed in terms of divestment/investment triggers. We follow the risk-neutral valuation approach — under the assumption of a complete market that precludes arbitrage opportunities. The firm faces a common, constant, risk-free rate r . The *payoff* received at the outset with initial state (x, K) and when the firm pursues strategy $\alpha \in \mathcal{A}$ is

$$(2.3) \quad J(x, K, \alpha) \triangleq \mathbb{E} \left[\int_0^\infty e^{-rt} \{ \pi(X_t, K_t) - a_t k(I_t, K_t) \} dt \mid x, K \right].$$

The optimal control policy yields, in state (x, K) , a *value*

$$V(x, K) \triangleq \max_{\alpha} J(x, K, \alpha).$$

Given the model assumptions, V is strictly increasing, twice continuously differentiable in the shock. Besides, $V_{xk} > 0$ and $V_{kk} < 0$.

We shall focus on Markov controls. We will see that when investment is costly reversible, the optimal Markov control ϕ involves lower and upper barriers: in the low region the firm divests; in the intermediary region, it stays put; in the high region, it invests. Since $\pi(\cdot, K)$ and $V(\cdot, K)$ are continuous and strictly increasing in the shock x , we could alternatively consider boundaries in terms of shock, profit, value, or marginal value. We shall adopt the original approach of Abel and Eberly (1994) involving the marginal value of capital.

Consider the payoff function (2.3). The resulting HJB equation is a second-order partial differential equation:

$$(2.4) \quad (r + \lambda) V(x, K) = \max_{a, I} \{ \pi(x, K) - a k(I, K) + \mathcal{D}V(x, K) \}.$$

The infinitesimal generator \mathcal{D} specializes here to

$$(2.5) \quad \mathcal{D}V(x, K) = V_K(x, K) (I - \lambda K) + V_x(x, K) \mu(x) + \frac{1}{2} V_{xx}(x, K) \sigma(x)^2.$$

Denote the marginal value or *shadow price of capital* by $v(x, K) \triangleq V_K(x, K)$. It is the marginal profitability of an additional unit of capital. $\{v_t\}_{t \geq 0}$ with $v_t = v(X_t, K_t)$ can be viewed as a controlled stochastic process. Let's now differentiate the HJB equation (2.4) by K . The optimal control trajectory $\{a_t, I_t^*\}_{t \geq 0}$ obtains to satisfy

$$(2.6) \quad (r + \lambda) v(x, K) = \max_{a, I} \{ \pi_K(x, K) - a k_K(I, K) + \mathcal{D}v(x, K) \}$$

In equation (2.6), the left-hand side term, $(r + \lambda) v(x, K)$, represents the required (gross) return on the value of the marginal unit of capital. The right-hand side comprises (i) the marginal operating profit, $\pi_K(x, K)$, (ii) the cost incurred by the firm if it adjusts its capital stock marginally, namely $a k_K(I, K)$, and (iii) the expected marginal instantaneous capital gain, $\mathcal{D}v(x, K)$. Applying the Feynman-Kac Theorem to equation (2.6) yields

$$(2.7) \quad v(x, K) = \max_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^\infty e^{-(r+\lambda)t} \{ \pi_K(X_t, K_t) - a_t k_K(I_t, K_t) \} dt \mid x, K \right] \right\}.$$

The expected capital gain term $\mathcal{D}v_s$ is of second order to reflect the unbounded first but bounded second variation of the stochastic process $\{v(X_t, K_t)\}_{t \geq 0}$. In the deterministic literature on the q theory of

investment, a first-order term suffices; it is given by the rate of change in value over time, namely $\dot{v} = dv/dt$. Alternatively, we could look at the dynamics of the shadow price directly in equation (2.7) instead of the original problem in (2.6).

A case of interest involves a simplified problem where the decision maker can acquire each unit of capital at a price $\bar{\kappa}$ (with no additional adjustment costs) but cannot reduce its capital stock. This stochastic control problem involves the value function

$$V(x, K) \triangleq \max_{\xi} \mathbb{E} \left[\int_0^{\infty} e^{-rt} \pi(X_t, K_t) dt - \bar{\kappa} d\xi_t \mid x, K \right].$$

Karatzas and Shreve (1984) show that the above singular control problem is related to an optimal stopping problem, called *Snell-envelope problem*, which consists in selecting, among the set of admissible stopping times \mathcal{T} , the stopping time that solves the optimization problem

$$\max_{\tau \in \mathcal{T}} \mathbb{E} \left[\int_{\tau}^{\infty} e^{-rt} \pi_K(X_t, K) dt - \bar{\kappa} e^{-r\tau} \mid x, K \right].$$

The dynamic optimization problem $V(x, k)$ can be decomposed in a sequence of optimal stopping problems: rather than deciding *how much* to invest at any given time, the firm can decide *when* infinitesimal particles of capital should be installed (see Bertola, 1998). This mathematical approach is useful and has been applied to provide further economic intuition to the underlying capital expansion problem. When, as in the present case, a firm may expand *and* reduce its capital stock, the “duality” between the two problems is not readily applicable.

2.2. Characterization of the optimal investment policy. By isolating terms depending on a and I in the HJB equation (2.4), the maximization program simplifies to the concave problem

$$(2.8) \quad f(v, K) = \max_{a, I} \{vI - ak(I, K)\},$$

where we drop the dependence of the marginal value v on the state (x, K) . We examine two cases in turn. In Case A, the firm invests ($a = 1$) incurring for sure the fixed adjustment costs and thereby determine the Markov investment rule. We subsequently compare in Case B if the firm is better off investing — following the investment rule determined in Case A — or staying put ($a = 0$).

In Case A, the optimization problem (2.8) becomes

$$(2.9) \quad f^A(v, K) = \max_{I \in \mathbb{R}} \{vI - k(I, K)\}.$$

Assume the solution to (2.9) is given by the Markov investment rule ϕ^A :

$$\phi^A(v, K) \in \arg \max_{I \in \mathbb{R}} \{vI - k(I, K)\}, \quad (x, K) \in \mathcal{X} \times \mathbb{R}.$$

Note that (i) the program in equation (2.9) is strictly concave in I and (ii) the function $k(\cdot, K)$ is differentiable in I everywhere on $(-\infty, 0)$ and $(0, \infty)$. In these regions, a set of first-order conditions are sufficient to determine the optimal Markov rule. Whenever $k(\cdot, K)$ is differentiable, the first-order condition is

$$v = k_I(\phi^A(v, K), K).$$

The above relationship states that the shadow price of capital, $v(x, K)$, must equal the marginal cost of investing, $k_I(\phi^A(x, K), K)$ whenever investment or divestment takes place. We want to characterize these two regions in terms of the shadow price of investment. Note that $k(\cdot, K)$ is strictly convex, so that the first-order derivative $k_I(\cdot, K)$ is strictly increasing. Therefore, $k_I(I, K) < \lim_{I \uparrow 0} k_I(I, K) \triangleq \underline{k}_I(K)$ for all $I \in (-\infty, 0)$ and $k_I(I, K) > \lim_{I \downarrow 0} k_I(I, K) \triangleq \bar{k}_I(K)$ for all $I \in (0, \infty)$. These two regions are then equivalent to $v(x, K) \in (-\infty, \underline{k}_I(K)]$ and $v(x, K) \in [\bar{k}_I(K), \infty)$, respectively. Theorem 11 characterizes the optimal Markov strategy ϕ^A further.

THEOREM 11. *The optimal investment intensity $\phi^A(\cdot, K)$ is strictly increasing on $(-\infty, \underline{k}_I(K)]$ and $[\bar{k}_I(K), \infty)$.*

PROOF. We have $k_{II}(\cdot, K) > 0$. For the equality $v = k_I(\phi^A(v, K), K)$ to hold, $\phi^A(\cdot, K)$ must be strictly increasing in v . The result for $v \in [\underline{k}_I(K), \infty)$ obtains similarly. \square

What happens if the adjustment cost function $k(\cdot, K)$ is not differentiable at $I = 0$? In this case, the set $(K) \triangleq (\underline{k}_I(K), \bar{k}_I(K))$ is non empty. The optimal investment rule prescribes not to invest: $\phi(v, K) = 0$ if $(K) \triangleq (\underline{k}_I(K), \bar{k}_I(K))$.

THEOREM 12 (Optimal Markov control in Case A). *The optimal investment rule in Case A thus satisfies:*

$$(2.10) \quad \phi^A(v, K) = \begin{cases} < 0 & \text{if } v \in (-\infty, \underline{k}_I(K)) & [\text{divest}] \\ = 0 & \text{if } v \in (\underline{k}_I(K), \bar{k}_I(K)) & [\text{stay put}] \\ > 0 & \text{if } v \in [\bar{k}_I(K), \infty) & [\text{invest}] \end{cases}$$

We can now consider Case B and discuss whether the firm should stay put to avoid incurring fixed adjustment costs. For $v(x, K) \in (\underline{k}_I(K), \bar{k}_I(K))$, the optimal investment rule in (2.10) prescribes to stay put, the firm receiving $f^A(v, K) = -k(0, K) < 0$. In Case B, the value $-k(0, K)$ must constitute a lower bound for $f(v, K)$; otherwise, the firm would choose to adjust capital ($a = 1$) while scrapping or adding no capital unit. Outside $(\underline{k}_I(K), \bar{k}_I(K))$, $f(\cdot, K)$ is continuously differentiable in v . By differentiating $f^A(\cdot, K)$ in equation (2.9) with respect to the shadow price v , it obtains

$$(2.11) \quad f_v^A(v, K) = \phi^A(v, K),$$

where the Markov control ϕ^A is given in equation (2.10). We denote the optimal Markov investment rule in Case B by ϕ . The optimal investment magnitude $\phi(\cdot, K) = \phi^A(\cdot, K)$ is increasing outside $(\underline{k}_I(K), \bar{k}_I(K))$. Besides, $\phi(\cdot, K)$ is differentiable. Hence, $f_{vv}^A(v, K) = \phi_v^A(v, K) > 0$ for all $v \notin (\underline{k}_I(K), \bar{k}_I(K))$. $f^A(\cdot, K)$ is a convex function attaining its nonpositive minimum $-k(0, K_s)$ for $v \in (\underline{k}_I(K), \bar{k}_I(K))$. Let $\underline{v}(K)$ and $\bar{v}(K)$ denote the smallest and largest roots of $f^A(\cdot, K)$. From equations (2.10) and (2.11), $f^A(\cdot, K)$ is strictly decreasing in $(-\infty, \underline{v}(K)] \subseteq (-\infty, \underline{k}_I(K)]$, strictly increasing in $[\bar{v}(K), \infty) \subseteq [\bar{k}_I(K), \infty)$, and attain a nonpositive minimum in $(\underline{v}(K), \bar{v}(K))$. The above situation is shown in Figure 1 for the case of positive fixed, nonprice adjustment cost, $k(0, K) > 0$. If $v \in (\underline{v}(K), \bar{v}(K))$, $f^A(v, K) \leq 0$, so that the firm should stay put receiving zero. From the above, it results that the optimal investment behavior in general for Case B is as given in the following theorem.

THEOREM 13 (Optimal Markov control).

$$(2.12) \quad \phi(v, K) = \begin{cases} \phi^A(v, K) < 0 & \text{if } v(x, K) \leq \underline{v}(K) & [\text{divest}] \\ = 0 & \text{if } v(x, K) \in (\underline{v}(K), \bar{v}(K)) & [\text{stay put}] \\ \phi^A(v, K) > 0 & \text{if } v(x, K) \geq \bar{v}(K_s) & [\text{invest}] \end{cases}$$

For any value of $v(x, k)$ in $(\underline{v}(K), \bar{v}(K))$, investment does not take place. The firm sells capital (“divest”) to prevent the marginal revenue product of capital v from falling below a certain lower profitability level $\underline{v}(K)$ and purchases capital (“invest”) to prevent v from rising above an upper profitability level $\bar{v}(K)$. The open interval between those thresholds is an inaction region (hysteresis region) where the firm neither invests nor divests (“stay put” or “waits”). Within the bands $\underline{v}(K)$ and $\bar{v}(K)$, the shadow price process is driven solely by the exogenous shock (i.e., no control is exercised). At the upper barrier $\bar{v}(K)$, firms will be indifferent between staying put or expanding capacity, while at the lower barrier $\underline{v}(K)$, firms will be indifferent between holding the current level of capital stock constant or downsizing. Expansion at $\bar{v}(K)$ and downsizing at $\underline{v}(K)$ maximize the present discounted value of profits.

Does the function $f^A(\cdot, K)$ admit roots or equivalently saying is the hysteresis region $(\underline{v}(K), \bar{v}(K))$ nonempty?

- (i) If $k(\cdot, K)$ is differentiable at $I = 0$ or $\underline{k}_I(K) = \bar{k}_I(K_s)$ and $k(0, K) = 0$, the function $f^A(\cdot, K)$ has a *unique* root;

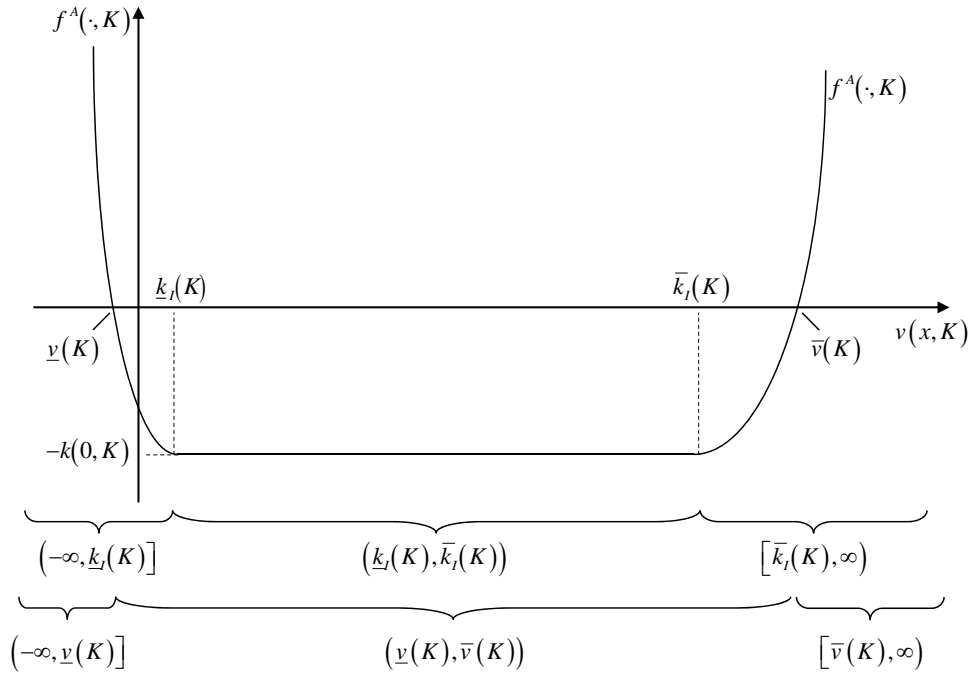


FIGURE 1. Convexity of reduced maximization program

NOTE: We assume $k(0, K) > 0$.

- (ii) If $k(0, K) > 0$, $f^A(\cdot, K)$ has *two* distinct roots (the differentiability of $k(\cdot, K)$ at $I = 0$ does not play a role here);
- (iii) If $k(\cdot, K)$ is not differentiable at $I = 0$ (or $\underline{k}_I(K) < \bar{k}_I(K)$ by strict convexity of $k(\cdot, K) \in C^2(\mathbb{R})$) and the fixed cost $k(0, K) = 0$, $f^A(\cdot, K)$ has a *continuum* of roots extending from $\underline{v}(K)$ to $\bar{v}(K)$.

If there are positive fixed adjustment costs — $k(0, K) > 0$ in case (ii) — or if $k(\cdot, K)$ is not differentiable at $I = 0$ — as in case (iii) — the set $(\underline{v}(K), \bar{v}(K))$ is nonempty. In other words, there is an inaction or hysteresis region where the firm stays put.

Consider now how the marginal value and capital stock evolve over time. At the *outset*, with initial state (X_{0-}, K_{0-}) , a discrete adjustment will occur if $v(X_{0-}, K_{0-}) \notin (\underline{v}(K_{0-}), \bar{v}(K_{0-}))$. Onwards, as the underlying stochastic shock $X = \{X_t\}_{t \geq 0}$ follows an unrestricted diffusion process, the investment and divestment decisions of the firm — as summarized in the control process ξ — implies that the endogenously determined process for the marginal value $\{v(X_t, K_t)\}_{t \geq 0}$ also follows a diffusion, with reflecting upper and lower barriers. The optimal policy involves adjusting the capital stock at each threshold, investing and divesting at the boundaries $\bar{v}(K_t)$ and $\underline{v}(K_t)$ respectively. The firm has to be very fast in adjusting its capital stock to maintain, at time t , the profitability level $v(X_t, K_t)$ within a certain range $(\underline{v}(K_t), \bar{v}(K_t))$. Since $\{v(X_t, K_t)\}_{t \geq 0}$ is a diffusion (with positive diffusion term), its sample paths have infinite first variation: though continuous almost everywhere, these sample paths are differentiable almost nowhere. At the points of investment or divestment, ξ moves infinitely fast, though it never increases by a discrete amount (except possibly at the outset): the rate of control is infinite. Hopefully, control only takes place at distinct moments in time, i.e., never throughout a nonempty time interval. In other words, this set of points has Lebesgue measure zero. The stochastic control process ξ is said to evolve in a *singular* manner: though continuous in the time interval $(0, \infty)$, it only increases on a time set — a collection of distinct points — that have Lebesgue measure zero. Properly speaking, since ξ is not absolutely continuous with respect to the Lebesgue measure, there is no rate of investment. Therefore, a naive application of Hamiltonian analysis is not appropriate. The engineering literature,

however, solves such abstract optimization problems by HJB methods using a limit argument to allow the rate of control to become infinite.

To summarize, in the presence of a convex adjustment cost the model delivers the conclusion from Tobin's (1969) q theory as modified by Abel (1983): in each state (x, K) , the investment intensity $\phi(\cdot, K)$ is increasing with the shadow price of capital, $v(x, K)$. With no adjustment cost, the model is the one in Jorgenson (1963). Namely, if the investment is costlessly reversible in the sense that the purchase and sale prices for capital are equal — $\underline{\kappa} = \bar{\kappa}$ such that $(\underline{v}(K), \bar{v}(K)) = \emptyset$ for all t almost everywhere — and that there are no nonprice adjustment costs ($k(0, K) > 0$), the model produces Jorgenson's (1963) result: at each date (in each state), the capital stock K is adjusted to equate the (ever-changing) marginal revenue product of capital $v(x, K)$ with the (constant) user cost of capital. Investment is completely irreversible - divestment never occurs - if the adjustment cost for divestment is prohibitively large and the sale price relatively small. A firm will find appropriate to sell installed capital only if the cost incurred to do so is negative or if the firm disposes of capital at a positive resale price. If there is no value of I for which $k(I, K)$ is negative, it will never be optimal for a firm to undertake downsizing initiatives.

Eberly and Van Mieghem (1997) consider dynamic investments in a *portfolio* of resources under uncertainty. Under concave operating profit functions and convex adjustment costs, the value function of the firm's discounted profits is concave. This implies the existence of a hysteresis zone and the optimality of invest/stay-put/divest (ISD) policy — as the authors name it. Similar to the above, an ISD policy indicates that, given the capacities of all resources, the optimal capacity decision for a given resource is defined by a lower and upper threshold (which depend on all other capital levels). Capacity is increased to the lower bound if the period's starting capacity is below it; it is decreased to the upper bound if starting capacity is above it, and it is kept constant if starting capacity is between the two bounds.

We now present two cases admitting closed-form solutions that can be subsumed under previous analysis of capital adjustment decisions.

EXAMPLE 5 (From Abel and Eberly, 1997). *Consider the investment cost function from Example 3 in Chapter 1, $k(I) = \kappa I + \kappa_1 I^{\frac{2n}{2n-1}}$. This cost function does not depend on the level of capital held and involves no fixed nonprice adjustment costs ($k(0) = 0$). In this case, the simplified optimization program in equation (2.9) reads*

$$f^A(v, K) = f^A(v) = \max_{I \in \mathbb{R}} \left\{ (v - \kappa) I - \kappa_1 I^{\frac{2n}{2n-1}} \right\}$$

The program is concave and differentiable everywhere in I (including at $I = 0$). From the first-order condition, the optimal investment rule is

$$\phi(v, K) = \phi(v) = \left(\frac{2n-1}{2n\kappa_1} \right)^{2n-1} (v - \kappa)^{2n-1}.$$

In the above equation, it is readily apparent that the firm adjusts its capital stock if the marginal value of capital, $v(x, K)$, exceeds the purchase (and resale) price of capital κ . In the quadratic case ($n = 1$), the optimal investment policy is characterized by

$$\begin{aligned} \phi(v) &= \frac{v - \kappa}{2\kappa_1} \\ &= \begin{cases} < 0 & \text{if } v < \kappa & [\text{divest}] \\ \geq 0 & \text{if } v \geq \kappa. & [\text{invest}] \end{cases} \end{aligned}$$

The above case describes costlessly reversible investments.

We now consider an example that makes it possible to derive — under certain restriction for the production function — the value to the firm that may adjust its capital stock $K = \{K_t\}_{t \geq 0}$ in reaction to exogenous market shocks $X = \{X_t\}_{t \geq 0}$.

EXAMPLE 6 (From Dixit and Pindyck, 1994, Chapter 11). Consider that $X = \{X_t\}_{t \geq 0}$ follows the geometric Brownian motion of equation (3.4), with $\mu < r$. Denote by $\delta \triangleq r - \mu (> 0)$ a measure of the opportunity cost of waiting. Besides, capital does not depreciate or $\lambda = 0$ in equation (3.1). The variable production costs are negligible. The technology offers constant returns to scale. The reduced-form profit flow reads $\pi(x, K) = x\pi(K)$ in state (x, K) . Consider investment perfectly irreversible with $\kappa(I) = \kappa I$ for $I \geq 0$; here, $\bar{\kappa}_I(K) = \kappa$. The optimal investment policy in equation (2.12) becomes

$$\phi(v) = \begin{cases} = 0 & \text{if } v(x, K) \leq \kappa, \\ > 0 & \text{if } v(x, K) > \kappa. \end{cases}$$

For $v(x, K) \leq \kappa$, as no investment takes place ($I_s = 0$), equation (2.4) becomes Feynman-Kac formula:

$$rV(x, K) = x\pi(K) + V_x(x, K)\mu x + \frac{1}{2}V_{xx}(x, K)\sigma^2 x^2.$$

The general solution to this ordinary differential equation has the form

$$V(x, K) = A(K)x^{\beta_1} + B(K)x^{\beta_2} + \frac{\pi(K)}{\delta}x;$$

here, $\beta_1 (> 0)$ and $\beta_2 (< 0)$ are, respectively, the positive and the negative root of the fundamental quadratic function $\mathcal{Q}(\cdot)$ in (1.4). The constant expressions $A(K)$ and $B(K)$ are determined by boundary conditions. To ensure that the value function V is finite as x approaches zero, we set $B(K) = 0$ as $\beta_2 < 0$. Denote by $\bar{X}(K)$ the (optimal) upper boundary for the shock that triggers investment. We now have

$$V_K(x, K) = A_K(K)x^{\beta_1} + \frac{\pi_K(K)}{\delta}x.$$

The value-matching condition states that $v(\bar{X}(K)) = V_K(\bar{X}(K)) = \kappa$, which yields now

$$A_K(K)\bar{X}(K)^{\beta_1} + \frac{\pi_K(K)}{\delta}\bar{X}(K) = \kappa.$$

The third boundary condition is the super-contact condition, a second-order condition for instantaneous control problem (Dumas, 1991). It is

$$v_x(\bar{X}(K), K) = \beta_1 A_K(K)\bar{X}(K)^{\beta_1-1} + \frac{\pi_K(K)}{\delta} = 0$$

From the previous equations, we obtain the threshold shock value $\bar{X}(K)$ together with the threshold shadow price $\bar{v}(K)$

$$\bar{X}(K) = \frac{\delta\kappa}{\pi_K(K)}\Pi^* \Leftrightarrow \bar{v}(K) = \bar{X}(K)\frac{\pi_K(K)}{\delta} = \Pi^*,$$

where $\Pi^* = \frac{\beta_1}{\beta_1-1}$. The optimal investment policy is

$$\phi(v, K) = \begin{cases} 0 & \text{if } v(x, K) < \bar{v}(K) & [\text{wait}] \\ > 0 & \text{if } v(x, K) \geq \bar{v}(K). & [\text{invest}] \end{cases}$$

The investment policy prescribes to wait until the marginal value of an additional unit of capital exceeds the required investment cost by the multiple $\Pi^* > 1$. By substitution, it obtains after a few algebra steps,

$$A_K(K) = -\left(\frac{\pi_K(K)}{\delta\beta_1}\right)^{\beta_1} \left(\frac{\beta_1-1}{\kappa}\right)^{\beta_1-1} (< 0)$$

Note that $A(K)x_1^\beta$ is the value, in state (x, K) , of the subsequent call options on future capital additions. As the marginal return is declining, the term $A_K(K) (< 0)$ corresponds to the opportunity cost of capital addition when the firm already holds K units of capital. We want to obtain a well-defined value for $A(K)$:

$$\begin{aligned}
A(K) &= \int_K^\infty [-A_K(k)] dk, \\
&= \left(\frac{\beta_1 - 1}{\kappa} \right)^{\beta_1 - 1} \left(\frac{1}{\delta \beta_1} \right)^{\beta_1} \int_K^\infty \pi_K(k)^{\beta_1} dk.
\end{aligned}$$

For the integral to converge, $\pi_K(K)$ must decrease sufficiently rapidly. To ensure that the production function satisfy the Inada conditions, we consider a Cobb-Douglas production function, which leads us to consider the reduced-form profit function of Example 1 in Chapter 1 where $\pi(K) = K^\alpha$; $\pi_K(K) = \alpha K^{\alpha-1}$, $\alpha \in (0, 1)$. For convergence of the integral, $\beta_1(\alpha - 1) > 1$. In other words, the elasticity α must be sufficiently less than one or the marginal return diminishes at a sufficiently quick rate.

3. References

McDonald and Siegel (1986) analyze under uncertainty when a firm should enter a market if it has no possibility to reverse its investment decision once it is operating in the market. Dixit (1989) analyzes the problem of entry and exit, generalizing insights obtained in McDonald and Siegel (1986). Relaxing the assumption of fixed lump sizes and considering instead perfectly divisible capital stocks may be advisable to achieve analytical elegance. Tobin (1969) analyzed the problem of capital accumulation in a deterministic setting; here, the optimal investment rate increases with the ratio of the market value to the replacement cost of the firm's capital (*average q*). Abel (1983) and Abel and Eberly (1994, 1996) modify the framework of Tobin to a stochastic environment. This variant of Tobin's *q* approach is referred to as *marginal q* in investment theory.

Social Optimality, Myopia, and Perfect Competition

A standard approach employed in economics to derive the equilibrium in *perfect competition* is to consider the problem of optimal resource allocation faced by a benevolent *social planner*. This logic applies in the context of investment under uncertainty as well. We here present some results concerning the social optimum and perfect competition under uncertainty, with a focus on the notion of myopic Nash equilibrium. Following Leahy (1993), there exists a connection between the aggregate behaviors of myopic firms and the competitive benchmark, a property known as *optimality of myopia*.

In Section 1, we derive the social optimum. In Section 2, we determine a number of equilibrium properties of the myopic Nash equilibrium. Section 3 discusses the connection between the social optimum, the myopic Nash equilibrium, and perfect competition. In Section 4, we introduce completion delays (of fixed, known duration) in a perfect-competition setting and explain how a modeler may derive the perfect competition equilibrium by a “simple” change of state space. Section 5 provides references to related academic works.

1. Social optimum

The equilibrium in perfect competition is connected with the problem faced by the social planner, whereby the social optimum is the solution of a one-agent dynamic optimization problem. We first discuss the stochastic control problem faced by the social planner. As before, we concentrate on trigger policies.

A firm’s profit depends on the aggregate capital stock held in the industry. For simplicity, we consider a one-dimensional capital stock $k \in \mathbb{R}_+$. Suppose an exogenous shock — modeled as a diffusion — affects a firm’s profit. If the exogenous state is $x \in \mathcal{X} \subseteq \mathbb{R}$ and the industry capital stock k , the market-clearing price is $D(x, k)$ with $D(x, 0) = 0$ for all $x \in \mathcal{X}$. The inverse demand function $D \in C^{2,2}(\mathcal{X}, \mathbb{R}_+)$ is increasing in the shock and concave in the capital stock. In state (x, k) , the variable production costs are $c(x, k)$. $\pi(x, k) \triangleq D(x, k) - c(x, k)$ is a firm’s instantaneous (reduced-form) profit obtained in state (x, k) . Besides, the firms employ a constant-returns-to-scale production technology, so that the firms produce at full capital (or $g(k) = k$). Assume investments and divestments can be made instantaneously with no fixed adjustment cost other than the purchase or resale price; the investment cost function is of the kinked piece-wise linear form:

$$(3.3) \quad \kappa : I \mapsto \bar{\kappa} I^+ + \underline{\kappa} I^-.$$

In other words, the firm may acquire capital at unit cost $\bar{\kappa}$ or scrap unused capital units for a price of $\underline{\kappa}$. To preclude arbitrage opportunities, we assume $\underline{\kappa} < \bar{\kappa}$. We simplify the analysis and assume that capital does not depreciate ($\lambda = 0$).

We now look at the aggregate problem faced by the social planner. As usual, the consumer surplus corresponds to the area under the demand curve; it is given by

$$U(\cdot, k) \triangleq \int_0^k D(\cdot, K) dK.$$

The marginal social utility is $U_k(\cdot, k) = D(\cdot, k)$. The total variable production costs is

$$C(\cdot, k) \triangleq \int_0^k c(\cdot, K) dK.$$

The flow of social operating surplus is thus

$$\Pi(\cdot, \cdot) \triangleq U(\cdot, \cdot) - C(\cdot, \cdot).$$

$U(\cdot, \cdot)$, resp. $C(\cdot, \cdot)$, is concave, resp. convex, in its second argument. $\Pi(\cdot, \cdot)$ is decreasing and concave in k ; it inherits the continuity and integrability properties of D and c . The social surplus and firm profits are connected via the relationship

$$(1.1) \quad \Pi(\cdot, k) = \int_0^k \pi(\cdot, K) dK.$$

In other words, the marginal social contribution of an additional unit of capital is the profit of an atomic firm: $\Pi_k(\cdot, k) = \pi(\cdot, k)$. The following example gives a specific case of social surplus function used in real options analysis.

EXAMPLE 7 (From Grenadier, 2000b). *Consider a firm that produces at full capacity faces isoelastic demand $D(x, k) = x k^{-1/\eta}$ of equation (2.2). Production costs are negligible. The social surplus is given by*

$$\Pi(x, k) = \int_0^k x K^{-1/\eta} dK = x \left[\frac{K^{g(1, \eta)}}{g(1, \eta)} \right]_0^k = x \frac{k^{g(1, \eta)}}{g(1, \eta)},$$

with $g(1, \eta) \triangleq 1 - 1/\eta$.

From a dynamic perspective, the social planner receives at each instant t a social surplus $\Pi(X_t(\omega), K_t(\omega))$ depending on the shock realization $X_t(\omega)$, and the prevalent aggregate industry capital stock, $K_t(\omega)$. $X = \{X_t\}_{t \geq 0}$ and $K = \{K_t\}_{t \geq 0}$ are stochastic processes. In this case, the (singular) control of the social planner is the unique process ξ that describe the capital stock evolution beyond the initial stock, namely

$$K_t(\omega) = k + \xi_t(\omega).$$

Equivalently, $\xi_t(\omega)$ can be viewed as the cumulated investment flows that occurred before time t . The positive (resp., negative) part of ξ is noted ξ^+ (ξ^-). The decomposition $\xi \triangleq \xi^+ - \xi^-$ is unique, whereby the processes ξ^+ and ξ^- are left-continuous, nondecreasing and their sample paths have bounded variations. Ξ is the space made up of such processes. The social planner adjusts the industry capital stock with a view to maximizing expected social surpluses. Investment in capital stock has no effect on the mean and variance coefficients $\mu(x)$ and $\sigma(x)$ of the diffusion. The objective function of the social planner is:

$$(1.2) \quad J(x, k, \xi) = \mathbb{E} \left[\int_0^\infty e^{-rt} \Pi(X_t, k + \xi_t) dt - \int_0^\infty e^{-rt} \underline{\kappa} d\xi_t^- - \int_0^\infty e^{-rt} \bar{\kappa} d\xi_t^+ \mid x, k \right],$$

where the two last integrals are understood (pathwise) as Lebesgue-Stieljes integrals. If a social optimum exists, it yields value

$$(1.3) \quad S(x, k) \triangleq \max_{\xi \in \Xi} J(x, k, \xi).$$

Alternatively, the social planner may decide not to invest at all. The perpetuity value of social surpluses if the industry capital stock stays at its initial level k is

$$(1.4) \quad \hat{S}(x, k) = J(x, k, 0) = \mathbb{E} \left[\int_0^\infty e^{-rt} \Pi(X_t, k) dt \mid x, k \right].$$

The social optimum is characterized by a system of equations (1.5a)–(1.5c) that the value function $S(x, k)$ must satisfy. In the low-demand region, i.e., if $(x, k) \in \{(x, k) \mid x \leq \underline{X}(k)\}$, we have

$$(1.5a) \quad \begin{cases} S_k(x, k) &= \underline{\kappa}, \\ S_{kx}(x, k) &= 0. \end{cases}$$

Besides, the ordinary differential equation

$$(1.5b) \quad r S(x, k) = \Pi(x, k) + \mu(x) S_x(x, k) + \frac{1}{2} \sigma(x)^2 S_{xx}(x, k)$$

must be satisfied in the hysteresis region, i.e., if $(x, k) \in \{(x, k) \mid x \in (\underline{X}(k), \overline{X}(k))\}$. Finally, for high demand, i.e., if $(x, k) \in \{(x, k) \mid x \geq \overline{X}(k)\}$,

$$(1.5c) \quad \begin{cases} S_k(x, k) &= \bar{\kappa}, \\ S_{kx}(x, k) &= 0. \end{cases}$$

2. Myopic Nash equilibrium

A *strategic* firm takes account of the impact of future investments on its value function. Firm i maximizes its value and assuming rational capital adjustments by its rivals, receives at time t project value

$$(2.1) \quad V(x, k) \triangleq \max_{\xi^i \in \Xi^i} \mathbb{E} \left[\int_0^\infty e^{-rt} \pi(X_t, K_t) dt - e^{-rt} \underline{\kappa} d\xi_t^{i-} - e^{-rt} \bar{\kappa} d\xi_t^{i+} \mid x, k \right].$$

One way to approach the oligopoly problem is to consider the notion of myopia in investment behaviors. A myopic firm is farsighted in calculating present values but shortsighted in that it decides as if the price process were not affected by other firms' investments. The myopic firm has static expectations regarding the capital stock but rational expectations for the exogenous shock fluctuations. A *myopic* firm is shortsighted and invests as if no further investment were to occur in the industry, i.e., as if K_t will remain fixed at its current level k for all $t \in \mathbb{R}_+$. It behaves rationally under this restriction, receiving for initial state (x, k)

$$(2.2) \quad \hat{V}(x, k) \triangleq \max_{\xi^i \in \Xi^i} \mathbb{E} \left[\int_0^\infty e^{-rt} \pi(X_t, k) ds \mid x, k \right].$$

Consider first an oligopoly with n symmetric firms ($n < \infty$) producing a single, homogeneous, nonstorable good. Each firm i adopts an investment strategy ξ^i . Firm i 's capital stock and the industry capital stock at time t (on path $\omega \in \Omega$) are, respectively,

$$K_t^i(\omega) = k^i + \xi_t^i(\omega), \quad \text{and} \quad K_t(\omega) = k + \sum_{i=1}^n \xi_t^i(\omega).$$

We note the strategy profile by $\xi = (\xi^i, \xi^{-i})$. At places, K_t is a n -dimensional random variable. The payoff to firm i is

$$(2.3) \quad J^i(x, k, \xi) \triangleq \mathbb{E} \left[\int_0^\infty e^{-rt} \pi^i(X_t, K_t) dt - \int_0^\infty e^{-rt} \underline{\kappa} d\xi_t^{i-} - \int_0^\infty e^{-rt} \bar{\kappa} d\xi_t^{i+} \mid x, k \right].$$

Firm i chooses its investment strategy ξ^i to maximize its payoff. A *Nash equilibrium* in myopic strategies is a n -tuple $(\xi^{1*}, \dots, \xi^{n*})$ such that, for each player $i \in N$,

$$\xi^{i*} \in \arg \max_{\xi^i \in \Xi^i} J(x, k; \xi^i, \xi^{-i*}).$$

Since the market-clearing price generally declines following investment, a firm will invest for values greater than the perpetuity value of capital in order to cover (partly) sunk capital-adjustment costs. Denote by $\hat{v}^i(\overline{X}(k), k)$ the upper level of profitability that triggers investment and $\hat{v}^i(\underline{X}(k), k)$ the divestment trigger. Since the capital-adjustment cost function in equation (3.3) is not differentiable at $I = 0$, the hysteresis set $(\underline{X}(k), \overline{X}(k))$ is nonempty. The value received by firm i if a Nash equilibrium in myopic strategies exists is $\hat{V}^i(x, k)$. In the hysteresis set, the marginal value $\hat{v}^i(x, k)$ solves the ordinary differential equation

$$(2.4a) \quad r\hat{v}^i(x, k) = \pi_{k^i}^i(x, k) + \mu(x) \hat{v}_x^i(x, k) + \frac{1}{2} \sigma(x)^2 \hat{v}_{xx}^i(x, k), \quad x \in (\underline{X}(k), \overline{X}(k)).$$

At the boundaries of the hysteresis set, we have by value matching

$$(2.4b) \quad \begin{cases} \hat{v}^i(\overline{X}(k), k) = \bar{\kappa}, \\ \hat{v}^i(\underline{X}(k), k) = \underline{\kappa} \end{cases}$$

Optimality is obtained by the high-contact or super-contact conditions

$$(2.4c) \quad \begin{cases} \hat{v}_x^i(\bar{X}(k), k) = 0, \\ \hat{v}_x^i(\underline{X}(k), k) = 0. \end{cases}$$

We consider, in Examples 8 and 9, two cases of myopic Nash equilibria that admit closed-form solutions.

EXAMPLE 8 (From Grenadier, 2002). *Assume the n firms expand their capital all at the same time and in the same proportion with $d\xi_t^i = d\xi_t/n$, with no possibility of scrapping. Assume a constant-returns-to-scale production technology and the isoelastic demand function*

$$D(x, k) = x k^{-1/\eta},$$

with $\eta > 1$. Firm i has negligible production costs and faces a profit function $\pi(x, k) = D(x, k) k^i$. The multiplicative demand shock X follows the geometric Brownian motion of equation (3.4), with $\delta \triangleq r - \mu (> 0)$. β_1 and β_2 are the positive and negative roots of the fundamental quadratic in equation (1.4). We assume $\eta > \beta_1$. Firm i 's marginal profit obtained from increasing its capital stock is

$$\pi_{k^i}(x, k) = x g(s^i, \eta) k^{-1/\eta},$$

where $s^i \triangleq k^i/k$ is firm i 's market share and $g(s^i, \eta) \triangleq 1 - s^i/\eta$. The differential equation (2.4a) then specializes to

$$r\hat{v}^i(x, k) = x g(s^i, \eta) k^{-1/\eta} + \mu x \hat{v}_x^i(x, k) + \frac{1}{2}\sigma^2 x^2 \hat{v}_{xx}^i(x, k),$$

which, in the hysteresis region, has a solution of the form

$$v(x, k) = \frac{\pi_{k^i}(x, k)}{\delta} + A(k) x^{\beta_1} + B(k) x^{\beta_2}.$$

To ensure finiteness of the value function at the origin ($\lim_{x \downarrow 0} v(x, K) = 0$), we must have $B = 0$. From the value-matching condition (2.4b) at the upper trigger,

$$\frac{\bar{X}(k) g(s^i, \eta) k^{-1/\eta}}{\delta} + A(k) \bar{X}(k)^{\beta_1} = \bar{\kappa}.$$

Applying the super-contact condition (2.4c) to the general solution, it follows

$$\frac{g(s^i, \eta) k^{-1/\eta}}{\delta} + \beta_1 A(k) \bar{X}(k)^{\beta_1 - 1} = 0 \quad \Leftrightarrow \quad A(k) = -\frac{\bar{X}(k)^{1 - \beta_1} g(s^i, \eta) k^{-1/\eta}}{\delta \beta_1}.$$

Substituting the expression for $A(k)$ back, it results

$$\frac{g(s^i, \eta) k^{-1/\eta}}{\delta} \left[\bar{X}(k) - \frac{\bar{X}(k)^{1 - \beta_1}}{\beta_1} \right] = \bar{\kappa},$$

eventually obtaining (cf. Grenadier, 2002, equation 21)

$$\bar{X}(k) = \Pi^* g(s^i, \eta) \bar{\kappa} \delta k^{\frac{1}{\eta}} \Leftrightarrow \frac{\bar{X}(k) g(s^i, \eta) k^{\frac{1}{\eta}}}{\delta \bar{\kappa}} = \Pi^*$$

with $\Pi^* \triangleq \beta_1/(\beta_1 - 1)$. The closed-form solution for the marginal value function is

$$v(x, k) = \frac{\pi_{k^i}(x, k)}{\delta} - \frac{\pi_{k^i}(\bar{X}(k), k)}{\delta \beta_1} \left(\frac{x \wedge \bar{X}(k)}{\bar{X}(k)} \right)^{\beta_1}.$$

EXAMPLE 9 (From Baldursson, 1998). *Consider the capital-adjustment cost function (3.3) with $\underline{\kappa} = 0$. Firm i faces linear demand (equation 2.1 with $a = 1$) and has unit production cost c . $X = \{X_t\}_{t \geq 0}$ follows the geometric Brownian motion as per equation (3.4). Firm i 's profit function is*

$\pi(x, k) = (x - bk)k^i - ck^i$, obtaining marginal profit $\pi_{k^i}(x, k) = x - 2bk^i - \sum_{j \neq i} bk^j - c$. To ensure that $\lim_{x \downarrow 0} v(x, K) = 0$, $B(k) = 0$. Solving for the value-matching and the super-contact conditions, (2.4b) and (2.4c), gives

$$\bar{X}(k) = \Pi^* \frac{\delta}{r} (c + r + 2bk^i + bk^i),$$

and the marginal value function

$$v(x, k) = \frac{-2bK^i - bk^{-i} - c}{r} + \frac{x}{\delta} - \frac{x^{\beta_1}}{\delta \beta_1 \bar{X}(k)}.$$

3. Perfect competition

The analysis of investment behaviors in perfect competition is typically build upon the premises that capital is perfectly divisible. Thereby, one does not explicitly look at the number of incumbents firms but rather consider *atomic* firms that each holds an negligible share of the aggregate industry capital stock with no material influence on the price-setting process.

In perfect competition, the industry capital stock process ξ is *not* a decision parameter of the atomic firm. It only reflects the belief about the change in the industry capital stock. Recall that a stopping time is a mapping $\tau : \Omega \rightarrow \mathbb{R}_+$ that do not allow foresight into the future, namely $\{\tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. For an arbitrary investment process ξ , the stopping times $\bar{\tau}(k) \triangleq \inf \{t \geq 0 | \xi_t^+ > 0\}$ and $\underline{\tau}(k) \triangleq \inf \{t \geq 0 | \xi_t^- > 0\}$ indicate, respectively, when capital is added and reduced for the first time. A change in the capital stock occurs for the first time at (stopping) time $\tau(k) \triangleq \underline{\tau}(k) \wedge \bar{\tau}(k)$. We can decompose the payoff function (2.3) as

$$\begin{aligned} J(x, k, \tau, \xi) &= \mathbb{E} \left[\int_t^\infty e^{-rt} \{ \pi(X_t, K_t) dt - \underline{\kappa} d\xi_t^{i-} - \bar{\kappa} d\xi_t^{i-} \} \mid x, k \right] \\ &= \mathbb{E} \left[\int_0^{\tau(k)} e^{-rt} \pi(X_t, k) ds + \int_{\tau(k)}^\infty e^{-rt} \pi(X_t, k^i + \xi_t^i) dt - \right. \\ &\quad \left. e^{-r\tau(k)} \underline{\kappa} \mathbb{1}_{\{\tau(k)=\underline{\tau}(k)\}} - e^{-r\tau(k)} \bar{\kappa} \mathbb{1}_{\{\tau(k)=\bar{\tau}(k)\}} \mid x, k \right] \\ &= \hat{V}(x, k) + \gamma(x, k), \end{aligned}$$

where the term

$$\begin{aligned} \gamma(x, k) &\triangleq \mathbb{E} \left[\int_{\tau(k)}^\infty e^{-rt} [\pi(X_t, K_t) - \pi(X_t, K_{\tau(k)})] ds - e^{-r\tau(k)} \underline{\kappa} \mathbb{1}_{\{\tau(k)=\underline{\tau}(k)\}} \right. \\ &\quad \left. - e^{-r\tau(k)} \bar{\kappa} \mathbb{1}_{\{\tau(k)=\bar{\tau}(k)\}} \mid x, k \right]. \end{aligned}$$

Suppose a strategy profile $\xi = (\xi^i, \xi^{-i})$ and payoff functions $J_1^i, J_2, J_3^i : \otimes_{i=1}^n \Xi^i \rightarrow \mathbb{R}$. $J_2(\cdot)$ is called *fictitious objective function* if there exists $J_3^i(\cdot)$ such that $J_1^i(\xi) = J_2(\xi) + J_3^i(\xi^{-i})$. Suppose now the n -tuple ξ^* maximizes $J_2(\cdot)$. Then,

$$\begin{aligned} \sup_{\xi^i \in \Xi^i} J_1^i(\xi^i, \xi^{-i*}) &= \sup_{\xi^i \in \Xi^i} J_2(\xi^i, \xi^{-i*}) + J_3^i(\xi^{-i*}) \\ &= J_2(\xi^*) + J_3^i(\xi^{-i*}). \end{aligned}$$

Hence, ξ^* is a Nash equilibrium. In other words, if a strategy profile ξ maximizes the fictitious objective function $J_2(\cdot)$, then ξ is a Nash equilibrium of the game $\Gamma = \{N, (\Xi_i)_{i \in N}, (J_1^i)_{i \in N}\}$.

From the above, we obtain:

$$\xi^{i*} \in \arg \max_{\xi^i \in \Xi^i} V(x, k) \Leftrightarrow \xi^{i*} \in \arg \max_{\xi^i \in \Xi^i} \hat{V}(x, k).$$

The above establishes the duality between the myopic firm's problem and the social planner's, stating that myopic investors invest at the same time(s) as a social planner would do. Although a competitive firm's deferral option is worthless, the option is still valuable for society as a whole. The issues raised in real options analysis do not disappear in perfect competition, they are simply shifted to the aggregate.

We now investigate the investment strategy a benevolent social planner would impose on decentralized firms. To a given social planner's strategy choice ξ , we associate the stopping times $\underline{\tau}(k) \triangleq \inf \{t \geq 0 \mid \xi_t^+ > 0\}$ and $\bar{\tau}(k) \triangleq \inf \{t \geq 0 \mid \xi_t^- > 0\}$ that indicate respectively when investment and divestment takes place. Note the equivalence for $t \in \mathbb{R}_+$ between $\bar{\tau}(\omega, K) < t$ and $\xi_t^+(\omega) > K$ on the one hand and $\underline{\tau}(\omega, K) < t$ and $\xi_t^-(\omega) > K$ on the other hand.

The myopic case is easier to handle. Applying Itô-Doebelin formula and considering the drift part, we have

$$\mathcal{D}_{X,K}\Pi(x, k) = \Pi_k(x, k)dK + \mu(x) \Pi_x(x, k) + \frac{1}{2}\sigma(x)^2 \Pi_{xx}(x, k)$$

From Dynkin's formula, we have

$$\begin{aligned} & \mathbb{E}\left[\Pi(X_t, K_t) \mid x, k\right] \\ &= \Pi(x, k) + \mathbb{E}\left[\int_0^t \mathcal{D}_x \Pi(X_s, K_s) ds \mid x, k\right] + \mathbb{E}\left[\int_0^t \int_k^\infty \pi(X_s, K) dK ds \mid x, k\right] \\ &= \mathbb{E}[\Pi(X_t, K_t) \mid x, k] + \int_k^\infty \mathbb{E}\left[\int_0^t \pi(X_s, K_s) ds \mid x, k\right] dK. \end{aligned}$$

Applying Fubini's theorem, we have

$$\begin{aligned} & \mathbb{E}\left[\int_0^\infty e^{-rt} \Pi(X_t, K_t) dt \mid x, k\right] \\ &= \int_0^\infty e^{-rt} \mathbb{E}[\Pi(X_t, K_t) \mid x, k] dt \\ &= \int_0^\infty e^{-rt} \mathbb{E}[\Pi(X_t, k) \mid x, k] dt + \int_k^\infty \mathbb{E}\left[\int_0^t \pi(X_s, K_s) dt \mid x, k\right] dK \end{aligned}$$

Applying Dynkin's formula, we now obtain:

$$(3.1) \quad S(x, k) = \hat{S}(x, k) + \int_k^\infty \hat{V}(x, K) dK.$$

The second right hand term captures the effect of investment and divestment to maintain at each time the social profit within a certain range of profitability. The socially optimal investment process ξ is expressed in terms of the optimal myopic investment policy. The social surplus obtained by a benevolent social planner equals the social surpluses earned when selfish, *myopic* investors invest and divest at their optimal thresholds. The payoff to the central planner is obtained by integrating the payoffs to the atomic firms. The social planner's problem can be seen either as a one-player stochastic control problem where the social planner adjusts the industry capital stock to maximize the expected social surpluses (see Section 1), or it can be seen as a series of stochastic control problems where each myopic firm chooses its investment and divestment triggers to maximize its own value (see Section 3).

Back and Paulsen (2009) prove that the myopic Nash equilibrium does not correspond to the Markov perfect equilibrium, because overinvestment is a profitable deviation. As the market share of a firm decreases, however, the myopic equilibrium yields a value which converges to the firm value under perfect competition.

4. Completion delays

Grenadier (2000a) extends the model by Leahy (1993) to allow for completion delays. At the first sight, completion delays may cause a certain path dependency that altogether affect the investment strategy choice of a firm; a firm might be reluctant to invest in capital in favorable situations if it knows that it will have soon new capital units at disposal. Future market-clearing prices $\{D(X_s, K_s)\}_{s \geq t}$ would depend not only on the expectations about the shock realizations $\{X_s\}_{s \geq t}$ conditional on current information X_t but also on past decisions relating to future additions to the capital stock. To know the arrival times of units under construction, firms need to remember the investment decisions over the past D years. At any time t , one considers both the units of capital $K_t(\omega)$ that are completed and already

available for production and the units that are currently under construction $D_t(\omega)$. In other words, $D_t(\omega)$ is the number of units that have begun construction during the time interval $(t - D, t]$ if $t > D$ or $(0, t]$ otherwise. At any time t , the social planner needs to know the current state of the shock $X_t(\omega)$ and the current capital stock $K_t(\omega)$ because both $X_t(\omega)$ and $K_t(\omega)$ determine current (reduced-form) profit flow. However, it must also consider the number of units in the pipeline D_t and the times at which these units will be completed: $T_t(\omega) \triangleq \{s \in [t - D, t] \mid K_t^D(\omega) > K_{t-s}^D(\omega)\}$ memorizes when investment decisions have been taken on a given path $\omega \in \Omega$. These factors affect the social surpluses over the ensuing D years. The relevant states are $X_t(\omega)$, $K_t(\omega)$, $D_t(\omega)$, and $T_t(\omega)$. For notational simplicity, we set $Y_t(\omega) = (X_t(\omega), K_t(\omega), D_t(\omega), T_t(\omega))$.

This time dependency might make it impossible to pin down a well-defined solution to the control problem. The apparent difficulty in deriving optimal controls in such situations is that the state space can be of infinite dimension and Markov policies would be biased. Especially, the set $T_t(\omega)$ may be of infinite dimension. Without time to build, the payoff-relevant states are simply $X_t(\omega)$ and $K_t(\omega)$. Unless the model can be converted into one of finite dimension, the standard solution techniques cannot be employed. The amount and arrival times of units currently “in the pipeline” are relevant because they will affect the market prices over the ensuing D years of construction. The social planner’s problem is

$$(4.1) \quad S(y) \triangleq \max_{\xi \in \Xi} \mathbb{E} \left[\int_0^\infty \Pi(X_t, K_t) dt - \int_0^\infty e^{-rt} \underline{\kappa} d\xi_t^- - \int_0^\infty e^{-rt} \bar{\kappa} d\xi_t^+ \mid y \right].$$

One can simplify the state space by adding a new state variable coined “committed capacity.” The latter state corresponds to the aggregation of already existing capital stocks and capital units in the pipeline.

The solution approach is thus to consider an *artificial* industry in which all units currently in the pipeline are assumed to be completed. In this industry, two states are relevant for determining the future optimal controls: the current realization of the shock $X_t(\omega)$ and the total supply of *committed capital*. Let k^D denote the committed capital, i.e., $K_t^D(\omega) = K_t(\omega) + D_t(\omega)$. Note that $K_t(\omega) = K_{t-D}^D(\omega)$ for $t > D$ on path $\omega \in \Omega$. Knowledge of the current shock value $X_t(\omega)$ and of the level of committed capacity $K_t^D(\omega)$ obtains to be sufficient to determine the equilibrium investment strategy. This transformation of the state space was demonstrated by Bar-Ilan and Strange (1996) in a model of production capacity with time to build. Because a firm starting construction will not receive economic profits for D years, the firm decides based on profits to be received from D years onwards. The decision to build a new unit of capacity is analogous to the decision to exercise a call option on an asset $D (> 0)$ years from completion. Here, in contrast to the standard real options competitive equilibrium approach, where the process of the *actual* marginal values follows a diffusion process with reflections, the *expected* value prevailing at completion D years from now follows such a diffusion.

Consider again the social surplus in equation (4.1) obtaining

$$(4.2) \quad S(y) = \mathbb{E} \left[\int_0^D e^{-rt} \Pi(X_t, K_t) dt \mid y \right] + \max_{\xi_{[D, \infty)}} \mathbb{E} \left[\int_D^\infty e^{-rt} \Pi(X_t, K_t) dt - e^{-rt} \underline{\kappa} d\xi_t^- - e^{-rt} \bar{\kappa} d\xi_t^+ \mid y \right]$$

The choice of strategy $\xi_{[D, \infty)}$ by the social planner does not affect the social surplus $\Pi(X_t, K_t)$ over the interval $(0, D)$ since newly added capital units are not yet available. The second right-hand side term in equation (4.2) does depend on the current level of committed capital units $K_t^D(\omega)$ but not the state $T_t(\omega)$ because in D years all units currently under construction will then be completed. Therefore, some information in $Y_t(\omega)$ are not needed. Considering a state set $Y_t'(\omega) = (X_t(\omega), K_t^D(\omega))$ is sufficient without loss of information. Equation (4.2) can be transformed into

$$\begin{aligned} S(y) &= \mathbb{E} \left[\int_0^D e^{-rt} \Pi(X_t, K_t) dt \mid y \right] + \max_{\xi_{[D, \infty)}} \mathbb{E} \left[\int_D^\infty e^{-rt} \Pi(X_t, K_t) dt - \int_D^\infty \underline{\kappa} d\xi_t^- - \int_D^\infty \bar{\kappa} d\xi_t^+ \mid y' \right] \\ &= S(x, k, 0, \emptyset) - \mathbb{E} \left[\int_0^D e^{-rt} \Pi(X_t, K_t) dt \mid y' \right] + \mathbb{E} \left[\int_0^D e^{-rt} \Pi(X_t, K_t) dt \mid y \right]. \end{aligned}$$

Thus, to derive the optimal investment policy ξ^* for $S(y)$, we can turn the simpler problem $S(y')$. In other words, the original problem has been transformed into a simpler, Markov state space problem in an *artificial economy*.

5. References

Based on the rational-expectation framework introduced by Lucas and Prescott (1971), Leahy (1993) sets the perfect-competition benchmark for continuous-time analysis of investment in (infinitely divisible) capital stock, allowing firms facing kinked piecewise capital-adjustment cost function to invest and divest. Leahy proves the social optimality of *myopia* in perfect competition: myopic firms invest at the same time as firms ignoring potential future capital stock changes would. Dixit (1991) discusses a perfect-competition model in lines with Leahy (1993). Baldursson (1998) considers a similar problem employing the approach of a *fictitious* social planner. Baldursson and Karatzas (1996) consider capital expansion in a very general set-up for the underlying stochastic processes and build upon the connection between singular control and optimal stopping. Grenadier (2000b) introduces completion delays in the setting of Leahy (1993). Grenadier (2002) considers an oligopoly where n firms adopt a myopic investment behaviors. Back and Paulsen (2009) prove that the equilibrium in Grenadier (2002) fails Markov perfection and prove the perfect competition equilibrium is perfect. Grenadier (2000b) employs a real options approach to derive dynamic equilibrium for perfect competition for a market subject to time-to-build.

Part 3

Imperfect Competition

We considered so far strategic considerations as being inconsequential. When real options are available to a few firms, such an approach is ill advised. The benchmark cases have been studied with use of techniques from stochastic control theory. In the following three chapters — each a research contribution submitted to a refereed journal — we investigate generalizations of the previous insights to situations involving multiple firms that hold similar investment opportunities, called *shared options*. Most models of irreversible investment in perfectly or imperfectly competitive markets (see, e.g., Grenadier, 2002; Baldursson, 1998) have adopted a mixed open-loop/Markov approach without explicitly referring to it and without properly discussing the implications of the assumed information structure. They consider myopic strategies and use the Nash equilibrium as the appropriate solution concept to derive their market equilibrium.

In Chapter 7, we provide a review over other contributions addressing the intersection of investment under uncertainty and industrial organization; this overview is aimed at categorizing and relating the research streams. We highlight there managerial insights concerning the nature of competitive advantage (first- versus second-mover advantage), the manner in which information is revealed, firm heterogeneity, capital increment size, and the number of competing firms. Finally, we conclude and discuss research avenues yet to be explored to extend the reach of contribution in this literature. Chapter 8 discusses a market-entry game under uncertainty. Chapter 9 investigates strategic interactions in a setting where competing firms have to decide on how much to invest or divest.

Strategic Investment under Uncertainty: A Synthesis

ABSTRACT. Investment is a central theme in economics, finance, and operations management. Traditionally, the focus of analysis has been either on assessing the value of flexibility (investment under uncertainty) or on describing commitment effects in competitive settings (industrial organization). Research contributions addressing the intersection of investment under uncertainty and industrial organization have become numerous in recent years. In this paper, we provide an overview aimed at categorizing and relating these research streams. We highlight managerial insights concerning the nature of competitive advantage (first- versus second-mover advantage), the manner in which information is revealed, firm heterogeneity, capital increment size, and the number of competing firms.

1. Introduction

Investment theory seeks to equip decision makers with quantitative tools for assessing investment projects. As emphasized by ?, the development of valuation models featured three stages. In *static* models, an investment project is completely described by a specified stream of expected cash flows with no managerial flexibility. A noted method anchored in this perspective is the “discounted cash flow” approach. In *dynamic* models, projects can be actively managed in response to the resolution of exogenous uncertainty. This approach is embedded in decision-tree analysis, stochastic dynamic programming, stochastic programming, and real options analysis (ROA). Dixit and Pindyck (1994) and Trigeorgis (1996) address such issues using the tools of stochastic dynamic programming and contingent claims analysis. Even though the theory has been refined over time to incorporate many real-world features, the lack of explicit consideration given to strategic interactions often creates a partial picture in competitive settings. del Sol and Ghemawat (1999) highlight the risk of ignoring competition in investment decisions. In combined real options and game-theoretic models, firms can presumably condition their decisions not only on the resolution of exogenous uncertainties but also on the (re)actions of outside parties (e.g., competitors). Future cash flows can be understood as the payoffs of a game involving several decision makers and “nature.” The third approach — commonly called *option games* — was first proposed in the early 1990s and has received attention in academia ever since. Option games analysis examines the trade-off that arises, under uncertainty, between managerial flexibility and commitment in dynamic competitive settings.

A firm can effectively make an early strategic investment that alters the later game structure by inducing asymmetry among firms. In deterministic settings, the sign of the *strategic effect* depends on the intent of the commitment and the type of competitive reaction.¹ Early “overinvestment” is optimal when the strategic effect is positive: when the entry of rivals acts as a negative externality for incumbents, committing to an irreversible investment typically has a positive strategic effect because the leading firm acquires a first-mover advantage and can foreclose the market to potential entrants. When the strategic effect is negative, however, a firm should refrain from investing. In stochastic settings, uncertainty may erode the incentive to make a sunk investment given that future prospects are risky. Any initial investment decision requires the committing firm to weigh the opportunity cost of a commitment against

¹Fudenberg and Tirole (1984) decompose the drivers that affect the sign of the strategic effect. An investment is either “tough” or “soft” depending on whether the investment is meant to make the rival worse or better off, respectively. If firm (re)actions are “strategic substitutes,” then firm j will respond “less” to an aggressive action by firm i . Conversely, firms’ (re)actions can be “strategic complements.”

the expected future strategic benefits resulting from it. Dealing with such a dilemma is commonplace in business practice, and decision makers therefore require appropriate tools. Dyson (2000) suggests that operational research should provide a means for such strategic decision making. Ferreira, Kar, and Trigeorgis (2009) stress the relevance of option games analysis for strategic management formulation.

This paper provides an overview of the literature and stresses new insights from a selection of articles.² We classify these research contributions and structure the paper as follows. Section 2 deals with lumpy investment decisions — for example, whether or not to enter a market. Section 3 addresses capital expansion models wherein it is recognized that firms may continuously adjust their capital stocks (e.g., capacity). Section 4 deals with more complex investment problems in which several investment decisions are intertwined. Section 5 summarizes the results, emphasizes key managerial insights, and offers suggestions for future research.

2. Lumpy investment decisions: Entry, improvement, and exit

2.1. Exogenous competition and random entry. An early approach modeled competition exogenously and helped identify certain major drivers. Trigeorgis (1991) studies the impact of competition on investment timing using standard ROA techniques for the geometric Brownian motion (GBM),

$$(2.1) \quad dX_t = \mu X_t dt + \sigma X_t dB_t,$$

for drift parameter μ , volatility σ , and $B = \{B_t\}_{t \geq 0}$ a standard Brownian motion.³ “Competitive arrivals” may reduce the value of a firm’s own investment opportunity by taking away significant market share. Such competitive entries are modeled as either anticipated or random Poisson arrivals. The former case is analogous to an increased constant opportunity cost of waiting, whereas the latter reflects the risk of a sudden drop in profitability; both interpretations suggest earlier investment. This framework is limited in that it does not explain what drives competitors’ entry decisions. An endogenous, game-theoretic approach explaining the incentives of firms to enter is more appropriate.

2.2. Discrete-time analysis of new-market models. A common practice in operational research and industrial organization is to assume that decision makers decide at discrete moments in time. Applied to option games, this approach provides an intuitive introduction to some key insights derived in more complex settings.

Spencer and Brander (1992) consider a duopoly with quantity competition and characterize in closed form the trade-off between flexibility and commitment. Demand is linear and subject to an additive stochastic shock X_t (demand intercept):

$$(2.2) \quad P(X_t, Q) = X_t - Q,$$

where $Q = q_i + q_j$ is the total industry output and where the random variable X_t has mean a (the standard demand intercept parameter) and variance σ^2 . Firms’ cost functions are linear, with variable production cost $c = c_i = c_j$. In a *Stackelberg* setting, the leader – say, firm i – may precommit to a choice of output prior to uncertainty resolution or may instead play a Cournot game after X_t is realized. In equilibrium, firm i precommits if the demand variance is sufficiently low – namely, if $\sigma^2 < (a - c)^2/8$. With higher variance, the flexibility to adjust output *after* uncertainty is resolved is more valuable than committing to secure a first-mover advantage. The authors then allow both firms to decide whether they will commit. Three game structures may result from the first-stage commitment game: (i) only one firm commits (*Stackelberg* game), (ii) both commit (*committed Cournot* game), or (iii) neither does (*flexible*

²Grenadier (2000b) and Huisman (2001) introduce the subject in continuous time, whereas Smit and Trigeorgis (2004) mainly deal with discrete-time models. Chevalier-Roignant and Trigeorgis (2011) discuss both modeling approaches. Boyer, Gravel, and Lasserre (2004) review the literature but do not stress commonalities of the models. Huisman, Kort, Pawlina, and Thijssen (2004) are restrictive in their scope, focusing on lumpy problems in continuous time.

³Throughout, we follow the risk-neutral valuation approach under the assumption of a complete market that precludes arbitrage opportunities (see Cox and Ross, 1976; Harrison and Kreps, 1979).

Cournot game). In the symmetric cost case with pure strategies, either both firms commit or neither does – that is, we have equilibrium (ii) or (iii).⁴ From a firm welfare perspective, the committed Cournot game is detrimental because both firms are affected by demand uncertainty yet have no opportunity to adjust their output. Even so, this outcome may obtain in equilibrium for sufficiently low variance, $\sigma^2 < (a - c)^2/8$: both firms strive to secure the lucrative leader role, which results in the inefficient equilibrium (ii). For $\sigma^2 > 7(a - c)^2/36$, the advantage from output flexibility dominates the commitment advantage and so (iii) obtains in equilibrium. In the intermediate region, the two Nash symmetric equilibria (ii) and (iii) coexist.

Smit and Ankum (1993) use a binomial lattice to analyze a duopoly investment game. Their results corroborate some (analytical) findings of Spencer and Brander (1992). Inefficient mutual preemption may arise when firms try to become the leader, yet firm asymmetry can result in a natural ordering or sequencing of firms' investment. Although the numerical approach provides only limited economic intuition, it does allow one to investigate richer model settings. Zhu and Weyant (2003b,a) use a similar approach to investigate the effect of information asymmetry on investment games for the case in which one of the two firms ignores its rival's production cost.⁵ For simultaneous commitment decisions, the authors derive results in line with those of Spencer and Brander (1992). If investment decisions are made sequentially, then information may be revealed depending on the order of investments. If the less-informed firm moves first, no new information is revealed to its rival. However, if the better-informed firm invests first, its rival learns the production cost of the leader. This gives rise to strategic learning that may dampen the preemption effect.

2.3. Continuous-time analysis of new-market models. Whereas the discrete-time modeling approach is readily implementable, continuous-time models generally provide more clear-cut economic interpretations. Next we discuss such models under the premise that no firm is active at the outset.

2.3.1. Complete-information case. Dixit and Pindyck (1994) propose a continuous-time symmetric duopoly model that simplifies the original approach of Smets (1991). An investment opportunity is available to firms i and j at a fixed (sunk) investment cost I . The stochastic profit function consists of a deterministic reduced-form profit — denoted π_L in monopoly or π_C in duopoly, $\pi_L > \pi_C$ — and a multiplicative shock X that follows a GBM, as in Equation (2.1), starting at level X_0 . The common discount factor is r . Since firm payoffs depend on the firms' entry decisions, the situation is analogous to a multiplayer optimal stopping problem. For sequential investments, the *follower* is faced with a single-agent optimal investment timing problem. The follower's (gross) value if it enters at level x is $V_F(x) \equiv x\pi_C/\delta$ with $\delta \equiv r - \mu$. Following the standard real options approach (see McDonald and Siegel, 1986), the follower's entry threshold X_F is obtained such that

$$(2.3) \quad \frac{V_F(X_F)}{I} = \Pi^* \quad \text{or} \quad \frac{\pi_C X_F}{I} = \Pi^* (r - \mu),$$

where

$$(2.4) \quad \Pi^* \equiv \frac{\beta_1}{\beta_1 - 1} \quad (> 1)$$

and

$$(2.5) \quad \beta_1 \equiv \frac{-(\mu - \frac{1}{2}\sigma^2) + \sqrt{(\mu - \frac{1}{2}\sigma^2)^2 + 2r\sigma^2}}{\sigma^2} \quad (> 1).$$

The term Π^* can be interpreted as a required profitability level that must be exceeded by the return on the investment project, $V_F(X_F)/I$, in order for the firm to invest. The follower's value $F(\cdot)$ as a function

⁴However, if the firms' marginal costs are *asymmetric* or privately known (*incomplete information*) then an asymmetric Bayesian Nash equilibrium, in which the more cost-efficient firm invests first, may arise endogenously. This helps illustrate how a static competitive (e.g., cost) advantage can create a dynamic strategic advantage.

⁵The analysis in Zhu and Weyant (2003a) focuses on a single decision point, whereas that in Zhu and Weyant (2003b) consider multiple time steps in an approach similar to the one used in Smit and Ankum (1993).

of the initial demand level X_0 is

$$(2.6) \quad F(X_0) = \begin{cases} V_F(X_0) - I & \text{if } X_0 \geq X_F \text{ [invest]}, \\ \mathbb{E}_0[e^{-k\tau_F}] [V_F(X_F) - I] & \text{if } X_0 < X_F \text{ [wait]}. \end{cases}$$

Here $\mathbb{E}_0[e^{-k\tau_F}] = (X_0/X_F)^{\beta_1}$ is an expected discount factor and $\tau_F = \inf\{t \in \mathbb{R}_+ \mid X_t \geq X_F\}$ is a random variable indicating the first time the barrier X_F is reached from below. A firm entering the market early on, a *leader*, earns monopoly rents $X_t\pi_L$ as long as the demand shock remains below X_F . After the follower's entry, the firms will form a duopoly, each earning $X_t\pi_C$. Hence, the value of being the leader can be expressed as

$$(2.7) \quad L(X_0) = \begin{cases} V_F(X_0) - I & \text{if } X_0 \geq X_F, \\ V_L(X_0) - I + \mathbb{E}_0[e^{-k\tau_F}] [V_F(X_F) - V_L(X_F)] & \text{if } X_0 < X_F; \end{cases}$$

here $V_L(x) \equiv x\pi_L/\delta$. Upon the follower's entry at random time τ_F , the leader "exchanges" its monopoly perpetuity value $V_L(X_F)$ for the duopoly perpetuity value $V_F(X_F)$. Figure 1 plots the values of leader and follower as a function of the state regions considered. At X_P , firms are indifferent between the leader and the follower role. For low initial values $X_0 < X_P$, the leader value is lower than the follower value; no entry will occur in this demand region. In (X_P, X_F) , there is an incentive to invest as leader because the leader's value exceeds the follower's. As a result, each firm plans to invest an ϵ -increment

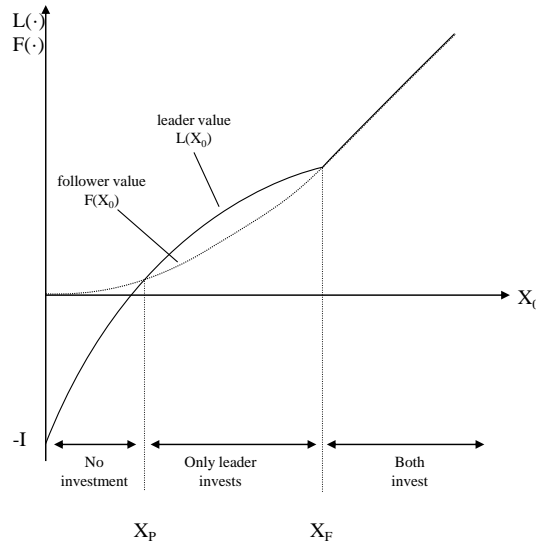


FIGURE 1. Values of leader and follower in duopoly (Dixit and Pindyck, 1994)

earlier than its rival. The firms thereby dissipate the first-mover advantage $L(\cdot) > F(\cdot)$. Preemptive behaviors are no longer profitable when firms are indifferent between the leader and the follower roles at X_P .⁶ This phenomenon corresponds to the rent equalization principle explained by Fudenberg and Tirole (1985) in a deterministic setting. In the intermediate region, (at least) one firm invests;⁷ for large values $X_0 > X_F$, both firms operate in the market. Here option premia are positive for both firms, but the threat of preemption will cause the leader to invest earlier than in the monopoly case.⁸

⁶The subscript P in X_P stands for "preemption."

⁷As shown by Huisman and Kort (1999), an undesired simultaneous investment outcome ("coordination failure" or "mistake") may result from strategic interaction with positive probability if $X_0 \in (X_P, X_F)$.

⁸Joaquin and Butler (2000) and Dias and Teixeira (2010) discuss new-market models with asymmetry in terms of production costs. Joaquin and Butler take an open-loop, pure-strategy approach; Dias and Teixeira use a closed-loop, mixed-strategy approach and prove that Joaquin and Butler's open-loop equilibrium obtains for a large cost differential. Similarly, Takashima, Goto, Kimura, and Madarame (2008) use a new-market model to analyze technology choices for

Bouis, Huisman, and Kort (2009) extend the duopoly model just described by considering a larger number of symmetric firms that have the option to (irreversibly) enter a new market. A multiplicative shock X that follows the GBM of Equation (2.1) affects the profit streams, where the deterministic profit components π_n are decreasing in the number of incumbents n . To derive their main insights, the authors focus on the three-firm case and provide numerical analysis for larger oligopolies. Investment thresholds in case of sequential equilibria are derived in an analogous fashion. The investment trigger of the first and the second entrant, X_1 and X_2 , are determined by rent equalization. If the triopoly profit is reduced, the third entrant will invest later and set a larger investment trigger X_3 . The second investor then has an incentive to invest earlier because it can enjoy duopoly rents longer, thus leading to a decrease in X_2 by rent equalization. The first investor then faces earlier entry by the follower and enjoys monopoly profits for a shorter period; hence its entry threshold X_1 is increased. The opposite direction of the change between X_2 and X_3 on the one hand, and between X_1 and X_2 on the other hand, is referred to as the *accordion effect*. Note also that the leader in the three-firm case invests after the leader in duopoly yet before the monopolist; therefore, increased competition (three firms rather than two) actually delays rather than hastens investment. For larger oligopolies, Bouis, Huisman, and Kort argue that the number of expected future entrants is critical (especially whether this number is odd or even) and illustrate how the accordion effect sustains.

Weeds (2002) analyzes a patent race between duopolists and the effect of competitive pressure on firms' research activities. Two symmetric firms have the opportunity to launch a research and development (R&D) project at a cost I . The first firm to succeed gains an exclusive patent while the other firm is left with nothing. Firms face two sources of uncertainty: the patent value evolves as the GBM of Equation (2.1), and the research outcome is random with mean Poisson arrival rate λ .⁹ A firm's R&D strategy implies a trigger X_F at which research activity is initiated. The follower's investment threshold X_F is such that

$$(2.8) \quad \frac{X_F}{I} = \Pi^* \frac{r + 2\lambda - \mu}{\lambda}.$$

Equation (2.8) resembles Equation (2.3) but with an adjustment for the Poisson arrival (and with profit flow normalized to unity). The leader's threshold X_P is obtained from rent equalization. Weeds also derives the optimal behavior imposed by a social planner in two related, cooperative settings: (A) two identical research units with launch cost I and success arrival rate λ ; and (B) one common large research institute with doubled launch cost $2I$ and success arrival rate 2λ . In case A, the social planner would optimally choose to phase research: one firm initiates research when the patent value reaches a threshold $X_L \in (X_P, X_F)$ that maximizes the leader's value; the other firm begins conducting research later, when the threshold $X'_F = X_F \cdot (r + \lambda - \mu)/(r - \mu) > X_F$ is first attained. Because the follower participates in the potential research success of the leader, it waits for a patent profitability that is larger than in the competitive case. In case B, the large research institute commences research activities — at threshold $X_C \in (X_F, X'_F)$ — later than would a follower in a competitive setting.

Mason and Weeds (2010) assume that the follower's investment can create both positive and negative payoff externalities, where the type of the payoff externality affects the equilibrium outcome. The authors derive conditions for the existence of a simultaneous investment equilibrium in which both firms invest simultaneously at a later threshold X_S than the preemptive threshold X_P . In the presence of negative externalities, the threshold X_P is increasing in volatility σ but is bounded from above while X_M (the monopolist threshold), X_F , and X_S are increasing without bound in σ . Under positive externalities,

power generation investments. Kong and Kwok (2007) analyze a new-market model where duopolist firms face asymmetric investment costs and revenues; the authors investigate the effects of these asymmetries as well as the presence of negative versus positive externalities on the type of investment schedules: open-loop, preemptive, and simultaneous. We discuss similar effects for investment in an already existing market in Section 2.4.

⁹Weeds (2002) assumes that the project value at time 0 is below I ; therefore, simultaneous investment at the outset is precluded as an equilibrium.

however, the leader investment threshold can actually decrease with higher volatility. This latter result stands in sharp contrast to the standard ROA story.

2.3.2. Incomplete-information case. Lambrecht and Perraudin (2003) consider the effect of incomplete information on investment policies. Consider a duopoly where firm j selects an investment trigger X_j .¹⁰ Here, firm i conjectures a distribution function $F_j(\cdot)$, with density $f_j(\cdot)$, for its rival's trigger X_j . Firm i can update its beliefs in view of whether its rival invests at new, higher levels $\hat{X}_t \equiv \{X_s \mid 0 \leq s \leq t\}$ with conditional distribution function $F_j(\cdot \mid \hat{X}_t)$. The market is incontestable once a leader has entered. Upon entry at level x , the leader's gross value is $V_L(x) = x/\delta$. Given the risk of preemption, firm i 's optimal investment threshold X_P^i is such that

$$(2.9) \quad \frac{V_L(X_P^i)}{I_i} = \Pi^{*'} \quad \text{or} \quad \frac{X_P^i}{I_i} = \Pi^{*'}(r - \mu),$$

where the profitability index of (2.4), applying for myopic firms, is adjusted for the hazard rate $h_j(x) = xf_j(x) / [1 - F_j(x)]$ (≥ 0):

$$\Pi^{*'} \equiv \frac{\beta_1 + h_j(X_P^i)}{\beta_1 - 1 + h_j(X_P^i)} \in (1, \Pi^*].$$

Owing to information asymmetry, the threshold X_P^i is located between the full preemption threshold (zero net present value threshold with associated profitability index 1) and the myopic threshold X_L^i (with profitability index Π^*). The presence of incomplete information thus preserves some option value of waiting ($\Pi^{*'} > 1$). In particular, assume that no firm knows the exact realization of its rival's investment cost but that each firm has a prior belief $G(\cdot)$. Since there exists a mapping from firm j 's investment cost I_j to its investment threshold $X_j(I_j)$, it follows that, in a perfect Bayesian equilibrium, an exercise strategy involves a map from the distribution $G(\cdot)$ of I_j to the distribution $F_j(\cdot)$ of the rival's entry trigger. For a special case of Pareto distribution for $G(\cdot)$, a closed-form solution obtains for the mapping from I to the optimal trigger X_P^i . Hsu and Lambrecht (2007) build upon Lambrecht and Perraudin's (2003) model and examine a patent race where a potential entrant has complete information about the incumbent but the latter does not know the realization of the challenger's entry cost.

Nishihara and Fukushima (2008) consider a duopoly where profits are subject to a multiplicative shock that follows the GBM described in Equation (2.1). The firms face asymmetric entry costs I_i and I_j : firm i does not know its rival's cost realization I_j but does know the cost's distribution. By assumption, firm i is necessarily the leader (firm j cannot profitably operate as a leader). Firm j may nevertheless enter at a later point, earning $\pi_C^j > 0$ and driving firm i out of the market. The follower threshold X_F solves a single-player optimization problem. Under perfect information, the leader invests in the demand region $(\underline{X}_L, \bar{X}_L)$. Because firm i cannot be preempted, the lower investment trigger \underline{X}_L is selected irrespective of the follower's actions. The value of the upper bound \bar{X}_L reflects the possibility that firm i might be driven out of the market when firm j invests. Given incomplete information, firm i forms a belief for the upper threshold \tilde{X}_L . If $\tilde{X}_L > \bar{X}_L$, then firm i will invest at high demand levels and will promptly be driven out of the market by its rival. But if $\tilde{X}_L < \bar{X}_L$, then firm i will forgo a profitable investment opportunity in the region (\tilde{X}_L, \bar{X}_L) . Firm i 's decision not to invest for low demand remains unaffected by the presence of information asymmetry.

Hoppe (2000) and Thijssen, Huisman, and Kort (2006) consider a timing game in which the profitability of the project is uncertain. Both firms have an initial belief $p_0 \in [0, 1]$ that the project is good. After a firm has invested, the project profitability becomes immediately known to both firms: these information spillovers may create a second-mover advantage. Hoppe (2000) considers a duopoly in which the belief is exogenously given, fixed, and common knowledge. Preemption (and rent equalization) takes place if the probability of success is high. The timing game switches from a preemption game to a war of attrition as this probability decreases. In Thijssen et al. (2006), Poisson signals convey information about the project's profitability: a "good" signal indicates a profitable project whereas a "bad" signal

¹⁰The analysis can also be extended to n firms.

indicates a project that will generate a loss. Firms update their belief p_t in view of new Poisson arrivals and the rival's investment decisions. If the quality and quantity of the signals is sufficiently high, the first-mover advantage dominates: the industry equilibrium exhibits preemption. Yet if information spillovers overwhelm the first-mover advantage then, at least initially, a war of attrition prevails. In this case, more competition does not necessarily lead to greater social welfare. Décamps and Mariotti (2004) discuss a similar problem for firms with asymmetric information about their rivals' investment costs.

In Grenadier (1999), each firm receives an independent, private signal about a parameter that drives the project value. A private signal is informative but imperfect, so firms condition their investment strategies not only on their private signals but also on information conveyed by rivals' observed investment decisions. Firms are assumed to know the distribution of their rivals' signals but not their realizations. Firms with signals that are more informative will invest earlier, often in a suboptimal manner. Situations where firms invest immediately — because their private information is overwhelmed by the information conveyed by others — may arise endogenously. Such situations are called “information cascades.”

2.4. Improvement. Huisman and Kort (1999) provide an advanced analysis of cases in which firms are already operating in a given market yet contemplate investing I in order to improve their current profit flows (a so-called existing-market model). Firm i 's stochastic profit consists of two components: a *multiplicative* exogenous shock $X = \{X_t\}_{t \geq 0}$ (modeled as the GBM of Equation 2.1) and a nonnegative, deterministic, reduced-form profit under some industry structure. At the outset, firms earn profit π_0 . Upon making a new investment, a firm may earn a higher incumbency profit $\pi_L > \pi_0$. Firms' investments typically impose a negative externality on their rivals' profit (by a known discrete amount). Once the leader has invested, the follower experiences a lower profit, $\pi_F < \pi_0$, as long as it has not itself invested. Once both firms have invested, they make duopoly profits with $\pi_C \in (\pi_0, \pi_L)$. Note that the value increment from leadership, $\pi_L - \pi_0$, is greater than the value increment received by the follower upon investing, $\pi_C - \pi_F$. These differences drive the somewhat different results for existing-market models as compared with the new-market models analyzed in Section 2.3.1. Firms' value functions account for the different industry structures that the firms experience. Huisman and Kort identify two distinct equilibrium scenarios for a *symmetric duopoly*. The first one, which we have already encountered, involves a preemptive investment sequence wherein the leader's investment threshold X_P is determined by rent equalization: $L(X_P) = F(X_P)$. In a second equilibrium scenario, firms implicitly agree to invest at a later time that maximizes joint value as part of *tacit collusion*.¹¹ They pursue this collaborative approach if they fear the loss of existing rents. If cash flows are highly volatile then firms are reluctant to invest early, which makes tacit collusion likely to prevail. The probability of each industry structure is endogenously determined. Huisman and Kort (1999) and Thijssen, Huisman, and Kort (2002) extend Fudenberg and Tirole's (1985) strategy and equilibrium definitions to allow for stochastic changes in the environment. Pawlina and Kort (2003) apply such techniques to study a situation where duopolist firms compete in quantities and may replace their existing technology for a more cost-effective one. Demand is linear (per Equation 2.2) with an *additive* shock that follows the GBM of Equation (2.1). As before, investments may take place either in sequence (preemption) or simultaneously (tacit collusion) depending on demand characteristics. Although the profit functions are convex in the uncertain demand parameter, increased volatility always favors a delayed technology replacement regardless of equilibrium type. The underlying assumption of GBM enables the authors to study the expected time until replacement and to analyze the impact of uncertainty.

Grenadier (1996) develops a duopoly model involving *completion delays* that provides insights into the behavior of property developers. Two symmetric real estate developers have opportunity to increase

¹¹This equilibrium was already identified by Fudenberg and Tirole (1985): if the payoffs from late and joint investment are sufficiently high then there exists a continuum of tacit collusion equilibria, one of which coincides with the cooperative (Pareto) optimum.

their profit stream by refurbishing their building for an investment outlay of I .¹² During the D years until completion, the owner receive no rents from the building. The follower's threshold X_F , which is obtained by standard techniques, satisfies

$$(2.10) \quad \frac{V_C(X_F)}{I'} = \Pi^* e^{-\delta D} \quad \text{or} \quad \frac{\pi_C X_F}{I'} = \Pi^* (r - \mu);$$

here Π^* is as given in (2.4) and I' is the total opportunity cost of investing — that is, investment cost I plus the value π_F/δ of the forgone perpetual income stream from the old building. For low values of the process (below X_P), the follower's value strictly exceeds the leader's and no improvement takes place. For intermediate values $[X_P, X_F)$, only a single firm decides to invest; in this case, each firm is equally likely to become the leader.¹³ For large process values $[X_F, \infty)$, both firms invest. The tacit collusion scenario may occur for certain parameter values of the underlying process. The real estate market is often characterized by sudden large development efforts (investment cascades) or building booms in the face of declining demand and property values (recession-induced construction booms). Grenadier determines factors that explain such anomalies: the occurrence of investment cascades is driven by market volatility, and completion delay is the major reason behind recession-induced construction booms.

Pawlina and Kort (2006) analyze the effects of firm *heterogeneity* on optimal timing (asymmetric investment costs). The advantaged firm pays I upon entry but its rival faces a higher investment cost αI ($\alpha \geq 1$). As shown in Figure 2, three types of equilibria may arise depending on the magnitude of the first-mover advantage π_L/π_C , and the level α of firm asymmetry. Preemptive equilibria, which are characterized by rent equalization, emerge for moderate asymmetry. Cooperative equilibria, which involve simultaneous investments at a socially optimal later time, obtain for sufficiently homogeneous firms that have no real possibility of gaining a first-mover advantage. *Sequential (open-loop) equilibria*,

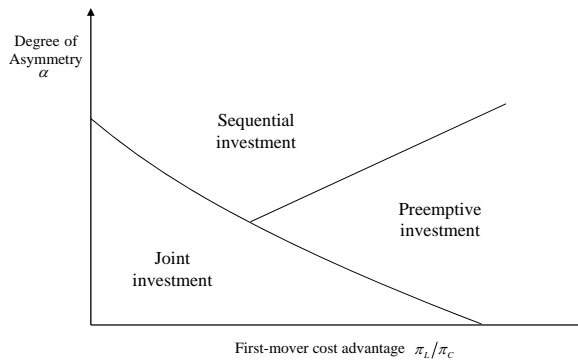


FIGURE 2. Equilibrium regions (Pawlina and Kort, 2006)

in which the advantaged firm does not fear preemption and invests at the monopolist threshold, obtain for high levels of asymmetry. The latter type of equilibrium is specific to the asymmetric case. The relationship between firm values and the degree of asymmetry among firms is not clear-cut: a higher degree of homogeneity may give rise to less efficient industry equilibria for both firms owing to preemption effects.

¹²Grenadier (1996) departs from standard existing-market assumptions by assuming that profits π_0 and π_F are not subject to the multiplicative shock X .

¹³Like Dixit and Pindyck (1994), Grenadier (1996) assumes that the leader is selected by the flip of a fair coin. Huisman and Kort (1999) point out that this result holds only if $X_0 < X_P$; if $X_0 \in (X_P, X_F]$, then there is a positive probability that a duopoly emerges immediately. This outcome is not desirable from the viewpoint of the two firms and constitutes a “coordination failure” or “mistake.”

2.5. Repeated lumpy capacity expansions. In the cases considered so far, each firm had a single investment opportunity. Boyer, Lasserre, Mariotti, and Moreaux (2004) and Boyer, Lasserre, and Moreaux (2010) consider, in continuous time, a duopoly where *symmetric* firms may sequentially add capacities in lump sums. Firms initially have low capacities, and the only possible equilibrium at initial stages of the industry involves preemption. As a result of this preemption, the first industry-wide investment occurs earlier than the socially optimal time (from the viewpoint of industry participants). This distortion implies riskier entry and lower expected returns. Rent equalization occurs irrespective of the volatility or speed of market development. When firms already hold substantial capacity, tacit collusion may be sustainable as an industry equilibrium because firms fear the loss of existing rents; such equilibria are more likely in highly volatile or fast-growing markets.¹⁴ The possibility of collusion is more attractive to symmetric firms than to asymmetric ones. Boyer et al. (2004) assume that reduced-form profits are the outcome of Bertrand price competition, whereas Boyer et al. (2010) consider Cournot quantity competition.

Murto, Näsäkkälä, and Keppo (2004) consider, in discrete time, the case where firms can expand capacity by a lumpy amount ΔQ at a cost I . The authors use stochastic dynamic programming and Monte Carlo simulation to derive a numerical characterization of the equilibrium outcome. Demand is isoelastic,

$$(2.11) \quad \pi(X_t, Q_t) = X_t \cdot Q_t^{-1/\eta},$$

with constant elasticity $\eta > 1$. The investment trigger $X(q_t^i, Q_t)$ is increasing in firm i 's total installed capacity q_t^i . Therefore, a smaller firm is more likely than a larger one to respond to small demand shocks by expanding capacity. The authors analyze a form of *firm asymmetry* whereby one firm invests in smaller lumps but for a larger unit cash outlay. Despite its higher investment cost per capacity unit, this firm may be better off because it can react more quickly to changes in demand and is less vulnerable to decreases in demand.

Novy-Marx (2007) allows for (a continuum of) *asymmetric* firms that differ in their initial capacities, where the logarithms of the capacities are uniformly distributed between the largest and the smallest firm. The assets produce a nonstorable output stream Q_t that is sold at the market-clearing price $P_t \equiv P(X_t, Q_t)$, and each firm realizes a revenue $\pi(X_t; q_t^i, Q_t) = q_t^i P_t$. Inverse demand is isoelastic (Equation 2.11), and the shock follows a GBM (as in Equation 2.1). A firm with capacity q can redevelop at a cost of $I(q) = q^\gamma$ ($\gamma > 1$), although doing so involves an increasing cost-to-scale ratio. Redevelopment requires abandoning the profit stream from the current assets in place. In equilibrium, a firm with capacity q^i expands to κq^i when the market-clearing price reaches a certain (upper) threshold $\bar{P}(q^i, Q)$. Both the intensity and the timing of redevelopment depend on the parameter κ (> 1). Remarkably, κ is industry specific and stationary, which ensures that the distribution of capacities is preserved over time. At any point, the next firm to expand capacity will be the smallest firm. Simultaneous investment does not occur, and firm heterogeneity results in a natural ordering of firm investments.

Carlson, Dockner, Fisher, and Giammarino (2010) study the effect of rivals' expansion and contraction options on incumbent firms' risk exposure. Heterogenous duopolists face isoelastic demand of Equation (2.11) and the shock follows the GBM of Equation (2.1). The firms may expand or contract capacity (by lump sums) in response to market realizations. A rival's investment decisions act as a natural hedge against variation in the exogenous state variable: growing demand may induce a rival's capacity expansion and thus limit the firm's upside potential; a bearish market increases the likelihood of a rival's contraction in capacity, reducing the firm's downside risk. These effects must be considered when using peer betas to proxy a firm's risk.

¹⁴In such a context, the conventional real options result — that high volatility leads to postponing investment — is reinforced because higher volatility may result in a switch from the preemption equilibrium to a tacit collusion equilibrium involving later investment and higher values.

2.6. Exit. Fine and Li (1989) supplement deterministic models of exit (e.g., Ghemawat and Nalebuff, 1985; Fudenberg and Tirole, 1986; Ghemawat and Nalebuff, 1990) by allowing for the stochastic decline of a duopoly market. The authors show (in discrete time) that, owing to “jumps” in the demand process, the sequence of exit is not unique.

Sparla (2004) analyzes, in continuous time, the divestment or exit of duopolists when firms face a stochastically declining market. At the outset, firms compete in the marketplace and earn π_0 . As soon as one firm (the leader) divests, its profits drop to $\pi_L < \pi_0$ and the profit of its rival (the follower) rises to $\pi_F > \pi_0$. Once both firms have divested, each earns $\pi_C \in [\pi_L, \pi_0)$. At the time of divestment, the drop in gross profit for the follower is lower than that for the leader, $\pi_F - \pi_C < \pi_0 - \pi_L$; this creates a second-mover advantage. For exit situations, we have $\pi_L = \pi_C = 0$. The multiplicative shock (a GBM in Equation 2.1) evolves at such a low rate $\mu < \frac{1}{2}\sigma^2$ that it eventually becomes optimal for the firms to exit. When a firm exits, it receives a salvage value $S > 0$.¹⁵ If a firm has already exited the market, then its rival, the “follower,” becomes a monopolist with (gross) value $V_F(x) = x\pi_F/\delta$. The follower’s exit problem is of a decision-theoretic nature: its divestment threshold X_F satisfies

$$(2.12) \quad \frac{V_F(X_F)}{S} = \Pi_* \quad \text{or} \quad \frac{\pi_F X_F}{S} = \Pi^*(r - \mu),$$

where Π_* is the profitability level that must be attained. It is given by

$$(2.13) \quad \Pi_* \equiv \frac{\beta_2}{\beta_2 - 1},$$

where

$$\beta_2 \equiv \frac{-(\mu - \frac{1}{2}\sigma^2) - \sqrt{(\mu - \frac{1}{2}\sigma^2)^2 + 2r\sigma^2}}{\sigma^2} \quad (< 0).$$

In this *war of attrition* or *chicken game*, both firms have an incentive to wait until the rival exits and subsequently operate as a monopolist, or until market conditions deteriorate to such an extent that both firms are better-off leaving the market, regardless of their rival’s action. In the unique symmetric perfect equilibrium in mixed strategies, both firms exit simultaneously the first time the threshold X_F in Equation (2.12) is reached from above. When firms have *asymmetric* production costs, a second equilibrium may also obtain: the high-cost firm exits first and the low-cost firm later. Under moderate cost differences, both types of equilibria coexist. If the degree of heterogeneity is sufficiently high, the “diffusion” equilibrium prevails. Murto (2004) considers asymmetry in terms of production scale. The exit threshold of the leader is X_L , such that

$$(2.14) \quad \frac{V_L(X_L)}{S} = \Pi_* \quad \text{or} \quad \frac{\pi_0 X_L}{S} = \Pi^*(r - \mu);$$

here $V_L(x) = x\pi_0/\delta$, and the follower’s threshold is again given by Equation (2.12). For a given degree of firm heterogeneity, the equilibrium — whereby the smaller firm exits first at X_L and the largest firm (j) later at $X_F < X_L$ — is unique if the volatility is sufficiently low, but under increased uncertainty it will coexist with an equilibrium that has the reverse ordering. The first exit sequence will also prevail when the largest firm is sufficiently larger (for a given volatility level).¹⁶

¹⁵For enhanced comparability, we simplify Sparla’s (2004) model by assuming exit rather than partial closure and by using a single aggregated salvage value S . Both Sparla (2004) and Murto (2004) decompose the salvage value into cost savings made upon exiting the market and costs incurred to effect the exit (e.g., layoff costs).

¹⁶Unfortunately, these models cannot be readily compared. Sparla (2004) formulates the problem with mixed strategies, whereas Murto (2004) uses pure (Markov) strategies and stopping sets. In the symmetric case, Murto (2004) obtains two equilibria involving a sequence of exits (i.e., the firms are interchanged with respect to their position as leader and follower).

3. Incremental capacity expansion

Another stream of research focuses on situations characterized by firms that can increase their capacity (or, more generally, their capital stock) by infinitesimal increments. As Pindyck (1988) suggests, this approach is extreme (since most investments are lumpy) but does offer the convenience of analytical elegance.¹⁷ Section 3.1 discusses oligopoly models of incremental capacity expansion, and Section 3.2 elaborates on the social optimality of “myopic” investment strategies under perfect competition.

3.1. Oligopoly. Grenadier (2002) describes a *symmetric* oligopoly with n firms. Profit streams are affected by a fairly general aggregate shock modeled as a diffusion process:

$$(3.1) \quad dX_t = \mu(X_t) dt + \sigma(X_t) dB_t,$$

where $\mu(X_t)$ and $\sigma(X_t)$ are, respectively, the drift and diffusion terms and where $B = \{B_t\}_{t \geq 0}$ is a standard Brownian motion. Each firm possesses q_t^i units of capacity at time t and can increase capacity incrementally at any time at a cost of I per capacity unit. Building new capacity requires D years until completion. Given constant returns to scale and negligible production cost, firms produce at full capacity and sell at the market-clearing price $P_t \equiv P(X_t; Q_t)$, where $Q_t = \sum_i q_t^i$ is the total industry output. If we assume symmetric investment strategies, then the firms expand simultaneously and in the same proportion; that is, $dq_t^i = dq_t = dQ/n$. The equilibrium implies an upper trigger for the price level $\bar{P}(Q) \equiv P(\bar{X}(Q), Q)$ at which the firms will increase their aggregate capacity by dQ . For $D > 0$, the firms will trigger investment at time t based not on the current price level P_t but rather on their expectation concerning the price at time $t + D$. Grenadier (2002) formulates the problem in a tractable way. If the shock follows the GBM of Equation (2.1), then the investment trigger $\bar{X}(Q)$ satisfies

$$(3.2) \quad \frac{P(\bar{X}(Q), Q) + \frac{Q}{n} P_Q(\bar{X}(Q), Q)}{I} = \Pi^*(r - \mu);$$

here the numerator on the left-hand side is the firm’s *marginal* profit at the investment trigger. For isoelastic demand in Equation (2.11), this marginal profit is given by $\bar{X}(Q) [(n\eta - 1)/n\eta] Q^{-1/\eta}$. Closed-form solutions also exist for several combinations of the arithmetic or geometric Brownian motion and square-root process with the linear or the isoelasticity demand function. As the market approaches perfect competition, option values become less valuable.

Aguerrevere (2003) extends Grenadier’s (2002) model by considering capacity utilization u as a strategic variable. Demand is linear and the production cost function is quadratic in u . Aggregate industry capacity is increasing in the number of firms but remains finite under perfect competition ($n \rightarrow \infty$). The level of capacity utilization is independent of the number of firms. The output price paths exhibit mean reversion due to discretion in the capacity utilization rate, with significant spikes during times of extremely high capacity utilization. Owing to time-to-build delays, completion of capacity expansion is preceded by a phase of high utilization and high prices. Moreover, the investment lag also leads to an ambiguous effect of increased volatility σ on the optimal level of capacity held and still under construction (“committed capacity”). If construction is instantaneous, then committed capacity is strictly decreasing in σ : firms can be surprised only by low levels of demand yet can always respond to high levels. Consequently, lower capacity levels are optimal. In contrast, if time-to-build is strictly positive then firms face the risk of having too little installed capacity on hand in buoyant markets. This is most likely in situations with high demand and high σ and firms have an incentive to hold reserve capacity.¹⁸ Hence, committed capacity is decreasing in σ for sufficiently low demand (underutilization is more likely) but increasing for high demand (firms anticipate capacity shortage and expand).

¹⁷Such models rely on *instantaneous control* techniques, whereas the models discussed previously deal with *impulse controls*.

¹⁸The utilization discretion is key for this argument. In a setting of constant utilization, firms will not hold reserve capacity.

Aguerrevere (2009) uses Aguerrevere's (2003) model to examine a firm's systematic risk (beta), which is determined as a function of the riskiness of the assets in place and of its expansion options. Weights are determined based on the present values of assets in place and of expansion options. An expansion option beta obtains to be constant (independent of industry capacity, demand level, and the number of firms), whereas the beta of assets in place is increasing in the number of firms.¹⁹ As in Grenadier (2002), the value of expansion options decreases with the number of firms. Assets in place are generally less risky when demand is high, since then capacity utilization is increased. In particular, the beta of assets in place is zero when demand reaches the capacity expansion threshold. At this point, the firm's beta is solely determined by the weight and beta of expansion options. Because expansion options are more valuable in more concentrated markets, a firm's beta under conditions of high demand is decreasing in the number of firms. But with low demand (and correspondingly a low value of expansion options), the firm beta is mostly driven by the beta of assets in place; therefore, a firm's beta under low demand is increasing in the number of firms.

3.2. Myopia and social optimality. Building upon the rational expectations framework, Leahy (1993) sets the benchmark case for continuous-time analysis of (infinitely divisible) capital stock expansion and scrapping under perfect competition. A large market is composed of infinitesimal firms that decide on entry and exit. Firms face an uncertain price $P(X_t, Q_t)$ as a function of the capacity held in the industry Q_t and a demand shock $X = \{X_t\}_{t \geq 0}$ modeled as a diffusion. Even if the demand shock follows an (unrestricted) diffusion process, the entry and exit of infinitesimal firms implies that the price process $\{P_t\}_{t \geq 0}$, which is endogenously determined, is also a diffusion process, with reflecting lower and upper barriers $\underline{P}(Q)$ and $\bar{P}(Q)$. Each firm's entry and exit thresholds also coincide with these barriers. In equilibrium, the firms' Markov investment strategies exhibit some form of myopia in that firms invest at the same time as would a firm that ignores potential future capacity expansions. In comparison with the investment strategy formulated by a social planner and imposed on decentralized firms, myopic investment policies are socially optimal.²⁰ Grenadier (2000a) extends Leahy's (1993) model to allow for completion delays.

Baldursson and Karatzas (1996) consider capacity expansion and allow for stochastic processes that are not necessarily of the diffusion type. The correspondence between perfect competition, myopic strategies, and social optimality is again established based on a probabilistic approach to stochastic control theory that relies on the connection between singular control and optimal stopping theory. This approach helps identify the model's relation to lumpy investment problems.

Baldursson (1998) discusses both expansion and downsizing. The result that obtains in the Nash equilibrium is compared with the choice of a social planner through the use of a "fictitious" objective function. The author proposes a general model for the inverse demand function (price as a function of shock and capacity). Some special cases admit closed-form solutions.

Back and Paulsen (2009) challenge the assumption underlying myopic investment behaviors. The authors discuss the appropriateness of the Nash or open-loop equilibrium concept employed in most models of oligopoly and perfect competition (e.g., Baldursson and Karatzas, 1996; Baldursson, 1998; Grenadier, 2002). Open-loop strategies allow firms to respond to the resolution of uncertainty with respect to the exogenous shock but do not allow firms to respond to the observed actions by rivals. Optimal open-loop strategies must form a Nash equilibrium as part of the open-loop equilibrium. Back and Paulsen (2009) discuss that if firms could, in effect, respond to their rivals' actions (i.e., could formulate closed-loop strategies) then the equilibrium strategies derived by Grenadier (2002) would not

¹⁹Since capacity is increasing in the number of firms and since utilization is constant, it follows that aggregate output adapts more rapidly in a more competitive market. Thus, assets in place are riskier because their cash flows are more sensitive to demand shocks.

²⁰Leahy (1993) also points out that the impossibility of increasing one's capital by (infinitely) small amounts might explain the emergence of excess returns for firms behaving myopically (in lumpy investment models). Dixit (1991) discusses a perfect competition model along the lines of Leahy (1993).

be perfect.²¹ Formulating the dynamic capacity expansion problem in closed-loop strategies is difficult; nonetheless Back and Paulsen manage to show that, in the limit, the perfect competition outcome derived in Leahy (1993) is part of a (perfect) closed-loop equilibrium.

4. Complex investment decisions under uncertainty and competition

So far we have considered models that focus solely on the *timing* of investment, exit, expansion, or contraction. In reality, firms rarely limit their investment strategy to determination of the entry or expansion time. They often determine the staging of entry, the appropriate scale of production, the type and quality of product and so on. For simplicity, one frequently abstracts from market entry strategy in order to focus on the timing issue, which explains why most option games adopt this simple perspective. However, richer situations have also been explored. Section 4.1 discusses strategic growth options whereby the attractiveness of an early investment may depend on the opportunity it creates for the future. Section 4.2 explores the issues involved in compound option settings under competition, and Section 4.3 presents models in which firms decide on multiple criteria (multidimensional strategy sets).

4.1. Strategic growth options. Kulatilaka and Perotti (1998) assess the possibility of a firm making an early commitment investment that confers greater “capability” to take advantage of future growth opportunities. Such early investment opportunities are “strategic growth options.” Demand is linear (as in Equation 2.2) with X_t distributed on $(0, \infty)$. Firms select the quantity they supply. A firm’s early investment helps to reduce variable production cost from C to $c < C$ at a sunk investment cost I . For a given industry structure, the profit functions are convex in demand X_t because firms respond to improved market conditions by increasing output and prices. A strategic investment leads to greater convexity of ex post profits than in the case without early investment. In the original symmetric case, both firms invest as duopolists in the second stage if $X_t \geq C$. In an asymmetric *Stackelberg* game, the leader’s (ex post) entry threshold is c . The disadvantaged firm would earn $\pi_C(X_t; c, C) = (X_t - 2C + c)^2 / 9$ in an asymmetric Cournot duopoly; it will invest only if $X_t \geq 2C - c$. In other words, in the intermediate demand region, $c \leq X_t \leq 2C - c$, the leader ends up in a monopoly position.²² Each firm’s profit function is continuous and increasing in X_t . For simplicity, we normalize c to zero. The leader assesses the net expected gain $G_D(X_0)$ from commitment, which is given by

$$G_D(X_0) = \mathbb{E}_0 [\pi_C(X_t, C) \mathbb{1}_{\{X_t \geq C\}} - \{\pi_C(X_t, C) \mathbb{1}_{\{X_t \geq 2C\}} + \pi_L(X_t, 0) \mathbb{1}_{\{X_t < 2C\}}\}] - I.$$

A unique threshold level X_D makes the Stackelberg leader indifferent between committing or not, $G_D(X_D) = 0$. If demand exceeds X_D , then the strategic investment should be undertaken. The Stackelberg leader’s profit function is “more convex,” which favors overinvestment. The *gross* expected value of the strategic growth option thus increases with uncertainty. However, since the payoff function in $G_D(\cdot)$ is neither convex nor concave, it follows that the overall *net* effect of uncertainty is ambiguous. The incentive to precommit depends critically on the magnitude of the competitive (cost) advantage gained from committing. For large strategic effects, higher dispersion will tend to decrease X_D and thereby encourage commitment. For weak strategic effects, increased volatility deters precommitment. Kulatilaka and Perotti then consider the special case where X_t is log-normally distributed and obtain Black–Scholes-like closed-form solutions. If *simultaneous precommitment* is allowed, then a number of situations may emerge in the first stage: simultaneous commitment, simultaneous inaction, or precommitment by only one firm resulting in either monopoly or asymmetric Cournot duopoly. The firms may use mixed strategies; three regions for the mixed strategies are distinguished for which boundaries are obtained at indifference points. For low demand value, no firm precommits. For high demand, both firms invest to reduce their production cost. For intermediate demand, the strategies are fully mixed.

²¹A firm that pursues the equilibrium open-loop strategy, while its rivals may revise their strategy choice over time, would run the risk of being preempted.

²²The authors implicitly assume that asymmetric Cournot duopolists cannot mix their second-stage entry strategies.

Smit and Trigeorgis (2001, 2007, 2009) discuss similar problems by utilizing a framework adapted from Fudenberg and Tirole (1984). Using simulations, the authors assess different R&D strategies and infrastructure investment decisions in competitive environments. In contrast to Kulatilaka and Perotti (1998), Smit and Trigeorgis consider American-type options and consider situations where the strategic effect is positive because strategic actions are complements (as in price competition) and where the commitment is beneficial to the rival (“soft commitment”).

4.2. Staged investment appraisal and sleeping patents. Completion delays affect the (optimal) investment strategies of firms. For R&D projects (or in bio-tech, IT, and oil exploration), the length of such delays is rarely known in advance owing to *uncertain innovation success*. Weeds (2000) examines this source of uncertainty and its impact on the duopolist’s adoption of technology. The firm invests in research with the aim of acquiring a patent that will give it exclusive access to a new market. If the innovation is successful, then the firm has the option to make an additional sunk investment to adopt the new technology. The entire R&D investment opportunity can be assessed as a compound option where the value of the (first-stage) research option derives, in part, from the (second-stage) commercial investment option. The framework provides a rational explanation for the existence of sleeping patents — that is, patents acquired but kept in a standby or “sleep” mode. Policy makers typically regard sleeping patents as anticompetitive devices employed by dominant firms to erect entry barriers (blockaded entry). However, in the context described by Weeds, sleeping patents may arise when options coexist with completion uncertainty and there is no need for incumbent firms to coordinate their actions. By restricting a firm’s ability to let patents sleep, antitrust authorities may actually reduce — via compulsory licensing — option values and weaken firms’ incentives to conduct research in the first place.

Lambrecht (2000) derives optimal investment strategies for two symmetric firms sharing the option to make a two-stage sequential investment under *asymmetric information* on the rival’s profit. In the first stage, each firm is competing to acquire a patent that will enable it to proceed to the second stage (commercialization). Lambrecht derives conditions under which inventions are likely to be patented without being put to immediate commercial use. Sleeping patents are more likely to exist in an R&D portfolio when interest rates are low, when volatility is high, and/or when the second-stage cost is high relative to the first-stage cost.

Miltersen and Schwartz (2004) analyze patent-protected R&D investment projects under imperfect competition in the development and commercialization of the product. Competition in R&D may not only increase production and reduce prices but may also shorten the time needed to develop the product and increase the probability of a successful development. These benefits to society are offset by increased R&D investment costs in oligopolistic markets and lower aggregate value of all R&D investment projects.

4.3. Hybrid investments. In Section 2 we discussed lumpy investment situations where the main problem was the optimal timing of an investment outlay. Huisman and Kort (2009) allow firms to choose optimally their capacity and production scale when they enter the market. Demand is linear and subject to a multiplicative shock modeled as the GBM of Equation (2.1):

$$(4.1) \quad D(x, Q) = x(1 - bQ).$$

Firms always produce at full capacity. Each firm pays an investment cost I per unit of capacity. Assume that the leader has chosen quantity Q_L and consider now the follower. Applying the standard techniques reveals that the follower’s investment trigger $X_F^*(Q_L)$, as a function of the leader’s quantity, must satisfy

$$(4.2) \quad \frac{V(X_F^*(Q_L))}{I} (1 - bQ_L) = \frac{\beta_1 + 1}{\beta_1 - 1};$$

where $V(x) = x/\delta$ and β_1 is as given in (2.5). The next step is to study the leader’s behavior. Given the initial state X_0 , the follower will invest later if the leader sets Q_L such that $X_F^*(Q_L) > X_0$; this is an “entry deterrence” strategy. Alternatively, the leader may “accommodate entry” by allowing the follower to invest immediately. The attractiveness of each strategic stance depends on the initial state

of the process, X_0 . For endogenous firm roles, the leader invests at X_P in capacity $Q_L^*(X_P)$ and the follower invests later at X_F in capacity $Q_F(Q_L^*(X_P))$; thus, the preemption point X_P is implicitly given by $L(X_P) = F(X_P, Q_L^*(X_P))$. For low volatility, the follower chooses a higher capacity than the leader and conversely for high volatility. Compared with the model without capacity choice, the monopolist and the follower invest later in more capacity for high volatility and the leader invests earlier in a higher capacity. Pennings (2004) examines entry timing, quality choice, and pricing in a vertically differentiated duopoly. Here X_t , the number of customers at future time t , is unknown, with $X = \{X_t\}_{t \geq 0}$ following a GBM as in Equation (2.1). Firm profits and investment costs depend on the selected product quality s : a firm's profit (investment cost) is increasing and concave (convex) in s . For tractability, the author assumes linear consumer utility U with uniformly distributed taste. The investment cost is given by $I(s) = \alpha s^\gamma$, with $\gamma > \Pi^*$. In a Stackelberg game, the leader (firm i) decides on its investment trigger X_L^i and product quality s^i as well as on the monopoly and duopoly prices p_L^i and p_C^i . The follower (firm j) chooses the entry trigger X_F^j , the product quality s^j , and the duopoly price p_C^j . The level of demand in each industry structure follows from market-clearing mechanisms and depends on the chosen prices and product qualities. A set of first-order conditions characterizes firms' optimal choices in terms of price, quality, and entry time. Under these assumptions, closed-form solutions exist. For low levels of volatility, firms have an incentive to invest immediately — in which case the leader supplies a product of higher quality than its rival's. For higher degrees of uncertainty, firms delay investments. In this case, the follower invests later for a better quality: $s^j = \kappa s^i$ with $\kappa > 1$. Finally, the degree of product differentiation is increasing with uncertainty because $\partial \kappa / \partial \sigma > 0$.

5. Conclusion and outlook

Van Mieghem (2008) considers the timing of investment and capacity sizing to be key elements of an operations strategy. In this paper, we described a number of research contributions that provide guidance on these issues in the face of uncertainty and competition. Table 3 summarizes the characteristics of these models. In Section 5.1, we highlight key managerial insights derived in the literature on strategic investment under uncertainty. We then provide an outlook for future research in Section 5.2.

5.1. Summary of key managerial insights. Many applications in operations management are meant to refine existing economic models, but they often fail to consider strategic interactions. Decades ago, Porter (1980) noted the importance of industry structure in determining optimal investment strategies. He identified economic and technological uncertainty, competitive (e.g., cost) advantage, capital divisibility, first-mover advantages, and competition intensity (number of incumbents) as major drivers. Taken together, the works discussed in this paper suggest that firms should consider the seven following factors when devising their investment strategies.

Static competitive advantage Pawlina and Kort (2006) demonstrate that, in a duopoly, a firm with a large comparative advantage (e.g., lower investment cost) may enter the market or expand production with little fear of preemption. Firm heterogeneity can thus explain a natural market entry sequence wherein each firm formulates its investment strategy myopically and selects the best entry time as a monopolist would. This sequence is also socially optimal from the viewpoint of the firms. The sequence is reversed in divestment situations: the stronger firm ends up as a monopolist after its weaker rival has exited (see Murto, 2004; Sparla, 2004).

First- versus second-mover advantage The existence of a first-mover advantage generally gives rise to preemption. As shown by Mason and Weeds (2010), however, the presence of second-mover advantages may mitigate the risk of preemption. Cottrell and Sick (2002) point out that managers often tend to underestimate second-mover advantages and consequently invest too early.

Complete versus incomplete information As emphasized by Lambrecht and Perraudin (2003), information asymmetry is not necessarily detrimental to firms because the risk of preemption may thereby be reduced.

Uncertainty about research outcomes may also affect investment behavior: information spillovers in terms of learning externalities may give rise to a second-mover advantage.

Size of capacity increments The size of the lump sums that firms can invest affects their responsiveness to economic changes. As underscored by Murto et al. (2004), the possibility of investing in smaller capacity increments may justify incurring higher (unit) investment costs.²³

Capacity utilization and returns to scale As demand declines, firms reduce capacity utilization to protect against downside risks. When demand is high and new capacity is not yet available, firms may choose high utilization levels despite increasing marginal costs. These effects, well illustrated in Aguerrevere (2003), may explain why output prices often exhibit mean reversion.

Number of competitors Preemption is more intense in highly concentrated markets, as demonstrated by Bouis et al. (2009). If the number of firms is large, then myopic investment behavior can be optimal as shown by Leahy (1993) and Back and Paulsen (2009). This property eases the quantitative analysis of oligopolies with a large number of incumbents (see Grenadier, 2002).

Completion delays Completion delays may explain why certain behaviors persist even under market conditions that prove earlier decisions wrong. Following Grenadier (1996), we can explain real estate construction booms in recession times from this perspective. According to Aguerrevere (2003), volatility may lead completion delays to have an ambiguous effect on aggregate capacity.

5.2. Directions for future research. Trigeorgis (1996) and del Sol and Ghemawat (1999) identify the inadequate treatment of strategic interactions in the analysis of investment under uncertainty as a significant gaps in the research literature. Although some of these gaps have been bridged, there remain many avenues yet to be explored. *Information* is a prime example. Dynamic strategic interactions are not systematically considered in the strategy space (open-loop versus closed-loop). The literature on (stochastic) differential games may help address this issue. Moreover, often firms initially ignore certain characteristics of the underlying market (e.g., the growth rate) but manage to learn it over time: this perspective on information revelation is central to filtering theory and strategic experimentation. In the literature on option games, strategy choices are generally defined in a very simple manner (e.g, solely in terms of timing). Yet real-life investment strategies encompass many other aspects, such as network design. The modeling approaches described in Section 4.3 hold some promise for future development. Most papers in this literature have considered horizontal competition in a given market. However, other parties along the value chain also represent threats or opportunities in investment situations, which means that the analysis of *vertical competition* is also a viable research field. The extensive research in operations management that deals with this subject may provide a good starting point. Most papers have been concerned with deriving analytical results at the expense of *applicability*. To increase the relevance of option games for management practice, new decision-making tool kits — suggested, for example, by Ferreira et al. (2009) — should be developed. This could be achieved by increasing research efforts dedicated to discrete-time approaches and by developing multiplayer stochastic programs and simulation tools. This is a promising approach to the improvement of investment decision-making tools for operational research.

^{23?} analyze the effect of uncertainty on the value of divisibility in a nonstrategic context.

Reference	Option		Industry		Stochastics			Investment type		Demand shock		Key features			
	Entry/expansion	Exit/disposal	Duopoly	Oligopoly	Perfect comp.	Heterogeneity ^a	Discrete-time	Continuous-time	Non-GBM	Lumpy	Repeated lumpy		Incremental	Linear ^b	Multiplicative ^c
Aguerrevere (2003, 2009)	✓	(✓)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Utilization choice turns capacity into call options on output
Baldursson (1998)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Fictitious objective function; myopia
Baldursson and Karatzas (1996)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Càdlàg processes; connection between singular control and optimal stopping
Bouis et al. (2009)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	“Accordion effects” in the investment sequence
Boyer et al. (2010)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Duopoly evolution over time
Dixit and Pindyck (1994)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Premption and rent equalization in new markets
Grenadier (1996)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Completion delays; application to real estate market
Grenadier (1999)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Private signals are revealed through option exercises
Grenadier (2002)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Capacity investment policies in oligopoly
Huisman and Kort (1999)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Formal discussion of new versus existing market
Kulatilaka and Perotti (1998)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Commitment increases payoff convexity
Lambrecht and Perraudin (2003)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	BNE with privately known investment cost
Leahy (1993)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Optimality of myopic strategies
Mason and Weeds (2010)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Analysis of positive competitive externalities
Murto (2004)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Exit (war of attrition) in duopoly
Murto et al. (2004)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Tradeoff between increment size and unit investment cost
Nishihara and Fukushima (2008)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Information asymmetry; exogenous entry ordering
Novy-Marx (2007)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Investment multiple κ characterizes time-invariant industry structure
Pawlina and Kort (2006)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Asymmetric investment costs
Pennings (2004)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Vertical differentiation in duopoly
Smit and Ankum (1993)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Binomial option game tree
Sparla (2004)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	“Closure options” and war of attrition
Spencer and Brander (1992)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Endogenous quantity precommitment
Thijssen et al. (2006)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Unknown project quality; Poisson signals
Weeds (2002)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Poisson process for technological uncertainty in R&D

^aHeterogeneity refers to competitive advantages or other differentiating characteristics among firms (including information asymmetries).

^bLinear demand corresponds to Equation (2).

^cMultiplicative demand shock encompasses isoelastic demand as in Equation (11).

FIGURE 3. Summary of selected contributions

Market-Entry Sequencing under Uncertainty

ABSTRACT. At the early stages of industry development, before operating in the market, firms may identify an opportunity to enter yet wait for the market to grow sufficiently (to justify the expense of market entry). We model this problem and show that, in oligopolies, market entry involving lumpy investments takes place in sequence under uncertainty regardless of whether or not firms can observe (and react to) their rivals' moves. Unlike the case in which firms disregard competition (the myopic or open-loop case), entry does not occur in a socially optimal manner when firms condition their moves on competitors' actions. Thus, additional information about rivals' play may lead to social loss.

1. Introduction

It is standard practice in finance and strategy to interpret real investment opportunities as being analogous to options. This view is well accepted among academics and practitioners alike and is the core of real options analysis (ROA). It allows one to capture the dynamic nature of decision making because it factors in management's flexibility to revise and adapt strategy in the face of market uncertainty. This approach is summarized in the works of Dixit and Pindyck (1994) and Trigeorgis (1996).

Standard ROA, however, has not adequately addressed strategic interactions. The exercise decisions of a financial option holder have little impact on the underlying asset's value dynamics. Since mainstream ROA draws from contingent-claims analysis (based on financial options), a key feature specific to real assets is often ignored — namely, the interplay or strategic interactions taking place among (real) option holders. This oversight may result in misestimating the value of strategic options and in suboptimal investment policies. The inadequate treatment of competitive dynamics among (real) option holders remains one of the most compelling research gaps (Trigeorgis 1996, p. 376). Bridging this gap requires concurrently taking into account both market and strategic uncertainties via the combined use of ROA and game theory.¹

In modeling multistage, multiplayer problems, one has to consider the possibility that players will react optimally in the future, conditional on the information they gather over time.² In order to analyze strategic interactions in multistage deterministic settings, one often distinguishes two types of information structures. In *closed-loop* strategies, the players can condition their actions (when called upon to play) on the observed sequence of rivals' moves. In contrast, *open-loop* strategies do not depend on rivals' previous play and instead set ex ante a certain investment path to be pursued regardless of the rivals' decisions. If the information structure allows players to condition their actions on previous plays, as in closed-loop strategies, then the subgame-perfect solution concept may yield stronger predictions.³ Nash

¹Smit and Trigeorgis (2004) discuss discrete-time option games, whereas Grenadier (2000b) and Huisman (2001) examine a number of continuous-time models. Chevalier-Roignant and Trigeorgis (2011) provide an overview of both types of "option games."

²This has different implications depending on whether a single player is faced with deviations from an expected market development or multiple players must predict what will be the outcome of strategic interactions over time. The first case is characterized by Bellman's principle of optimality, whereby the agent acts optimally in every (exogenous) state of the world; the second case is characterized by subgame perfection, whereby players act optimally (as part of a Nash equilibrium) both on and off the equilibrium path.

³In most cases, perfect equilibria are best because most dynamic games assume negligible information lags between a player's move and rivals' observations. However, Nash equilibria in open-loop strategies are best when information lags are infinite or firms are precommitted to a certain path of investment. See Fudenberg and Tirole (1991) for further details.

equilibria in open-loop strategy profiles often fail to be subgame perfect when firms can observe (and react to) their rivals' moves.⁴ The open-loop approach is often employed in economic analysis when many small players cannot condition their play on their opponents' actions. Fudenberg and Levine (1988) have shown that, as the number of players increases, the outcome of a perfect equilibrium in closed-loop strategies converges to a Nash equilibrium in open-loop strategies. Therefore, the Nash equilibrium in open-loop strategies provides a good benchmark and reasonable predictions for large oligopolies when firms formulate their strategy in isolation and their decisions do not materially affect the decision making of rivals.

The distinction between open-loop and closed-loop strategies has not always been clear in the emerging literature on "option games." Several authors (e.g., Leahy, 1993; Baldursson and Karatzas, 1996; Baldursson, 1998) have examined investment behavior in (perfectly) competitive markets under uncertainty where firms use capital stock as a control variable and may increase that stock by a small amount (infinitely divisible capital).⁵ These models assume that the Nash equilibrium (in open-loop strategies) is the appropriate solution concept and do not discuss which information structure is the most descriptive of the dynamic game considered. A new strand in the literature (see, e.g., Back and Paulsen, 2009; Novy-Marx, 2009) considers perfect equilibria in an oligopoly where firms make incremental capital investments. In contrast, several real options models (see, e.g., Smets, 1991; Grenadier, 1996; Huisman and Kort, 1999) have been developed in a duopoly setting and examine lumpy investments under closed-loop strategies.

In this paper, we address a discrete or lumpy investment decision in the context of an oligopoly, examining whether firms should enter the market now or later.⁶ This problem was previously studied in the field of industrial organization previously by Reinganum (1981a,b); Gilbert and Harris (1984); Fudenberg and Tirole (1985), and Reynolds (1987). However, these authors assume a deterministic market environment. In this paper, we link the two previous approaches by exploring optimal market-entry strategies that allow for uncertainty in market development as modeled by stochastic (Itô) processes. To the best of our knowledge, only Bouis et al. (2009) have addressed oligopoly (more than two firms) and lumpy investment. Our setting and modeling approach differs from theirs which uses numerical analysis to study the investment behaviors in large oligopolies. Here we shall deal concurrently with an exogenous shock and endogenous market-entry decisions.

The paper is organized as follows. Section 2 sets up the problem faced by would-be market entrants and presents the model assumptions. Section 3 examines the investment timing problem in a new market under uncertainty and demonstrates that investment takes place in sequence regardless of whether or not firms can condition their decision on previous rivals' moves. The open-loop equilibrium in Section 3.1 describes an oligopoly with a large number of incumbents (i.e., perfect competition); the closed-loop duopoly model is discussed in Section 3.2. Section 4 compares the market-entry sequences with the socially optimal benchmark, confirming that only the open-loop market-entry sequence is necessarily socially optimal; in the closed-loop case, where firms can observe and react to their rivals' previous moves, the market-entry sequence need not be socially optimal. Section 5 concludes.

⁴In open-loop models there is a single proper subgame (the game as a whole); this explains why the Nash equilibrium in open-loop strategies is subgame perfect if firms cannot condition their actions on rivals' play.

⁵As noted by Pindyck (1988), the assumption that firms can continuously add incremental amounts of capital is an extreme one; in fact, most business problems involve discrete or lumpy investment decisions.

⁶In reality, firms typically do not limit their market-entry strategy solely to determination of the entry time; rather, they determine the staging of entry, the appropriate scale of production, the type of product to manufacture, etc. For simplicity, economists frequently abstract market-entry strategy to focus on the timing issue in isolation.

2. Model setup and assumptions

We consider settings to involve complete information on historic market developments and firm investment decisions and in which (option-holding) firms have common priors about the probabilities of (future) exogenous events (homogeneous expectations). Consider the filtered probability space $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$, where the filtration $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t \geq 0}$ is a family of “tribes” that allow perfect recall: $\mathcal{F}_t \subseteq \mathcal{F}_s$ for all $s > t$. The tribe \mathcal{F}_t denotes for the possible histories (information set) on which the decision maker bases her decision at time t ; the information set keeps track of the exogenous market development and the evolution of industry structure. For tractability, we restrict ourselves to a Markov environment where the latest (i.e., current) state (shock and industry structure) is a sufficient statistic on which firms can condition their decision making.⁷ We consider an oligopoly with n identical firms, where each firm has an infinitely lived investment option (modeled as a perpetual American call, as in McDonald and Siegel 1986) but where no firm is active in the marketplace at the outset. When new firms enter the market, they incur a positive lumpy (sunk, exogenous) investment cost I .

Assume the market is subject to an exogenous (\mathbb{F} -adapted⁸) shock $\{X_t\}_{t \geq 0}$ that follow a time-homogenous Itô process characterized by the stochastic differential equation

$$(2.1) \quad dX_t = \mu(X_t) dt + \sigma(X_t) dB_t,$$

where $z = \{z_t\}_{t \geq t_0}$ is a standard Brownian motion and the process starts at $X_0 = x$ almost surely.⁹ This process covers a fairly broad family of stochastic processes used in economic analysis, such as geometric Brownian motion and Ornstein–Uhlenbeck process. Assume that the drift process exhibits a long-term positive growth. The state of the industry structure is captured by an integer-valued (\mathbb{F} -adapted) nondecreasing process $m = \{m_t\}_{t \geq 0}$ that indicates the number of incumbent firms as of time t .

The revenue function $r(\cdot, \cdot)$ of an incumbent firm at time t depends on the current value of the exogenous shock and also on the number of incumbent firms:

$$\rho : \mathbb{R} \times \mathbb{N} \rightarrow [0, \infty).$$

The (stochastic) revenue function is \mathcal{F}_t -measurable, twice continuously differentiable, and nondecreasing in the shock; it is also decreasing in the number of incumbent firms. Denote the cost incurred by an incumbent firm at time t by

$$c : \mathbb{R} \times \mathbb{N} \rightarrow [0, \infty).$$

The cost function $c(\cdot, \cdot)$ also is \mathcal{F}_t -measurable, twice continuously differentiable, and nondecreasing in the shock. When new entrants arrive, this cost might be reduced owing to spillover effects, best practices, or asset sharing with competitors. On the other hand, the cost might increase in the number of incumbent firms in response to higher demand for inputs.¹⁰ The instantaneous profit flow $\pi(\cdot, \cdot) \equiv \rho(\cdot, \cdot) - c(\cdot, \cdot)$ to an incumbent is \mathcal{F}_t -measurable, twice continuously differentiable, and strictly increasing in the shock. However, it is strictly decreasing in m_t because competitive arrivals are viewed as a negative externality.

⁷This restriction enables us to reduce the state space and avoid the “curse of dimensionality” by allowing the use of Markov strategies (i.e., strategies based on the latest values of states).

⁸An \mathbb{F} -adapted process is \mathcal{F}_t -measurable for all $t \in \mathbb{R}_+$. The assumption of adaptedness to the filtration means that no foresight about the future economic development is permitted; this is equivalent to the non-anticipativity constraint in operations research.

⁹For technical reasons (existence conditions), assume further that the drift and the volatility processes ($\mu : \mathbb{R} \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$, respectively) are \mathbb{F} -adapted and have finite variations and, assume there exist constants $M, M' \in \mathbb{R}_+$ such that: (i) $|\mu(X_t)| + |\sigma(X_t)| \leq M(1 + |X_t|)$, for all $X_t \in \mathbb{R}$, $t \geq 0$ (linear growth condition); (ii) $|\mu(X_t^1) - \mu(X_t^2)| + |\sigma(X_t^1) - \sigma(X_t^2)| \leq M'|X_t^1 - X_t^2|$, for all $X_t^1, X_t^2 \in \mathbb{R}$, $t \geq 0$ (Lipschitz condition).

¹⁰If $c(\cdot, \cdot)$ were decreasing in m_t , we would assume that the decrease in $c(\cdot, \cdot)$ is strictly less than the decrease in $\rho(\cdot, \cdot)$.

We assume that risk-neutral firms face a constant, risk-free rate r that is common to all market participants (including the central planner).¹¹ Assumption 1 reflects the market’s incipience, with no firm investing or operating at the outset. Firms that have not yet entered the market are not affected by rivals’ investment decisions.

ASSUMPTION 1. *No firm is active at the outset ($m_0 = 0$), and none has an incentive to immediately invest:*

$$\mathbb{E}_0 \left[\int_0^\infty e^{-rs} \pi(X_s, 1) ds - I \right] < 0.$$

The operator $\mathbb{E}_t[\cdot]$ is used henceforth as shorthand for $\mathbb{E}[\cdot | \tilde{h}_t]$, that is, the expectation conditional on the (payoff-relevant) history $\tilde{h}_t \equiv (X_t, m_t)$.

3. Market-entry sequencing

We now show that regardless of whether or not firms can condition their action on previous moves (closed-loop versus open-loop strategies), the equilibrium market entries are characterized by a time sequencing or ordering of the firms. We first examine a large oligopoly in which firms do not observe their rivals’ play (open-loop strategies). We then discuss the preemption that arises when firms observe their rival’s move in a duopoly model (closed-loop strategies). This latter approach has been formulated by Dixit and Pindyck (1994) for a multiplicative shock that follows a geometric Brownian motion and (implicit) mixed strategies. We shall derive an analogous result for the general Itô process with related behavioral strategies.

3.1. Market-entry sequencing for open-loop strategies. As already noted, the Nash equilibrium in open-loop strategies is a good approximation for large oligopolies. To adapt the game-theoretic notion of open-loop strategies (developed in a deterministic environment) to a setting where market features evolve stochastically, we must refine the notion of open-loop strategies to account for the exogenous shock. In this context, an *investment policy* is a decision rule based solely on the observed resolution of market uncertainty over time but not on the sequence of moves in the industry.¹² This assumption is reasonable in some circumstances, since it may be easier to gather information about the prospects of a given market than about rivals’ likely competitive moves.

DEFINITION 1. *An investment policy is a decision rule that maps, for every possibly state $x \in \mathbb{R}$ of the shock, an action $a_t^i \equiv a^i(x) \in A^i(h_t) \subseteq \{0, 1\}$. Here, a_t^i indicates whether firm i operates (1) or not (0), and A_t^i is the stage action set.*

An investment policy is a pure strategy — that is, a map from the information set (where the payoff-relevant history is the value of the exogenous shock in the Markov environment) to actions (to “enter” or “not”). Investment is assumed to be irreversible, so the firm continues to operate indefinitely once investment takes place. Given the infinite planning horizon, Bellman’s principle of optimality prescribes that the existence of a fixed threshold level X_i that partitions the state space for the exogenous shock such that $(-\infty, X_i)$ is the inaction region and $[X_i, \infty)$ is the action region. We can thus simplify the strategy formulation and define it as selecting the investment threshold X_i , rather than specifying the mapping in Definition 1. We associate to this threshold X_i a (\mathbb{F} -adapted) stopping time $\tau_i \equiv \inf \{t \geq 0 \mid X_t \geq X_i\}$ at

¹¹The assumption of risk neutrality can be relaxed if arbitrage opportunities in the market do not exist and the market is complete. Along the lines of Cox and Ross (1976) and Harrison and Kreps (1979), we could replace the drift in the stochastic differential equation (2.1) by the one that is prevalent in a risk-neutral environment. In this case, B would be a standard Brownian motion under the equivalent martingale measure.

¹²In this sense, an investment policy is open-loop with respect to the filtration of the strategy space but closed-loop with respect to nature’s moves.

which investment occurs.¹³ The assumption of irreversibility makes it possible to relate the stage action a_t^i in Definition 1 to the stopping time τ_i via the relationship

$$a_t^i = \mathbb{1}_{\{t \geq \tau_i\}}, \quad t \geq 0,$$

where $\mathbb{1}$ is the indicator function. The process that indicates the number of incumbent firms at time t is $m = \{m_t\}_{t \geq 0}$, where $m_t = \sum_{i=1}^n a_t^i$. Given Definition 1, we can now define the Nash equilibrium in investment policies for a generic payoff function $F^i(\cdot, \cdot)$. Let X_{-i} denote the strategy profile of all firms *except* firm i .

DEFINITION 2. *The profile of investment policies $X^* = (X_1^*, \dots, X_n^*)$ is a Nash equilibrium if and only if, for each firm $i = 1, \dots, n$,*

$$F^i(X_i^*, X_{-i}^*) \geq F^i(X_i, X_{-i}^*), \quad \forall X_i \in \mathbb{R}.$$

If a firm were to enter the market immediately, then it would receive (at time t)

$$(3.1) \quad W(X_t, m_t) = \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} \pi(X_s, m_s) ds - I \right].$$

3.1.1. *Myopic behavior in large oligopolies.* The investor's present expected value when firm i decides to enter the market at time τ_i is given by

$$(3.2) \quad V_0^i(X_i, X_{-i}) = V^i(x; X_i, X_{-i}) \equiv \mathbb{E}_0 \left[\int_0^\infty e^{-rs} a_s^i \pi(X_s, m_s) ds - e^{-r\tau_i} I \right].$$

Here $\{m_s\}_{s \geq 0}$ evolves over time with the arrival of new firms in the marketplace. For a given profile of investment policies by rivals X_{-i} , firm i chooses its investment policy to maximize its payoff $V_0^i(X_i, X_{-i})$. Determining the Nash equilibrium in investment policies corresponds to the multiplayer optimization problem

$$(3.3) \quad V_0^i(X_i^*, X_{-i}^*) \equiv \sup_{X_i \in \mathbb{R}} V_0^i(X_i, X_{-i}^*), \quad \forall i = 1, \dots, n.$$

Here firm i takes account of the impact of future investments on its value function when formulating its optimal strategy. We then refer to i as a *strategic* firm.

In contrast, a *myopic* firm ignores the investments of other players in its maximizing behavior. In other words, it invests as if no other investment occurs in the industry. The value that myopic investors expect to receive (i.e., their beliefs about payoff) is

$$(3.4) \quad v_0^i(X_i) = v^i(x; X_i) \equiv \mathbb{E}_0 \left[\int_0^\infty e^{-rs} a_s^i \pi(X_s, m_{\tau_i}) ds - e^{-r\tau_i} I \right].$$

The number of incumbents in equation (3.4) is hypothesized to remain constant (at level m_{τ_i}) after firm i 's entry, whereas the strategic firm's maximization problem in equation (3.2) takes account of future firm arrivals. Because firm arrivals induce negative externalities on the incumbent firm's profit, the myopic firm believes to it is entitled to a higher value than is a strategic firm for the same strategy profile $X = (X_i, X_{-i})$. This may be expressed formally as follows:

$$V_0^i(X_i, X_{-i}) \leq v_0^i(X_i) < \infty, \quad \forall X = (X_i, X_{-i}) \in \mathbb{R}^n.$$

Both $V^i(\cdot; \cdot, \cdot)$ and $v^i(\cdot; \cdot)$ are increasing in x . The myopic firm behaves rationally under the restriction that the effects of rivals' investments on its value can be ignored. In this case, the optimal strategy solves the following problem:

$$(3.5) \quad v_0^i(X_i^*) \equiv \sup_{X_i \in \mathbb{R}} v_0^i(X_i), \quad \forall i = 1, \dots, n.$$

Proposition 2 establishes that both strategic and myopic firms follow the same Nash equilibrium strategies. The intuition behind this claim involves the strategic firm's taking into account a (negative) value

¹³The finiteness of the first-hitting time typically involves conditions on the drift of the exogenous shock process $\{X_t\}_{t \geq 0}$. These conditions are specific to the case considered.

component that is not material for strategy formulation. This property simplifies the underlying problem, allowing one to derive the Nash equilibrium investment policy profiles of strategic firms based on the myopic firm's simpler value functions of equation (3.4).

PROPOSITION 2. *In a Nash equilibrium in investment policies, a strategic firm formulates the same investment policy as a myopic firm.*

PROOF. Define $CE(\cdot, \cdot) \equiv V_0^i(\cdot, \cdot) - v_0^i(\cdot)$, where CE denotes competitive erosion. From equations (3.2) and (3.4), it follows that

$$\begin{aligned} CE(X_i, X_{-i}) &= \mathbb{E}_0 \left[\int_0^\infty e^{-rs} a_s^i \{ \pi(X_s, m_s) - \pi(X_s, m_{\tau_i}) \} ds \right] \\ &= \mathbb{E}_0 \left[\int_{\tau_i}^\infty e^{-rs} \{ \pi(X_s, m_s) - \pi(X_s, m_{\tau_i}) \} ds \right]. \end{aligned}$$

Assume a weak ordering of firms, with firm i investing at $\tau_i < \infty$ (and firm $i + 1$ at $\tau_i \leq \tau_{i+1} < \infty$), such that $i = m_{\tau_i}$ firms are operating in the time interval $[\tau_i, \tau_{i+1}]$. Then

$$\begin{aligned} CE(X_i, X_{-i}) &= \mathbb{E}_0 \left[\int_{\tau_i}^{\tau_{i+1}} e^{-rs} \{ \pi(X_s, i) - \pi(X_s, i) \} ds \right. \\ &\quad \left. + \int_{\tau_{i+1}}^\infty e^{-rs} \{ \pi(X_s, m_s) - \pi(X_s, i) \} ds \right] \\ (3.6) \quad &= \mathbb{E}_0 \left[\int_{\tau_{i+1}}^\infty e^{-rs} \{ \pi(X_s, m_s) - \pi(X_s, i) \} ds \right]. \end{aligned}$$

This shows that $CE(X_i, X_{-i})$ depends not on X_i but only on X_{-i} , so $CE(X_i, X_{-i}) = CE(X_{-i})$ and $V_0^i(X_i, X_{-i}) = v_0^i(X_i) + CE(X_{-i})$. Furthermore, since the profit is decreasing in the number of firms, we have $CE(X_{-i}) \leq 0$. The term $CE(X_{-i})$ can thus be interpreted as the competitive value erosion incurred by the (strategic) investor. Following Slade (1994), call $v_0^i(\cdot)$ the *fictitious objective function*. Observe that $CE(X_{-i})$ does not affect the optimizing behavior of the strategic firm i :

$$X_i^* \in \arg \max_{X_i \in \mathbb{R}} V_0^i(X_i, X_{-i}^*) \iff X_i^* \in \arg \max_{X_i \in \mathbb{R}} v_0^i(X_i).$$

□

Proposition 2 can now be used to derive sufficient conditions for the Nash equilibrium based on the value function of the myopic firm in equation (3.4). These conditions are summarized in Proposition 3. For tractability, we omit the dependence of the value functions on the information set and on the investment trigger X_i^* .

PROPOSITION 3. *Firm i 's optimal investment policy X_i^* must solve the following system of equations:*

$$(3.7a) \quad rv^i - \mathcal{D}v^i = 0,$$

$$(3.7b) \quad v^i|_{x=X_i^*} = W^i|_{x=X_i^*},$$

$$(3.7c) \quad v_x^i|_{x=X_i^*} = W_x^i|_{x=X_i^*},$$

$$(3.7d) \quad \lim_{x \rightarrow 0} v^i(x; X_i^*) = 0.$$

The term $\mathcal{D}v^i$ in equation (3.7a) corresponds to the expected capital gain over an infinitesimally small time interval. It is given by

$$\begin{aligned} \mathcal{D}v^i &\equiv \lim_{h \rightarrow 0} \frac{\mathbb{E}_t[v_{t+h}^i] - v_t^i}{h} \\ (3.8) \quad &= v_x^i \mu(X_t) + \frac{1}{2} v_{xx}^i \sigma(X_t)^2. \end{aligned}$$

Here v_x^i and v_{xx}^i indicate respectively the first- and second-order derivatives of v_0^i with respect to $X_0 = x$ (\mathcal{D} is the infinitesimal generator in stochastic calculus). Equation (3.7a) is the Hamilton–Jacobi–Bellman (HJB) equation. Equation (3.7b) is the value-matching condition; it prescribes that, at the time of optimal exercise, the firm is indifferent between investing now (and receiving W) and waiting (and receiving obtaining value v^i). The “smooth pasting” condition represented by equation (3.7c) ensures that the first-order derivative of the value function is continuous at the optimal threshold. Finally, condition (3.7d) implies that the value function does not “explode” when the underlying shock has a low value. These conditions are fairly standard in real options analysis.

PROOF. The proof is equivalent to that for the Stefan problem in optimal stopping theory as derived by Peskir and Shiryaev (2006), for example. \square

3.1.2. *Market-entry sequencing as Nash equilibrium.* We have characterized the optimal Nash investment policies formulated by option-holding firms, but we have not yet characterized the timing of when market entry takes place. We next examine whether firms enter simultaneously, sequentially, or sequentially but with clustering effects.

Without loss of generality, we assume a weak ordering of investment with firm i denoting the i th investor. From equation (3.2), it follows that

$$(3.9) \quad V_0^i(X_i, X_{-i}) = \mathbb{E}_0 \left[\sum_{k=i}^n \int_{\tau_k}^{\tau_{k+1}} e^{-rs} \pi(X_s, k) ds - e^{-r\tau_i} I \right],$$

where (by convention) $\tau_{n+1} \equiv \infty$.

PROPOSITION 4. *Given a weak ordering of market-entry times, each firm has a unique optimal threshold X_i^* . The thresholds for $i = 1, \dots, n$ are distinct, and investment takes place in sequence. Formally:*

$$0 < \tau_1^* < \dots < \tau_n^* < \infty,$$

where $\tau_i^* = \inf \{t \geq 0 \mid X_t \geq X_i^*\}$.

PROOF. We know that $v_0^i(\cdot)$ is strictly increasing in $[X_{i-1}^*, X_{i+1}^*]$, so the function

$$(3.10) \quad \mathcal{L}v^i \equiv rv^i - \mathcal{D}v^i$$

is also strictly increasing. Here $\mathcal{D}v^i$ is as defined in equation (3.8), and by equation (3.7a) we have $\mathcal{L}v^i|_{x=X_i^*} = 0$. To see that $X_i^* \in (X_{i-1}^*, X_{i+1}^*)$, evaluate (3.10) at $X_i = X_{i-1}^*$. Since $\mathcal{L}v^i$ is strictly increasing $\mathcal{L}v^i|_{x=X_{i-1}^*} < 0$. Similarly, it follows that $\mathcal{L}v^i|_{x=X_{i+1}^*} > 0$ at $X_i = X_{i+1}^*$. The threshold X_i^* (characterized by the conditions in Proposition 3) is the unique maximum of $v_0^i(\cdot)$ in (X_{i-1}^*, X_{i+1}^*) . \square

PROPOSITION 5. *The Nash equilibria in investment policies are characterized by the following properties.*

- (i) *The set (n -tuple) of market-entry thresholds $X^* = (X_1^*, \dots, X_n^*)$ where X_i^* , $i = 1, \dots, n$, solves equations (3.7a)–(3.7d), is a Nash equilibrium in investment policies.*
- (ii) *There are in total $n!$ Nash equilibria in market-entry policies, and each is characterized as a permutation of the market-entry thresholds in part (i).*

PROOF. See Appendix I. \square

It cannot be readily predicted which Nash equilibrium is the most likely to occur among the $n!$ equilibria of Proposition 5.¹⁴ Example 10 considers a special case involving a multiplicative shock that follows the geometric Brownian motion and admitting a closed-form solution.¹⁵

¹⁴If firms were asymmetric, then we could reasonably predict that the more likely investment sequence would be the socially optimal one by employing focal-point or commonsense considerations.

¹⁵Proposition 5 is derived for the fairly general case of when the profit functions are twice continuously differentiable in the shock and are subject to a shock that follows a time-homogeneous Itô process. Existence of an analytical solution to

EXAMPLE 10. Suppose the shock $\{X_t\}_{t \geq 0}$ follows the geometric Brownian motion

$$dX_t = \mu X_t dt + \sigma X_t dB_t,$$

where $\mu, \sigma \in \mathbb{R}_+$ and $\mu < r$. The shock enters the profit function multiplicatively: $\pi(X_t, m_t) = X_t \pi(m_t)$ for all $X_t \in \mathbb{R}_+$ and all $m_t \in \mathbb{N}$. Then the optimal (Nash equilibrium) threshold X_i^* for the i th market entrant, $i = 1, \dots, n$, is such that

$$X_i^* \frac{\pi(i)}{\delta I} = \frac{\beta_1}{\beta_1 - 1}.$$

Here $\delta \equiv r - \mu > 0$, and β_1 is the positive root of the “fundamental quadratic” given by

$$\beta_1 \equiv \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + 2\frac{r}{\sigma^2}} \quad (> 1).$$

When firms pursue the Nash equilibrium investment policies, the value function for a (strategic) firm $i = 1, \dots, n$ is

$$(3.11) \quad V_0^i(X_i^*, X_{-i}^*) = \left(\frac{X_0}{X_i^*}\right)^{\beta_1} \left(X_i^* \frac{\pi(i)}{\delta} - I\right) + CE(X_{-i}^*),$$

where the last competitive erosion term is

$$CE(X_{-i}^*) = \sum_{m=i+1}^n \left(\frac{X_0}{X_m^*}\right)^{\beta_1} X_m^* \left(\frac{\pi(m) - \pi(m-1)}{\delta}\right) \quad (\leq 0).$$

PROOF. In the case of geometric Brownian motion, it follows from equations (3.7a) and (3.8) that the solution solves the second-order differential equation

$$rv^i - \mu X_t v_x^i + \frac{1}{2} \sigma^2 X_t^2 v_{xx}^i = 0.$$

This equation has solutions of the form

$$v^i(x; X_i) = Ax^{\beta_1} + Bx^{\beta_2},$$

where A and B are constants to be determined and

$$\beta_2 \equiv \frac{1}{2} - \frac{\mu}{\sigma^2} - \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + 2\frac{r}{\sigma^2}} \quad (< 0).$$

From the boundary conditions (3.7b)–(3.7d), one can derive the closed-form solutions. According to (3.7d), $B = 0$. By (3.7b) and (3.1), we have $v^i(X_i^*; X_i^*) = A(X_i^*)^{\beta_1} = X_i^* \frac{\pi(i)}{\delta} - I$ and so

$$(3.12) \quad A = (X_i^*)^{-\beta_1} \left[\frac{\pi(i)}{\delta} X_i^* - I \right].$$

From (3.7c), it follows that $v_x^i(X_i^*; X_i^*) = A\beta_1 (X_i^*)^{\beta_1-1} = \frac{\pi(i)}{\delta}$. By (3.12),

$$\beta_1 \left[(X_i^*)^{\beta_1-1-\beta_1+1} \times \frac{\pi(i)}{\delta} - (X_i^*)^{\beta_1-1-\beta_1} I \right] = \frac{\pi(i)}{\delta}.$$

Therefore, $X_i^* = \frac{\delta I}{\pi(i)} \frac{\beta_1}{\beta_1-1}$. The value of the strategic firm is obtained by specializing equations (3.2) and (3.6). \square

Proposition 6 establishes that early entrants into an emerging market are better-off than later entrants. It is interesting that, although firms are assumed to be symmetric, they receive asymmetric values in equilibrium (open-loop approach).

this problem is not ensured generally but only under restrictive assumptions for the process and the profit function. The geometric Brownian motion is standard in economic analysis because it is fairly descriptive of problems faced by economic agents and often yields closed-form solutions.

PROPOSITION 6. *In a Nash equilibrium in investment policies, firms' equilibrium values decline monotonically with their order of entry: an earlier investor is better-off (on average) than a later market entrant. Hence there exist early-mover advantages $V_0^i(X^*) > V_0^j(X^*)$, $i < j$.*

PROOF. For two firms i and j , assume a weak ordering of investments with firm i investing earlier than firm j , so that $X_i^* < X_j^*$ or $\tau_i^* < \tau_j^*$. It then follows from the uniqueness of the maximum X_i^* and the definition of the Nash equilibrium in (5) that

$$V_0^i(X_i^*, X_j^*, X_{-i,j}^*) > V_0^i(X_j^*, X_j^*, X_{-i,j}^*).$$

Because firms are symmetric, we have

$$\begin{aligned} V_0^i(X_j^*, X_j^*, X_{-i,j}^*) &= V_0^j(X_j^*, X_j^*, X_{-i,j}^*), \\ &\geq V_0^j(X^*). \end{aligned}$$

The last inequality follows because firm j is not worse-off if firm i invests later (at $X_j^* > X_i^*$). \square

It may seem counterintuitive that symmetric firms receive asymmetric values in Nash equilibrium (with open-loop investment policies) and that firms do not struggle to become the first investor. These results stem from the assumption that firms act myopically and do not revise their strategy over time. For closed-loop strategies, as discussed next, results may well differ because firms can react to their opponents' moves.

3.2. Market-entry sequencing for closed-loop strategies. In the Section 3.1, we assumed that the investment option was available to a large number of firms, so that potential entrants could not observe their rivals' moves. This assumption was sufficient for the Nash equilibrium in investment policies to be a reasonable solution concept and to provide the most accurate prediction concerning the industry dynamics. The optimal investment triggers were selected myopically in that they were not linked to the investment decisions of future entrants (thus, firms behaved as if they were the last investor to enter the market or like a monopolist with an exclusive right to enter). In this section, we look at the problem where (exogenous) entry barriers exist that prevent more than two firms from entering the market.

Next, we briefly summarize what happens when firms devise investment strategies that are closed-loop with respect to the filtration of (both) the strategy and the state space, as opposed to strategies that are open-loop with respect to the strategy space.¹⁶

3.2.1. *Markov perfection in investment strategies.* From a game-theoretic viewpoint, investment policies (as formulated in Definition 1) are pure strategies whereby players select an investment path $\{a_t^i\}_{t \geq 0}$ in response to the actual development of a market shock but ignore the industry structure evolution $\{m_t\}_{t \geq 0}$. Fudenberg and Tirole (1985) have shown that, in a deterministic duopoly setting, closed-loop investment strategies cannot be considered in a similar manner (i.e., via pure strategies). Only randomization allows us to give theoretical explanations for certain effects that are heuristically expected, such as preemption or a war of attrition.

In some circumstances or subgames, the pure-strategy formulation for deriving what to do (a binary decision) in each state of the world fails to result in a unique Nash equilibrium in pure actions; this failure is due to the lack of convexity of the instantaneous discrete action set. Hence, "convexifying" the instantaneous action set may help solve this problem by randomization and the formulation of market-entry strategies as behavioral strategies in continuous time. A behavioral strategy specifies a probability

¹⁶This problem has been discussed heuristically in Dixit and Pindyck (1994, Chap. 9) for the case of a multiplicative shock that follows a geometric Brownian motion. Here, we derive a solution for a general shock — not necessarily multiplicative — that follows a time-homogenous Itô process. For a discussion close to ours but employing mixed strategies, see Thijssen, Huisman, and Kort (2002).

distribution $\Delta A^i(\tilde{h}_t)$ over pure actions $A^i(\tilde{h}_t)$ for each history path $\tilde{h}_t = (X_t, m_t)$, where the probability distributions for different histories are independent.¹⁷

DEFINITION 3. A behavioral market-entry strategy is a decision rule that maps every possible history $\tilde{h}_t = (X_t, m_t)$ to a mixed action $\alpha_t^i \equiv \alpha^i(\tilde{h}_t) \in \Delta A^i(\tilde{h}_t) \subseteq [0, 1]$.¹⁸ The mixed action α_t^i measures the instantaneous probability of investing in the short time interval $[t, t + h]$ in state \tilde{h}_t as h approaches zero.¹⁹

Because investment is irreversible, the mixed action set may change over an industry's evolution. Formally, $\Delta A^i(\tilde{h}_t) \in \{1\}$ after firm i 's entry. In light of Definition 3, we can pin down the appropriate solution concept — namely, Markov perfect equilibrium.

DEFINITION 4. A Markov perfect equilibrium (MPE) is a profile of Markov investment strategies $\alpha^* = (\alpha_t^{i*}, \alpha_t^{j*})_{t \geq 0}$ that form a perfect equilibrium; in other words, $(\alpha_s^{i*}, \alpha_s^{j*})_{s \geq t}$ is a Nash equilibrium for all \tilde{h}_t , $t \geq 0$.²⁰

We consider next the values of strategic firms. The value of a leader investing at time t is

$$(3.13) \quad L(X_t) \equiv \mathbb{E}_t \left[\int_t^{\tau_F^*} e^{-r(s-t)} \pi(X_s, 1) ds + \int_{\tau_F^*}^{\infty} e^{-r(s-t)} \pi(X_s, 2) ds - I \right],$$

where X_F^* is the threshold at which the follower enters and $\tau_F^* = \inf \{t \geq 0 \mid X_t \geq X_F^*\}$ is the follower's entry time. The value of the follower (at time t) is

$$(3.14) \quad F(X_t) \equiv \begin{cases} \mathbb{E}_t \left[\int_{\tau_F^*}^{\infty} e^{-r(s-t)} \pi(X_s, 2) ds - e^{-r(\tau_F^*-t)} I \right] & \text{if } X_t \leq X_F^*, \\ \mathbb{E}_t \left[\int_t^{\infty} e^{-r(s-t)} \pi(X_s, 2) ds - I \right] & \text{if } X_t > X_F^*. \end{cases}$$

Alternatively, the option-holding firms might decide to invest at the same time, which would result in the following value of simultaneous immediate investment:

$$(3.15) \quad M(X_t) \equiv \mathbb{E}_t \left[\int_t^{\infty} e^{-r(s-t)} \pi(X_s, 2) ds - I \right].$$

For $t \geq \tau_F$, we have $M(X_t) = F(X_t)$.

PROPOSITION 7. Given that no firm invests at the outset (Assumption 1), the profile of symmetric market-entry strategies $\alpha^* = (\alpha_t^{i*}, \alpha_t^{j*})_{t \geq 0}$ written as

$$\alpha^*(x) = \alpha^{i*}(x) = \alpha^{j*}(x) = \begin{cases} 0 & \text{if } x \in (-\infty, X_P^*) & [\text{don't invest for low values}], \\ \phi(x) & \text{if } x \in [X_P^*, X_F^*) & [\text{mix for intermediate values}], \\ 1 & \text{if } x \in [X_F^*, \infty) & [\text{invest for high values}], \end{cases}$$

where

$$(3.16) \quad \phi(\cdot) = \frac{L(\cdot) - F(\cdot)}{L(\cdot) - M(\cdot)} \in [0, 1]$$

¹⁷A behavioral strategy differs from a mixed strategy in that a mixed strategy determines a probability distribution over pure strategies (i.e., over mappings from information sets to actions). A behavioral strategy randomizes, perhaps degenerately, the action of the player in each state. Kuhn's (1953) theorem establishes an equivalence between these two definitions of randomization (mixed and behavioral strategy) under certain conditions. In both cases, the action sets may differ for different histories \tilde{h}_t .

¹⁸Fudenberg and Tirole (1985) and Thijssen, Huisman, and Kort (2002) provide an alternative formulation using mixed strategies in continuous time rather than behavioral strategies. These authors consider a cumulative distribution function representing the cumulative probability that firm i has invested in state \tilde{h}_t or at any previous time, and α_t^i is a "probability of atoms" — a sort of density function. The equivalence between the definitions of mixed strategy and behavioral strategy in such settings is rigorously shown by Touzi and Vieille (2002).

¹⁹Here α_t^i is the limit (as the time length h approaches zero) of the discrete-time mixed-action measures in strategic-form games. When $\alpha_t^i > 0$, the firm invests immediately and enters the market.

²⁰The equilibrium strategy profile must be adapted to the payoff-relevant history filtration. For a formal definition of MPE, see Maskin and Tirole (2001).

constitutes a unique Markov perfect equilibrium in investment strategies. The thresholds X_F^* and X_P^* are such that the following statements hold.

(i) X_F^* solves the system

$$(3.17a) \quad rF - \mathcal{D}F = 0,$$

$$(3.17b) \quad F|_{x=X_F^*} = W|_{x=X_F^*},$$

$$(3.17c) \quad F_x|_{x=X_F^*} = W_x|_{x=X_F^*},$$

$$(3.17d) \quad \lim_{x \rightarrow 0} F(x) = 0.$$

(ii) X_P^* solves the system

$$(3.18a) \quad rL - \mathcal{D}L = 0,$$

$$(3.18b) \quad L|_{x=X_P^*} = W|_{x=X_P^*},$$

$$(3.18c) \quad L(X_P^*) = F(X_P^*),$$

$$(3.18d) \quad \lim_{x \rightarrow 0} L(x) = 0.$$

PROOF. See Appendix II. □

The equilibrium strategy described in Proposition 7 is interpreted as follows. In the period prior to the preemption time τ_P^* , there is no incentive for either firm to invest, so both firms stay out ($\alpha_t^i = \alpha_t^j = 0$). In the period immediately after the preemption time τ_P^* , one firm in the industry will invest first. At the optimal preemption time τ_P^* , each firm is indifferent between being the leader and being the follower (i.e., $L(X_P^*) = F(X_P^*)$) and the probability of investment for each firm at the optimal preemption time τ_P^* is given by $\alpha_{\tau_P^*}^i = \alpha_{\tau_P^*}^j = 0$. For t greater than the follower's optimal time of investment, both firms will operate in the marketplace with the second entrant "following suit". The only possible equilibrium is characterized by a sequence in which one of the firms invests at time τ_P^* and the other at a later time $\tau_F^* > \tau_P^*$. The probability of being the leader in the duopoly market is exactly one half. Proposition 8 summarizes these two properties.

PROPOSITION 8. *The Markov perfect equilibrium established in Proposition 7 has the following properties.*

(i) *Investment takes place in sequence:*

$$X_P^* < X_F^* \quad \text{or} \quad \tau_P^* < \tau_F^*.$$

(ii) *Each (symmetric) firm has a one-half probability of being the leader in the duopoly market — that is, of entering at time τ_P^* .*

PROOF. See Appendix III. □

The discussion so far has concerned a Markov perfect equilibrium in the case of a duopoly. Bouis et al. (2009) consider a similar problem in the case of a larger number of symmetric oligopolistic firms. The authors consider a multiplicative shock that follows the geometric Brownian motion in the context of reduced-form (deterministic) profits $\pi(n)$ that decrease in the number of incumbent firms n . An equilibrium arises where all firms invest sequentially.²¹ The investment trigger of market entrants (except for the last one) is determined by "rent equalization."²² Therefore, the inclusion of more option-holding firms does not critically affect the sequential occurrence of investments.

²¹Bouis et al. (2009) formalize explicitly the three-firm case and provide numerical analysis for larger oligopolies. The authors show that simultaneous investments may also occur if the starting value of the process is large. In our model, Assumption 1 ensures that such an equilibrium does not arise.

²²The authors also demonstrate the existence of an additional effect that they call the *accordion* effect. In the three-firm case, if the threshold of the third entrant X_3^* rises, then the second investor has an incentive to invest earlier and thus enjoy duopoly rents longer, thereby setting a lower threshold X_2^* . The first investor then faces earlier entry by the

4. Social optimality of myopic market-entry sequencing

We have seen that oligopolist firms invest in sequence. Next, we examine whether the entry decisions taken by firms acting in their own interest may lead to some form of social optimality. To set a benchmark, we consider the investment timing decisions that a central planner would impose on (decentralized) firms.

Assume the existence of a given investment threshold choice $X = (X_1, \dots, X_n)$ in the social planner's admissible strategy set \mathbb{R}^n . Then the expected social surplus is given by

$$(4.1) \quad S(X) = \mathbb{E}_0 \left[\int_0^\infty e^{-rs} \Pi(X_s, m_s) ds - \sum_{i=1}^n e^{-r\tau_i} I \right],$$

where $\Pi(\cdot, \cdot) \equiv R(\cdot, \cdot) - C(\cdot, \cdot)$ is the flow of social operating surplus. We have $R(\cdot, m_s) = \sum_{i=0}^{m_s} \rho(\cdot, i)$ for the flow of consumer surplus (revenues) and $C(\cdot, m_s) = \sum_{i=0}^{m_s} c(\cdot, i)$ for the flow of total production costs. The relationship between the (net) social benefit and firms' profits is

$$(4.2) \quad \Pi(\cdot, m_s) = \sum_{i=0}^{m_s} [\rho(\cdot, i) - c(\cdot, i)] = \sum_{i=0}^{m_s} \pi(\cdot, i) = \sum_{i=0}^n a_s^i \pi(\cdot, i).$$

The social planner would choose investment thresholds to maximize the expected social surplus $S(X)$. Thus, the social planner is faced with a (stochastic) control problem of the form

$$S(X^*) \equiv \sup_{X \in \mathbb{R}^n} S(X).$$

4.1. Oligopoly. Proposition 9 asserts that myopic firms — such as those pursuing open-loop investment policies — would invest at the time(s) expected by a social planner. Leahy (1993) uses instantaneous (rather than impulse) control to establish an equivalent result in the context of (infinitely divisible) capacity investment. Here we extend that result to binary (lumpy) market-entry decisions.

PROPOSITION 9. *A Nash equilibrium in investment policies (as obtained in Proposition 5) is socially optimal. That is, $S(X^*) = \sum_{i=1}^n v_0^i(X_i^*)$.*²³

PROOF. See Appendix IV. □

4.2. Duopoly. Following the analysis in Section 3.2, we show that a social loss arises when investment options are available to a small number of firms. Proposition 10 establishes that both firms receive lower payoffs under the Markov perfect equilibrium than under the equivalent Nash equilibrium in investment policies for two players.²⁴ Proposition 11 then asserts that the first market entry takes place too early to be socially optimal.

PROPOSITION 10. *In terms of expectations, no firm is better-off in the region $[X_P^*, X_F^*]$ because the expected value of each firm (including the actual leader) is equal to the value of the follower. In other words, rents are dissipated (on average).*

PROOF. See Appendix V. □

second entrant and so enjoys monopoly profits for a shorter time period; its entry threshold is thus increased. The opposite directions of the change in the “wedges” between X_1^* and between X_2^* and X_2^* and X_3^* is at the core of the accordion effect.

²³Whether the sequence of entry is optimal for society as a whole (i.e., both firms and consumers) or for firms considered jointly is a question that cannot be answered unequivocally here. The answer depends partly on how the revenue function $R(\cdot, m_t)$ is determined. If the optimal choice of revenue is made by a social planner who is looking out for the interest of all market participants, then the investment sequence would benefit society as a whole. However, if the planner is concerned solely with the interest of firms, the resulting Pareto optimality might actually be suboptimal from a social-welfare perspective.

²⁴The two-player Nash equilibrium in investment policies described by Joaquin and Butler (2000) is derived in a setting where the exchange-rate process is a multiplicative shock and evolves as a geometric Brownian motion, the investment cost is constant over time, and profit functions are obtained in Cournot quantity competition.

If firms cannot commit to sticking to their market-entry strategies, then identical firms receive equal expected values (in equilibrium); this result is some form of “rent dissipation”. The intuition behind this result is based on the rent-equalization principle determining the preemption point: If one firm planned to enter as leader at a time t in order to receive a greater value than the follower, then the follower could increase its value by preempting and investing just before t , at $t - \epsilon$ for small ϵ . This strategic interplay would be repeated inducing firms to preempt all the way down to the point X_P^* , beyond which there is no advantage to be gained from preempting.²⁵

PROPOSITION 11. *The optimal investment times arising from the MPE are ranked as follows,*

$$\tau_P^* < \tau_L^* < \tau_F^*,$$

where $\tau_L^* = \inf \{t \geq 0 \mid X_t \geq X_L^*\}$ for $X_L^* \equiv \arg \max_{X_L \in \mathbb{R}} L_0(X_L, X_F^*)$. *The first entrant in the (duopoly) MPE enters earlier, at time τ_P^* , and faces riskier returns and a higher probability of going bankrupt than is socially optimal.*

PROOF. The second inequality follows from Proposition 4 in the two-firm case. To derive the first inequality, suppose by way of contradiction that $\tau_P^* \geq \tau_L^*$. The rent-equalization principle implies that $L(X_P^*) = F(X_P^*)$. Since the follower’s value is nondecreasing in the shock, we have $F(X_P^*) \geq F(X_L^*)$. Since $F(x) > L(x)$ for $x \in (-\infty, X_P^*]$, it follows that $F(X_L^*) > L(X_L^*)$. Therefore, $L(X_P^*) > L(X_L^*)$, which contradicts the definition of X_L^* . The uniqueness of $\arg \max_{X_L \in \mathbb{R}} L(X_L, X_F^*)$ was established in Proposition 4. \square

In short, because myopic investment policies lead to social optimality (Proposition 9), more knowledge of the competition (closed-loop approach) results in social loss.

5. Conclusion

The choice of an appropriate solution approach (open-loop versus closed-loop) is critical to determining equilibria in real options models that involve competition under uncertainty. When firms can observe and react to rivals’ actions, the appropriate solution involves closed-loop strategies and a Markov perfect equilibrium. However, deriving the MPE in closed-loop strategies is usually more involved than deriving the Nash equilibrium for the case of open-loop strategies.

If there are no entry barriers and if a large number of firms can enter a growing market (ignoring rivals’ moves), then the sequence of investments is Pareto optimal (an open-loop equilibrium). A similar result was shown previously (e.g., Leahy 1993; Baldursson 1998) in a context where firms could invest incrementally in capital stock but without due reference to the assumed information structure. We have shown that this outcome applies to lumpy market-entry investments as well. When only a few firms are protected from new market entry, the sequence of investment is not necessarily Pareto optimal (closed-loop). In this case, Markov perfection implies investment at an earlier time than the timing a social planner would impose on the firms.

For sufficiently large oligopolistic industries, the optimal market-entry strategy can be determined while ignoring potential rivals. The value obtained, however, represents the effect of competition via an additional (negative) competitive erosion term, as proved in Proposition 2. In the case of a duopoly, firms that jointly share an investment option are subject to adverse effects that alter their incentive to erect entry barriers. Restricted availability of the investment option leads to preemption and rent dissipation (on average). Once both duopoly firms enter, the knowledge that no other firm can arrive ensures fairly high rents in the marketplace (with no competitive erosion term). This trade-off calls for

²⁵The difference here from the deterministic model in which this rent dissipation was first developed (cf. Fudenberg and Tirole, 1985) is that deviation from the expected market development scenario could lead to positive rents earned by the leader in favorable states and to losses if market development falls short of expectations. This is because entry decisions are now based on *expected* market development scenarios.

concurrent consideration of both the preemption risk that arises in industries with only a few firms and of the endogenous erection of entry barriers to deter further competitive arrivals.

Our modeling approach has proved useful for explaining dynamic market entry timing decisions and strategic interactions in the case of stochastically growing markets characterized by lumpy investments under uncertainty. We have adopted a fairly general approach that covers a large number of stochastic (Itô) processes and many reduced-form profit functions.

Appendices

Appendix I: Proof of Proposition 5. We first deal with part (i). From Proposition 4, we know that player i has no incentive to select a threshold in (X_{i-1}^*, X_{i+1}^*) other than the X_i^* characterized in Proposition 3. Suppose now that firm i selects an alternative $X_i \in [X_{k-1}^*, X_k^*]$, $k \leq i$, while seeking profitable deviations from the hypothesized Nash equilibrium. Let $\tau_i \equiv \inf \{t \geq 0 \mid X_t \geq X_i\}$. We have

$$\begin{aligned} V_0^i(X_i, X_{-i}^*) = & \mathbb{E}_0 \left[\int_{\tau_i}^{\tau_k^*} e^{-rs} \pi(X_s, i) ds + \sum_{m=k}^{i-2} \int_{\tau_m^*}^{\tau_{m+1}^*} e^{-rs} \pi(X_s, m) ds \right. \\ & + \int_{\tau_{i-1}^*}^{\tau_{i+1}^*} e^{-rs} \pi(X_s, i) ds + \sum_{m=i+1}^{n-1} \int_{\tau_m^*}^{\tau_{m+1}^*} e^{-rs} \pi(X_s, m) ds \\ & \left. - e^{-r\tau_i} I \right]. \end{aligned}$$

Selecting X_i to maximize this expression leads to solving the system of equations (3.7a)–(3.7d). This corresponds to the previous definition of X_i^* . Hence, the value function $V_0(\cdot, X_{-i}^*)$ increases monotonically on the interval $[X_{k-1}^*, X_k^*]$ and reaches a maximum at X_k^* .

Similarly, suppose that firm i deviates and selects $X_i \in [X_k^*, X_{k+1}^*]$ for $k \geq i$. Then

$$\begin{aligned} V_0^i(X_i, X_{-i}^*) = & \mathbb{E}_0 \left[\int_{\tau_i}^{\tau_{k+1}^*} e^{-rs} \pi(X_s, k) ds \right. \\ & \left. + \sum_{m=k+1}^{n-1} \int_{\tau_m^*}^{\tau_{m+1}^*} e^{-rs} \pi(X_s, m) ds - e^{-r\tau_i} I \right]. \end{aligned}$$

Maximizing this program yields the system of equations (3.7a)–(3.7d), so $X_i^* = X_k^*$. For $X_i > X_i^*$, we have that $V_0^i(\cdot, X_{-i}^*)$ is strictly decreasing on each interval of the form $[X_k^*, X_{k+1}^*]$ and reaches a maximum at the left boundary τ_k^* , $k = 1, \dots, n$. Because the value function is continuous at X_i^* (owing to the value-matching conditions), X_i^* is the unique maximum. For part (ii), observe that any Nash equilibrium implies a weak sequencing of market entries. Hence, X^* and the other permutations of firm labels in $X^* = (X_1^*, \dots, X_n^*)$ are the only Nash equilibria in market-entry thresholds ($n!$ permutations).

Appendix II: Proof of Proposition 7. We need to show that no firm has an incentive to deviate from the hypothesized MPE of Proposition 7. We shall consider several cases in turn.

Case 1: $x < X_P^*$. Here $L(x) < F(x)$, so investing is a strictly dominated strategy.

Case 2: $x \in [X_P^*, X_F^*]$. In this case, a coordination problem may arise. The characterization of α^i as the limit (as the time interval approaches zero) of discrete-time mixed action facilitates depiction of the problem in strategic form, as shown in Figure 1. The value of pursuing the mixed action $\alpha_t^i = \alpha^i(x)$ (while the rival invests with probability $\alpha_t^j = \alpha^j(x)$) is

$$\begin{aligned} V^i(\alpha^i, \alpha^j) = & \alpha^i \alpha^j M^i + (1 - \alpha^i) \alpha^j F^i + \alpha^i (1 - \alpha^i) L^i \\ & + (1 - \alpha^i) (1 - \alpha^j) V^i(\alpha^i, \alpha^j), \end{aligned}$$

resulting in

$$V^i(\alpha^i, \alpha^j) = \frac{\alpha^i \alpha^j M^i + (1 - \alpha^i) \alpha^j F^i + \alpha^i (1 - \alpha^j) L^i}{\alpha^i + \alpha^j - \alpha^i \alpha^j}$$

for $(\alpha^i, \alpha^j) \neq (0, 0)$. Since this expression is concave in α_t^i (i.e., $\frac{\partial^2 V_t^i}{\partial \alpha_t^{i2}}(\alpha_t^i, \alpha_t^j) < 0$), the following first-order condition is both sufficient and necessary for an optimal mixed action to obtain:

$$\frac{\partial V^i}{\partial \alpha^i}(\alpha^i, \alpha^j) = 0 \iff \alpha^i(x) = \phi^i(x),$$

where

$$\phi^i(\cdot) = \frac{L^j(\cdot) - F^j(\cdot)}{L^j(\cdot) - M^j(\cdot)}.$$

For identical firms following symmetric behavioral strategies ($\alpha_t^i = \alpha_t^j = \alpha_t$), the preceding equality simplifies to

$$\phi^i(\cdot) = \frac{L(\cdot) - F(\cdot)}{L(\cdot) - M(\cdot)}.$$

	Invest	Wait
Invest	$M_t^j(\cdot)$	$F^j(\cdot)$
Wait	$L_t^j(\cdot)$	$V^j(\cdot)$
	$M^i(\cdot)$	$L^i(\cdot)$
	$F^i(\cdot)$	$V^i(\cdot)$

FIGURE 1. Strategic-form representation of the coordination problem in (X_P^*, X_F^*)

It then follows from equations (3.14) and (3.15) that

$$\begin{aligned} M(X_t) &= \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} \pi(X_s, 2) ds - I \right] \\ &\leq \sup_{X_F \in \mathbb{R}} \mathbb{E}_t \left[\int_{\tau_F}^\infty e^{-r(s-t)} \pi(X_s, 2) ds - e^{-r(\tau_F-t)} I \right] \\ &\leq F(X_t). \end{aligned}$$

Case 3: $x \geq X_F^$.* Here the firm has a dominant strategy to invest immediately and receive $M(x)$. For symmetric firms, the value of the leader and of the follower are equal(ized) for process values higher than X_F^* : $L(x) = M(x) = F(x)$.²⁶ Firm i can do no better than to pursue the prescribed strategy. Firm i is faced with a decision-theoretic problem. Its optimal entry time is jointly determined by the HJB equation (3.17a), the “smooth pasting” condition (3.17c), and equation (3.17d). In this case, the strategy formulated in Proposition 3 is dominant.

Together these results indicate that the strategy profile given in Proposition 7 is a Markov perfect equilibrium whose uniqueness is shown by Fudenberg and Tirole (1985, Apx. A) in the deterministic context.

Appendix III: Proof of Proposition 8. We suppose the contrary and then derive a contradiction. So let $X_P^* \geq X_F^*$. For $X_P^* > X_F^*$, since $F(X_P^*, \cdot)$ is nondecreasing in X_F it follows that $F(X_P^*, X_P^*) \geq F(X_P^*, X_F^*)$, which means that $F(X_P^*, \cdot)$ is maximized for two distinct trigger values, X_F^* and X_P^* . This contradicts the uniqueness of the threshold as established by conditions (3.7a)–(3.7d). If $X_P^* = X_F^*$, then both firms enter at time τ_P^* and receive $M(X_P^*)$. Hence there is an investment trigger $X_i \in (-\infty, X_P^*)$ such that $L(X_i, X_P^*) > M(X_P^*, X_P^*)$ as $\pi(X_s, 1) > \pi(X_s, 2)$. Since firms are identical, $M(X_P^*, X_P^*) = F(X_P^*, X_P^*)$. As a result, there is an incentive to invest before the preemption point X_P^* but this contradicts the rent-equalization principle of equation (3.18c). Thus, we have established part (i) of the proposition.

²⁶In new-market models, $L(x) = F(x) = M(x)$ for all $x \geq X_F^*$.

For part (ii), observe that the probability that one of the firms ends up being the leader the first time X_t is in (X_P^*, X_F^*) is given by $p_L^i = \alpha_t^i (1 - \alpha_t^j) + (1 - \alpha_t^i) (1 - \alpha_t^j) p_L^i$; therefore, $p_L^i = \frac{\alpha_t^i (1 - \alpha_t^j)}{\alpha_t^i + \alpha_t^j - \alpha_t^i \alpha_t^j}$ for $\alpha_t^i, \alpha_t^j \neq 0$. Then, in the symmetric case, we have

$$(5.1) \quad p_L = p_L^i = p_L^j = \frac{1 - \alpha_t}{2 - \alpha_t},$$

which admits a right limit at $\alpha_t = 0$. Part (ii) now follows from equations (3.16) and (3.18c).

Appendix IV: Proof of Proposition 9. By equation (4.2), we have

$$(5.2) \quad \begin{aligned} \sum_{i=0}^n \int_{\tau_i}^{\infty} e^{-rs} \pi(\cdot, m_s) ds &= \sum_{i=0}^n \int_0^{\infty} e^{-rs} a_s^i \pi(\cdot, m_s) ds \\ &= \int_0^{\infty} e^{-rs} \left[\sum_{i=0}^n a_s^i \pi(\cdot, m_s) \right] ds \\ &= \int_0^{\infty} e^{-rs} \Pi(\cdot, m_s) ds \end{aligned}$$

because $\Pi(\cdot, 0) = 0$. It now follows from Fubini's theorem and equations (4.1), (5.2), and (3.4) that

$$\begin{aligned} S_0(X) &= \mathbb{E}_0 \left[\int_0^{\infty} e^{-rs} \Pi(X_s, m_s) ds - \sum_{i=0}^n e^{-r\tau_i} I \right] \\ &= \sum_{i=0}^n \mathbb{E}_0 \left[\int_{\tau_i}^{\infty} e^{-rs} \pi(X_s, m_s) ds - e^{-r\tau_i} I \right] \\ &= \sum_{i=0}^n v_0^i(X_i). \end{aligned}$$

The proposition is obtained by taking the supremum X^* over \mathbb{R}^n .

Appendix V: Proof of Proposition 10. The proof proceeds in two steps: first, we derive the probability of certain scenarios; then, we calculate the firm's expected value in the considered region as the weighted average of values obtained in each scenario. Here we examine the symmetric case.

The probability that a firm becomes the leader in the region $[X_P^*, X_F^*]$ was obtained in (5.1) as $p_L = \frac{1 - \alpha_t}{2 - \alpha_t}$. The probability of simultaneous investment, should the process X be located for the first time in the preemption region $[X_P^*, X_F^*]$, is equal to $p_C = \frac{\alpha_t^i}{2 - \alpha_t^i}$. Hence, the value in this region is

$$\begin{aligned} V^i(\alpha^{i*}, \alpha^{j*}) &= p_L(L + F) + p_C M \\ &= \frac{L + (1 - \alpha^*)F - \alpha^*(L - M)}{2 - \alpha^*} \\ &= \frac{L + (1 - \alpha^*)F - L + F}{2 - \alpha^*} = F, \end{aligned}$$

where the second equality follows from the probabilities just described and the third equality from equation (3.16).

Preemptive Capacity Investment under Uncertainty

ABSTRACT. The incentive to “overinvest” in capital may be eroded in dynamic, competitive settings if firms face uncertainty and irreversibility. In this paper, we derive the stationary Markov perfect equilibrium for a dynamic, infinite-horizon capacity investment game formulated in continuous time in which reduced-form profits are subject to industry shocks. We show that the theory of marginal Tobin’s q still holds in this setting if the strategic externalities of rivals’ investment are properly accounted for.

1. Introduction

It is standard practice in finance and strategy to interpret real investment opportunities as being analogous to options. This view is well accepted among academics and practitioners alike and is the core of real options analysis (ROA), as summarized in the works of Dixit and Pindyck (1994) and Trigeorgis (1996). ROA has some connection with the literature on marginal Tobin’s q (see, e.g., Abel, 1983; Abel and Eberly, 1994). Following the latter approach, marginal firm value fluctuates along with industry shocks as long as the marginal value stays within a certain range of profitability characterized by lower and upper barriers. When the value process hits one of these boundaries, the firm either invests or divests, so that the marginal value is reflected off the boundaries.

Since standard ROA draws from contingent-claims analysis based on financial options, a key feature specific to real assets is often ignored — namely, the interplay or strategic interactions taking place among (real) option holders. Bridging this gap requires the use of additional notions borrowed from game theory. Chevalier-Roignant and Trigeorgis (2011) develop further the notion of “option games,” while Chevalier-Roignant et al. (2011) review the relevant literature. We deal here with a subset of this literature where firms adjust their capital stocks or production capacity in the face of uncertainty and competition (see, e.g., Leahy, 1993; Baldursson, 1998; Baldursson and Karatzas, 1996; Grenadier, 2002).

In dynamic problems, one often distinguishes several types of controls or strategies. *Open-loop* or precommitment strategies set ex ante a certain investment path to be pursued regardless of the state realization. In case of *closed-loop* strategies, the decision makers can condition their actions on the entire state paths. Inbetween, decision makers can employ *Markov or feedback* strategies, whereby the current state value is considered a sufficient statistic for decision making. This latter type of control is typically employed in standard ROA which builds upon dynamic programming. Such notions originating from control theory also find application in dynamic game theory, as discussed in Başar and Olsder (1999) and Fudenberg and Tirole (2002). Capital accumulation games — studied in continuous time by, e.g., Spence (1979), Fudenberg and Tirole (1983) and Reynolds (1987) — hinge on the distinction between these different notions of strategies. These models are typically built as differential games that assume a deterministically growing market. Spence (1979) cautions:

“Uncertainty may alter the qualitative conclusions drawn here, even if firms are not risk-averse. Because of the irreversibility, uncertainty acts in part like a further constraint on, or cost of, growth. This aspect of the subject is complicated and in need of further theoretical work as well.”

To the best of our knowledge, this research gap has not been bridged yet. The “option games” literature dealing with investment in capital stocks has typically overlooked the inclusion of the capital stock held by rivals in the characterization of the investment strategy, in effect employing strategies

that are open-loop with respect to the industry state (see Leahy, 1993; Baldursson and Karatzas, 1996; Baldursson, 1998).¹ This approach may be ill-advised because Nash equilibria in open-loop strategies often fail to be subgame perfect (see Fudenberg and Tirole, 1991, section 13.4). An emerging strand in the literature turns to closed-loop investment strategies instead. Back and Paulsen (2009) show that the Nash equilibrium obtained in Grenadier (2002) fails subgame perfection and discuss the difficulties in formulating dynamic capacity-expansion problems in closed-loop strategies. Novy-Marx (2009) derives closed-loop equilibrium strategies for duopolistic capacity competition under more restrictive assumptions (than we do). This paper is an attempt to bridge this research gap. We model market uncertainty via Itô processes as standard in ROA and supplement the analysis with techniques from the literature on stochastic differential games. We particularly focus on Markov capital-adjustment strategies and characterize the Markov perfect equilibria. This allows us to quantify the reaction to rivals' strategic capacity adjustment decisions and modify the investment rule on marginal Tobin's q to account for negative strategic externalities on the firm's marginal value.

The paper is organized as follows. Section 2 presents the model assumptions. Section 3 discusses the appropriate choice of strategy space and proposes the use of Markov strategies. Section 4 formulates the problem with Markov controls and derives the Markov perfect equilibria. Section 5 concludes.

2. Model primitives and assumptions

Consider a complete filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t \geq 0}$ is a family of "tribes" that satisfies the usual conditions: $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s < t$, $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ and the \mathbb{P} -null set belongs to \mathcal{F}_0 . The tribe \mathcal{F}_t denotes the information available to decision makers at time t . The market is subject to an exogenous shock $X = \{X_t\}_{t \geq 0}$ modeled as a diffusion. To preclude foresight about future economic developments, we assume that X is \mathbb{F} -adapted, namely that X_t is \mathcal{F}_t -measurable for all $t \in \mathbb{R}_+$. The process X starts at $X_0 = x \in \mathcal{X}$ almost surely, with $\mathcal{X} = \mathbb{R}$ for simplicity; X is the unique (strong) solution to the stochastic differential equation (SDE),

$$(2.1) \quad dX_t = \mu(X_t) dt + \sigma(X_t) dB_t,$$

where $B = \{B_t\}_{t \geq 0}$ is a standard Brownian motion. For existence, we assume that $\{\mu(X_t)\}_{t \geq 0}$ and $\{\sigma(X_t)\}_{t \geq 0}$ are \mathbb{F} -adapted with finite variations and that $\mu, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are Lipschitz continuous and satisfy the growth condition. This SDE covers a fairly broad family of stochastic processes used in economic analysis, such as arithmetic Brownian motion, geometric Brownian motion, the Ornstein–Uhlenbeck mean-reverting process and the geometric Ornstein–Uhlenbeck process.

We consider a symmetric duopoly, where at time t on path $\omega \in \Omega$, firm i (j) has capital stock $K_t^i(\omega)$ ($K_t^j(\omega)$).² Firm i (j)'s initial capital stock is $K_{0-}^i = k^i$ ($K_{0-}^j = k^j$) almost surely.³ The vector of initial industry capital stocks is $k = (k^i, k^j)^\top$, where \top denotes the transpose. The realization of the capital stocks at time t is given by the two-dimensional vector $K_t(\omega) = \left(K_t^i(\omega), K_t^j(\omega) \right)^\top$.

At any instant t , firm i can adjust its capital stock, i.e., acquire new capital or sell unnecessary units. Assume firms face a symmetric, strictly convex capital-adjustment cost function $\kappa : \mathbb{R} \rightarrow \mathbb{R}$. The function $\kappa(\cdot)$ is differentiable almost everywhere except at $I = 0$. We denote the (partial) derivative of function f with respect to argument y by $f_y(\cdot)$. $\kappa_I(\cdot)$ is the first-order derivative of the capital-adjustment cost function $\kappa(\cdot)$. Denote the left (resp., right) limit of $\kappa_I(\cdot)$ by $\underline{\kappa}_I$ (resp., $\bar{\kappa}_I$). Besides, $\kappa(0) = 0$. An

¹An exception is most lump-sum investment models (see, e.g., Smets, 1991; Grenadier, 1996; Huisman and Kort, 1999) that do not suffer from the same weakness as they build on Fudenberg and Tirole's (1985) closed-loop approach.

²To reduce mathematical complexity, we here consider for each firm a one-dimensional capital stock. Considering multiple resources would require to examine whether the resources are complements or substitutes in addition to the factors analyzed herein.

³We distinguish between time $0-$ and 0 because a lump-sum investment usually occurs at time 0 immediately adjusting the capital stock held.

example of such an adjustment cost function is the often-used kinked, piecewise linear cost function,

$$(2.2) \quad \kappa(I) = \underline{\kappa}I^- + \bar{\kappa}I^+,$$

with $(0 \leq) \underline{\kappa} < \bar{\kappa}$ (to preclude arbitrage opportunities) and where, by definition, $I^+ \triangleq \max\{0, I\}$ and $I^- \triangleq \max\{0, -I\}$. The quadratic investment cost function is

$$(2.3) \quad \kappa(I) = aI + \frac{b}{2}I^2,$$

with $a, b \in \mathbb{R}_+$. Function (2.3) is differentiable at $I = 0$, while function (2.2) is not, having left-limit $\kappa_I = \underline{\kappa}$ and right-limit $\bar{\kappa}_I = \bar{\kappa}$.

Following Baldursson (1998), an investment strategy is a capital-accumulation process ξ^i with decomposition $\xi^i \triangleq \xi^{i+} - \xi^{i-}$ where the negative and the positive parts of the process, ξ^{i-} and ξ^{i+} , respectively, are right-continuous with left-limits, nondecreasing, and \mathbb{F} -adapted.⁴ Besides, $\xi^{i+}(0-) = \xi^{i-}(0-) = 0$ almost surely. The process ξ^i has (pathwise) bounded variation as the difference between two nondecreasing processes. Following investment strategy ξ^i , firm i 's capital stock at time t on path $\omega \in \Omega$ is

$$(2.4) \quad K_t^i(\omega) = k^i + \xi_t^{i+}(\omega) - \xi_t^{i-}(\omega).$$

An instantaneous change in capital stock is denoted $d\xi_t(\omega) = \xi_t^i(\omega) - \xi_{t-}^i(\omega)$. $d\xi_t(\omega)$ is a Lebesgue-Stieltjes differential and cannot properly be interpreted in terms of investment rate since it admits no time derivative. Similarly to Leahy (1993), the strategy choice has no impact on the exogenous state dynamics given in equation (2.1).

If it were possible to determine an “investment rate,” we could define — as in capital-accumulation (differential) games — a control policy prescribing for firm i an investment rate I_t^i at time $t \in \mathbb{R}_+$. An instantaneous change at time t in firm i 's capital stock is then I_t^i , with the capital stock evolving according to the ordinary differential equation:

$$(2.5) \quad \dot{K}_t^i = I_t^i.$$

The industry-wide shock $X = \{X_t\}_{t \geq 0}$ follows the Itô process of equation (2.1), while the capital stocks evolve according to equation (2.4). At each instant t , firm i receives an unknown revenue flow $\pi^i(X_t(\omega), K_t(\omega))$ depending on the shock realization $X_t(\omega)$ and the prevalent capital stocks of the rivals, $K_t^i(\omega)$ and $K_t^j(\omega)$. The profit function π^i is strictly increasing and twice continuously differentiable in x ; it is continuously differentiable, strictly decreasing in k^j , and concave in k^i . We assume that the marginal profit is increasing in the shock ($\pi_{k^i x}^i > 0$); this ensures that the optimal strategy can be expressed in terms of divestment/ investment triggers.⁵ Also assume that the profit function is submodular, with $\pi_{k^i k^j} < 0$. Although not necessary, we can specialize the reduced-form profit function. Consider, for instance, that the price $P_t(\omega)$ that clears the market is given by $P_t(\omega) = D(X_t(\omega), K_t^i(\omega) + K_t^j(\omega))$, where the demand function $D(\cdot, \cdot)$ is strictly increasing and twice continuously differentiable in x and strictly decreasing in k^i and k^j . For negligible production costs, firm i 's instantaneous profit is $k^i D(x, k^i + k^j) - \kappa(I)$, which is increasing and twice continuously differentiable in x .

Firms are assumed risk-neutral, facing a common, constant risk-free rate r .⁶ The *payoff* to firm i if the strategy profile (ξ^i, ξ^j) is followed is

$$(2.6) \quad J^i(x, k \mid \xi^i, \xi^j) \triangleq \mathbb{E} \left[\int_0^\infty e^{-rt} \pi(X_t, K_t) dt - \int_0^\infty e^{-rt} \kappa(d\xi_t^i \mid x, k) \right],$$

⁴Back and Paulsen (2009), who consider capital expansion only, define an “open-loop strategy” in a similar way as a nondecreasing \mathbb{F} -adapted process ξ^{i+} .

⁵We also impose Inada conditions on the profit function: $\lim_{k^i \rightarrow 0} \pi_{k^i}^i(x, k) = \infty$ and $\lim_{k^i \rightarrow \infty} \pi_{k^i}^i(x, k) = 0$.

⁶The assumption of risk neutrality can be relaxed if arbitrage opportunities in the market do not exist and the market is complete (see Cox and Ross, 1976; Harrison and Kreps, 1979). Introducing capital depreciation in the model would not add much flavor; it only increases the discount factor by the depreciation rate.

where the second integral is understood as a (pathwise) Lebesgue-Stieljes integral. The firm adjusts its capital stock with a view to maximizing expected payoff. If a Nash equilibrium $\xi^* = (\xi^{i*}, \xi^{j*})$ exists, the *value* received by firm i in state (x, k) is

$$(2.7) \quad V^i(x, k) \triangleq \max_{\xi^i} J^i(x, k \mid \xi^i, \xi^{j*}).$$

Given the model assumptions, V^i is strictly increasing, twice continuously differentiable in the shock. Besides, $V_{xk^i}^i > 0$ and $V_{k^i k^j}^i < 0$.

In the above definition of strategy space, it is not clear how the arrival of new information about the rival's realized capital stock influences the capital adjustment strategy. This point is, however, essential in pinning down the appropriate strategy space, as discussed below.

3. Strategy-space choice and solution approach

In dynamic problems, the choice of the strategy space is closely related to the hypothesized information structure. Two information structures are common in the deterministic control and differential games literature. Open-loop or precommitment strategies prescribe a certain investment rate at each time t of the control horizon. Markov strategies enable the decision maker to condition her decision on the latest value of the state. In deterministic single-agent control problems, this distinction is not material since these assumptions lead to different representations of the same optimal policy (see Başar and Olsder, 1999, 5.6). When competition or uncertainty are involved, the different types of strategies may lead to different outcomes. The definition of ξ^i above is not precise enough to describe the hypothesized information set of player i at each time t . A crucial element in defining the strategy space relates to the behavioral conjectures on how firm j 's capital-stock adjustments feed back into firm i 's investment or divestment choices. Fudenberg and Tirole (1985) face a related problem in a deterministic case of multiplayer optimal investment timing: a player's strategy defined by a nondecreasing, right-continuous, $[0, 1]$ -valued function (interpreted as the cumulative probability of investing) is not rich enough to accommodate consideration of dynamic strategic interactions over the play of the game. One should thus consider as part of a player's strategy an additional "atoms function" that prescribes a certain intensity of investment at each time t (on path $\omega \in \Omega$ in the stochastic case).

The literature on incremental capital adjustment under uncertainty and competition fails to describe the assumed atoms function, relying in effect on the notion of a "myopic firm." A *myopic* firm essentially assumes that there will be no further changes in the industry capital stock other than those arising from its own investment or divestment decisions. Its value is $J^i(x, k \mid \xi^i, 0)$. Myopic strategies are Markovian in the exogenous state x but not in rival firm j 's capital stock. If a firm invests according to such a strategy, it either cannot observe the rivals' investment behavior or it chooses to precommit to a myopic investment path as a function of the actual exogenous state realization.

Grenadier (2002, proposition 3) and Back and Paulsen (2009, equation 8) appear to make an attempt to circumvent the notion of the myopic firm. There, firm i conjectures that, on the equilibrium path, the rival's capital stock is a deterministic function of firm i 's capital stock, $K_t^j(\omega) = h(K_t^i(\omega))$. In the n -firm case, they explicitly consider $K_t^{-i}(\omega) = (n-1)K_t^i(\omega)$. Then, what seems to be an investment trigger which depends on the realization of rival's capital stock $K_t^{-i}(\omega)$ is actually an investment trigger which depends on $K_t^i(\omega)$ solely. In other words, this firm assumes a parallel development of rival's capital stock with no possibility for the rival to deviate from the capital-stock development path conjectured by the myopic firm. This statement is not a derived equilibrium outcome but simply a modeling assumption. The actual capital stock of firm i 's rival does not feed back into firm i 's investment strategy. Owing to this behavioral assumption on firm i , the Nash equilibrium in which firms follow the latter type of strategy turns out to correspond to the "myopic equilibrium" whereby firm i explicitly ignores rival's capital-stock adjustments (Grenadier, 2002, section 2). As is standard in real options analysis, the above investment strategies are Markovian in the exogenous state. They nevertheless fail to allow the rivals'

actual capital stock to feed back into the firm's strategy. The literature on stochastic differential games provides mechanisms to allow states driven by the actions of rivals to influence a firm's own decisions. In the next section, we employ such methods.

4. Markov perfect equilibrium

Consider that all states are perfectly observable, including the level of capital stock held by the rival firm. Introducing an atoms function makes it possible to account for the rival's capacity adjustments through the dependence on the states x , k^i , and k^j . In this case, the history of the game matters only through the present consequences of all past decisions, as reflected in the state vector (x, k) . Because both the payoffs and the strategies depend on the current state realization, we call this a "state-space" game. The Nash equilibrium in Markov strategies is a Markov Nash Equilibrium.⁷ A Markov Nash equilibrium is perfect if it is a Markov Nash equilibrium for all state vectors (x, k^i, k^j) , even those off the equilibrium path. Setting $k^j = h(k^i)$ fails to ensure a Nash equilibrium for values of k^j off the path $h(k^i)$. Formally, a Markov Nash equilibrium profile is a Markov perfect equilibrium (MPE) if, for all states (x, k^i, k^j) , it holds that

$$V^i(x, k) = \max_{\xi^i} J^i(x, k | \xi^i, \xi^{j*}),$$

whereby the investment intensity $\xi_t^i - \xi_{t-}^i \neq 0$ is given by $\phi^i(X_t, K_t)$. Although it is not mathematically rigorous to define investment rate in the present setting, we can nevertheless use the heuristics in Abel and Eberly (1994): the change in capital stock is interpreted as analogous to an investment rate I , allowing the derivation of a Hamilton-Jacobi-Bellman (HJB) equation. We next derive sufficient conditions for existence of a Markov perfect equilibrium.

PROPOSITION 12. *Firm i 's value in state (x, k) , $V^i(x, k)$, is the solution of the HJB equation*

$$(4.1) \quad rV^i(x, k) = \max_{I^i \in \mathbb{R}} \{ \pi^i(x, k) - \kappa(I^i) + V_{k^i}^i(x, k) I^i + V_{k^j}^i(x, k) \phi^j(x, k) + \mathcal{D}_x V^i(x, k) \},$$

where $\mathcal{D}_x V^i(x, k)$ indicates an instantaneous value change for an infinitesimal change in the shock with the operator \mathcal{D}_x given by

$$\mathcal{D}_x = \mu(x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma(x)^2 \frac{\partial^2}{\partial x^2}.$$

PROOF. Bellman's principle of optimality builds upon the recursive nature of the payoff function in equation (2.6). For a short time interval of length t , the continuation payoff is

$$\begin{aligned} J^i(x, k | \xi^i, \xi^{j*}) &\approx \{ \pi^i(x, k) - \kappa(I) \} t + \mathbb{E} \left[\mathbb{E} \left[\int_t^\infty \{ \pi^i(X_s, K_s) - \kappa(I) \} ds \mid X_t(\omega), K_t(\omega) \right] \mid x, k \right] \\ &\approx \{ \pi^i(x, k) - \kappa(I) \} t + \mathbb{E} [J^i(X_t(\omega), K_t(\omega) | \xi^i, \xi^{j*}) \mid x, k] \end{aligned}$$

By taking the supremum,

$$\begin{aligned} V^i(x, k) &\approx \max_{I^i \in \mathbb{R}} \left\{ [\pi^i(x, k) - \kappa(I)] t + \max_{\xi^i} \mathbb{E} [J^i(X_t(\omega), K_t(\omega) | \xi^i, \xi^{j*}) \mid x, k] \right\} \\ &\approx \max_{I^i \in \mathbb{R}} \left\{ [\pi^i(x, k) - \kappa(I^i)] t + \mathbb{E} [V^i(X_t(\omega), K_t(\omega)) \mid x, k] \right\} \end{aligned}$$

Dividing the above equation by $t (> 0)$ and taking $t \rightarrow 0$, equation (4.1) obtains, whereby V^i is presumed twice continuously differentiable in x and once continuously differentiable in k^i and k^j with

$$\lim_{t \downarrow 0} \frac{\mathbb{E} [V^i(X_t(\omega), K_t(\omega)) \mid x, k] - V^i(x, k)}{t} = V_{k^i}^i(x, k) I^i + V_{k^j}^i(x, k) \phi^j(x, k) + \mathcal{D}_x V^i(x, k).$$

□

⁷Dynamic games of complete information are generally solved by means of the perfect Nash equilibrium solution concept. We adapt the usual definition of perfection to a state-space game by identifying information sets in the game with points in the state space.

It is desirable to give more intuition (via the stochastic maximum principle) concerning the interpretation of HJB equation (4.1) and further characterize the induced optimal behavior. Consider a vector of adjoint variables for firm i , $\lambda^i = (\lambda^{ii}, \lambda^{ij})^\top$. Define the deterministic Hamiltonian

$$(4.2) \quad H^i(x, k, I^i, I^j, \lambda^i) \triangleq \pi^i(x, k) - \kappa(I^i) + \lambda^{ii}I^i + \lambda^{ij}I^j.$$

Using the notation of equation (4.2) in HJB equation (4.1), player i 's value function V^i solves for each $(x, t) \in \mathcal{X} \times \mathbb{R}_+$ the HJB equation

$$(4.3) \quad rV^i(x, k) = \max_{I^i \in \mathbb{R}} \{H^i(x, k, I^i, \phi^j(x, k), V_k^i(x, k)) + \mathcal{D}_x V^i(x, k)\},$$

where $V_k^i = (V_{k^i}^i, V_{k^j}^i)^\top$ is the gradient of firm i 's value function with respect to the vector of capital stocks. In other words, the HJB equation (4.1) can be understood as a static optimization problem, whereby firm i is not only looking at maximizing its current cash flow $\pi^i(x, k) - \kappa(I^i)$ but considers the future, indirect value contribution of its current investment via the shadow price vector $\lambda^i = V_k^i$. Here, the rival's reaction to changing industry capital stock $k^i \mapsto \phi^j(x, k)$ with $k = (k^i, k^j)^\top$ is explicitly considered when determining the optimal investment intensity $\phi^i(x, k) \in \mathbb{R}$. Player i cares about how his opponent will react to a change in its own capital stock.

Owing to the second-order term in the HJB equation, it is generally difficult to find closed-form solutions for the value functions V^i , unless certain "smoothness" restrictions are imposed. In the present stochastic case with two players, a system of coupled second-order partial differential equations obtains, making finding a closed-form solution even more challenging. To obtain a Markov perfect equilibrium, we must be able at each (x, k) to find a pair $(\phi^i(x, k), \phi^j(x, k))$ that forms a Markov Nash equilibrium in the static, continuous two-player game with values satisfying (4.3).

To enable characterizing the equilibrium further, denote the marginal value of firm i to a change in its own capital stock by $v^i(x, k) \triangleq V_{k^i}^i(x, k)$. At places, for expositional purposes, we drop the dependence of the shadow price v^i on the state (x, k) when no confusion may arise. By isolating terms depending on I^i in the HJB equation (4.1), the maximum condition simplifies to the static optimization problem

$$(4.4) \quad f(v^i) \triangleq \max_{I^i \in \mathbb{R}} \{v^i I^i - \kappa(I^i)\} \quad \text{or} \quad \phi^i(x, k) = \arg \max_{I^i \in \mathbb{R}} \{v^i I^i - \kappa(I^i)\}.$$

The problem in equation (4.4) is strictly concave in I^i and $\kappa(\cdot)$ is differentiable everywhere on $(-\infty, 0)$ and $(0, \infty)$. In these regions, a set of first-order conditions is sufficient to determine the optimal investment rule. When firm i invests or divests ($\phi^i(x, k) \neq 0$), the MNE strategy in state (x, k) prescribes that

$$(4.5) \quad v^i = (\kappa_I \circ \phi^i)(x, k),$$

with the symbol "o" indicating compound functions. In other words, when firm i divests or invests, the Markov Nash equilibrium strategy prescribes to invest up to the level where the marginal cost of adding capital, $(\kappa_I \circ \phi^i)(x, k)$, equals the marginal value of capital, $v^i(x, k)$. This investment policy is analogous to the theory prescription on marginal Tobin's q with the additional consideration of competition in the market. By the chain rule, it obtains that

$$v_x^i = (\kappa_{II} \circ \phi^i)(x, k) \times \phi_x^i(x, k).$$

For a sufficiently smooth function V^i , we have $V_{xk^i}^i = V_{k^i x}^i$, with $V_{xk^i}^i > 0$ (by assumption). Since $\kappa(\cdot)$ is strictly convex ($\kappa_{II} > 0$) and $v_x^i > 0$, it obtains that $\phi_x^i > 0$ or firm i invests more for improved economic conditions or vice versa, all other things remaining equal. We readily note an unexpected aspect of competition on the firms' reactions to economic shocks. Define

$$R^i(x, k, I^j) \triangleq \arg \max_{I^i \in \mathbb{R}} \{H^i(x, k, I^i, I^j, V_k^i(x, k)) + \mathcal{D}_x V^i(x, k)\}.$$

$R^i(x, k, \cdot)$ is firm i 's reaction function to firm j 's investment intensity. In the Nash equilibrium of the static game, $\phi^i(x, k) = R^i(x, k, \phi^j(x, k))$. We wish to capture the influence of the shock on the investment

intensity of firm i . By total differentiation,

$$\frac{d\phi^i}{dx}(x, k) = \frac{\partial R^i}{\partial x} + \frac{\partial R^i}{\partial I^j} \frac{\partial \phi^j}{\partial x}(x, k)$$

The second right-hand term captures the strategic effect in reaction to a shock. Since $\frac{\partial R^i}{\partial I^j} < 0$ (strategic substitute) and $\frac{\partial \phi^j}{\partial x}(x, k) > 0$, the intensity of investment in state (x, k) is decreased for firm i compared to the myopic situation. The presence of a rival decreases the intensity with which a firm increases (resp., decreases) its capital stock in response to favorable (resp., unfavorable) industry shocks. This effect, arising under competition, creates a natural hedge for the firm. By applying the chain rule to equation (4.5), it obtains

$$(4.6a) \quad v_{k^i}^i(x, k) = (\kappa_{II} \circ \phi^i)(x, k) \times \phi_{k^i}^i(x, k),$$

$$(4.6b) \quad v_{k^j}^i(x, k) = (\kappa_{II} \circ \phi^i)(x, k) \times \phi_{k^j}^i(x, k)$$

Since V^i is concave in k^i , $\phi_{k^i}^i < 0$ follows from equation (4.6a), so that the firm is less responsive to a favorable shock when it already has a large capital stock. By submodularity of the profit function ($\pi_{k^i k^j}^i < 0$), we have $v_{k^j}^i < 0$. Thus from equation (4.6b), $\phi_{k^j}^i(x, \cdot) < 0$ or the investment intensity of firm i decreases in the capital stock of its rival, k^j . The firm that has the larger capital stock thus can discipline the smaller firm to refrain from investing. In other words, if a firm increases its capital stock, rivals react to an increased industry capital stock by reducing their own capital stocks. This preemptive effect is not captured when one considers myopic strategies.⁸

It is suitable to partition the state space in terms of hysteresis (wait), investment, and divestment regions. Since $\kappa(\cdot)$ is strictly convex, $\kappa_I(I) < \underline{\kappa}_I$ for all $I \in (-\infty, 0)$ and $\bar{\kappa}_I < \kappa_I(I)$ for all $I \in (0, \infty)$. The regions (x, k) where divestment or investment occurs are such that $v^i(x, k) \in (-\infty, \underline{\kappa}_I]$ and $v^i(x, k) \in [\bar{\kappa}_I, \infty)$, respectively. The optimal investment rule prescribes not to invest, $\phi(x, k^i) = 0$, if $v^i(x, k) \in (\underline{\kappa}_I, \bar{\kappa}_I)$. The investment policy thus reads:

$$(4.7) \quad \phi^i(x, k) = \begin{cases} < 0 & \text{if } v^i(x, k) \in (-\infty, \underline{\kappa}_I], \\ = 0 & \text{if } v^i(x, k) \in (\underline{\kappa}_I, \bar{\kappa}_I), \\ > 0 & \text{if } v^i(x, k) \in [\bar{\kappa}_I, \infty). \end{cases}$$

If $\kappa(\cdot)$ is differentiable at $I = 0$ as in the quadratic case of equation (2.3), then $\underline{\kappa}_I = \bar{\kappa}_I$ and $(\underline{\kappa}_I, \bar{\kappa}_I)$ is empty. If $\kappa(\cdot)$ is not differentiable at $I = 0$ — as in the case of the kinked, linear adjustment cost function in equation (2.2) — with $\underline{\kappa}_I < \bar{\kappa}_I$ (by strict convexity of the cost function), $(\underline{\kappa}_I, \bar{\kappa}_I)$ is nonempty.

Following the marginal Tobin's q approach, the investment policy characterized in equation (4.7) is driven by the marginal value from investing and the marginal cost incurred by the firm upon investing. The firm “sells” capital (divests) to prevent the marginal revenue product of capital v^i from falling below a certain lower profitability level $\underline{\kappa}_I$ and acquires capital (invests) to prevent v^i from rising above an upper profitability level $\bar{\kappa}_I$. If the open interval between these thresholds $(\underline{\kappa}_I, \bar{\kappa}_I)$ is nonempty — as in the case of equation (2.2) — the set of states (x, k) whereby $v^i(x, k) \in (\underline{\kappa}_I, \bar{\kappa}_I)$ is an inaction or *hysteresis* region where the firm neither invests nor divests (“stays put” or waits). Within the bands $\underline{\kappa}_I$ and $\bar{\kappa}_I$, the shadow price process is driven solely by the exogenous shock (i.e., no control is exercised). At the upper barrier $\bar{\kappa}_I$, firms will be indifferent between staying put or expanding capacity, while at the lower barrier $\underline{\kappa}_I$, firms will be indifferent between holding the current level of capital stock constant or downsizing. Expansion at $\bar{\kappa}_I$ and downsizing at $\underline{\kappa}_I$ maximize the present discounted value of firm i 's profits. Due to the nature of the problem (singular control), the marginal value is always located inside the hysteresis (wait) region.⁹

⁸In myopic strategies, the optimal myopic strategy is not adapted to the actual evolution of the rival's capital stock $K^j = \{K_t^j\}_{t \geq 0}$ and firms assume instead parallel capital stock development.

⁹The set of times at which investment or divestment occurs has Lebesgue measure zero, the marginal value process being located within the bands almost surely. If $v^i(X_t, K_{t-})$ were outside this region, an instantaneous control of large intensity $\xi_t^i - \xi_{t-}^i$ would adjust the capital stock up or down to ensure that $v^i(X_t, K_t)$ is located in $(\underline{\kappa}_I, \bar{\kappa}_I)$.

Consider next the dynamics of the shadow price. Differentiating HJB equation (4.1) w.r.t. k_i yields

$$(4.8) \quad rv^i(x, k) = \pi_{k^i}^i(x, k) + v_{k^i}^i(x, k) \phi^i(x, k) + V_{k^j}^i(x, k) \phi_{k^i}^j(x, k) + V_{k^i k^i}^i(x, k) \phi^j(x, k) + \mathcal{D}_x v(x, k).$$

The left-hand side is the required return on the value of a marginal unit of capital. The right-hand side of equation (4.8) is the expected return; it consists of the marginal operating profit and the marginal expected gain. Equation (4.8) is thus a no-arbitrage condition to be satisfied at each time and in each state (x, k) in Markov perfect equilibrium. If a strategy $\phi^i(x, k)$ solves the above maximization problem of equation (4.8), the equilibrium shadow price process obtains as the solution of a partial differential equation. The term $V_{k^j}^i(x, k) \phi_{k^i}^j(x, k)$ is the strategic effect term: it captures the impact on firm i 's value from the reaction (capital stock change) by firm j . Since $\phi_{k^i}^j < 0$, this term constitutes a negative externality on the marginal value of incremental investment in capital. Because of the cross influence, the evolution of the shadow price v^i is determined by a system of coupled partial differential equations. Applying Feynman-Kac formula to equation (4.8), it obtains

$$(4.9) \quad v(x, k) = \mathbb{E} \left[\int_0^\infty \{ \pi_{k^i}^i(x, k) + v_{k^i}^i(x, k) [\phi^i(x, k) + \phi^j(x, k)] + V_{k^j}^i(x, k) \phi_{k^i}^j(x, k) \} dt \right].$$

Consider now how the policy influences the dynamic behavior of the shadow price process. After a possible initial discrete adjustment at time 0 occurring if the initial state (x, k) is such that $v^i(x, k) \notin (\underline{\kappa}_i, \bar{\kappa}_I)$, the optimal policy of firm i involves adjusting the capital stock process $\{K_t^i\}_{t \geq 0}$ at the lower and upper thresholds so the firm sustains a certain level (region) of profitability $(\underline{\kappa}_I, \bar{\kappa}_I)$. The shadow price evolution thus admits a lower and upper reflecting barrier, whereby the shadow price evolves according to the partial differential equation (4.8) within the hysteresis region. The determination of the barriers is affected by the presence of competition in Markov perfect equilibrium.

5. Conclusion

The present modeling approach is useful for analyzing dynamic capacity investment decisions in industries subject to stochastic growth and competitive dynamics. We adopted a fairly general approach that covers a large family of stochastic processes used in economic analysis and many reduced-form profit functions. The characterization of the appropriate strategy space is critical to determining equilibria in real option models that involve competition under uncertainty. When firms can observe and react to rivals' actions over time, the equilibrium solution should form a Markov perfect equilibrium. However, deriving the Markov Perfect Equilibrium is usually more involved than deriving the myopic equilibrium since an additional partial derivative obtains in the HJB equation.

In the presence of a convex capital-adjustment cost, the investment rule on marginal Tobin's q still holds. If the investment is costlessly reversible, the model produces Jorgenson's (1963) classic result: at each date (in each state) the capital stock K^i is adjusted to equate the (ever-changing) marginal revenue product of capital $v^i(x, k)$ with user cost of capital. If the investment costs are partly unrecoverable then there exists a hysteresis region in which the firm waits to invest or divest. In Markov perfect equilibrium, these thresholds as well as the intensity of investment are affected by the presence of rivals. The firm with the larger capital stock disciplines its smaller rival away from investing. Besides, the presence of competition creates a hedging effect: under competition, a firm invests with less intensity than if it were ignoring competition.

Summary and Outlook

1. Contributions to capital budgeting

Investment decisions are generally made at early stages when the intensity of competition is not known and the market realizations are not yet observed. Since market developments are not perfectly predictable and managers have flexibility to adapt their decision making to new events, cashflows are state-contingent and often characterized by asymmetric or skewed probability distribution. Besides, in the presence of partly unrecoverable investment costs, assessing an investment strategy becomes a dynamic problem. In this light, valuing an investment opportunity is challenging. From a scientific perspective, the analysis should be performed with use of tools from dynamic optimization and allow considerations to market and strategic uncertainty. In real options analysis, a key feature specific to real assets is often ignored — namely, the interplay or strategic interactions among several firms having similar real investment opportunities. Continuous-time real options analysis mainly rests on stochastic control and often fails short of assessing endogenous strategic interdependencies. A proper theory of investment should thus leverage on both real options analysis and dynamic game theory that together help solve dynamic games in stochastic environments.

This monograph developed an innovative modeling approach to analyze strategic investment under uncertainty that should be appealing to the academic community in the areas of (mathematical) economics, finance and management science. From the perspective of financial theorists, the monograph expands real options analysis with the use of dynamic game theory in such a way that researchers can build upon its content to systematically approach problems of strategic investment under uncertainty. For researchers in industrial organization, it helps stress how the mainstream results may be invalidated under uncertainty. The discussion is deeply anchored in new developments made in the literature on control and game theory.

Strategic investment under uncertainty has not been systematically analyzed from this perspective before, either in its depth (e.g., information structure, connection with stochastic differential games) nor in its scope (e.g., lumpy investments and incremental capital adjustments). Grenadier (2000b) edited an academic book that brings together (published) research contributions that deal with strategy investment under uncertainty; by its very nature, this work does not provide a consistent frame for the analysis of investment decisions under market and strategic uncertainty. Huisman's (2001) doctoral thesis examines a number of continuous-time models of technology adoption, some of which discussing R&D adoption timing games. Smit and Trigeorgis (2004) discuss mainly option games but mainly in a discrete-time setting as a means to raise awareness to these issues among a practitioners' audience. While Ziegler (2010) also combines game theory and option pricing, it focuses on applications to the theory of corporate finance and financial intermediation, such as the design of debt contracts, capital structure choice, the structure of banking deposit contracts, and the incentive effects of deposit insurance. Chevalier-Roignant and Trigeorgis (2011) provide a general introduction to "option games," wherein both the discrete and the continuous-time perspectives are presented; it is essentially intended to an audience of nonaverted undergraduate and graduate students in strategy, economics, and finance. In contrast to these works, we discuss in depth stochastic models where several players interact. We present in detail the underlying probability concepts and the mathematical dynamic optimization tools developed in control and game theory. Throughout the manuscript, we have adopted a fairly general

approach involving diffusion processes and a large class of reduced-form profit functions. A key modeling assumption relates to whether firms react to changes about the market or the industry structure as new information unfolds. We pointed out that this assumption drives much equilibrium results. As argued, when firms are nonatomic decision makers with material influence on market developments (active management), the most reasonable hypothesis is that of closed-loop or Markov strategies.

2. Review of major results

The development of a theory that combines real options analysis with dynamic game theory leads to mathematical complexity that need be addressed first. Part I presents prerequisite notions and concepts from economics and applied mathematics. The main objective is to generalize notions and techniques borrowed from stochastic control theory to multiplayer settings, in so-called stopping and stochastic differential games. In Part II, we provide a short overview over continuous-time real options analysis with particular focus on the benchmark cases of monopoly and perfect competition. The core of the monograph, Part III, consists of three original research contributions offering new insights to the theory of finance and economics dealing with strategy investment under uncertainty. We reviewed recent research contributions dealing with such problems in Chapter 7 and discussed in detail the problem of lumpy investment and capital accumulation under competition and uncertainty, in Chapters 8 and 9, respectively.

More precisely, the review in Chapter 7 helps highlight several managerial insights about the nature of competitive advantage (first versus second-mover advantage), the manner in which information is revealed, firm heterogeneity, capital investment size, and the number of competing firms. It also sets a research agenda.

Chapter 8 discusses a market-entry game under uncertainty, explaining dynamic market entry timing decisions and strategic interactions in the case of stochastically growing markets. In oligopolies, market entries obtain to take place sequentially regardless of whether or not firms can observe and react to their rivals' moves. Besides, unlike the case in which firms disregard competition, entry does not occur in a socially optimal manner when firms observe their rivals' decisions before acting. If there are no entry barriers and a large number of firms can enter a growing market, firms formulate open-loop strategies; thereby, the sequence of investments is shown to be Pareto optimal. The value of the early movers are eroded owing to later market entries by rivals. When only a few firms are protected from new market entry, the sequence of investment is not necessarily Pareto optimal (closed-loop). In this case, Markov perfection implies investment at an earlier time than the timing a social planner would impose on the firms. Restricted availability of the investment option leads to preemption and rent dissipation.

In Chapter 9 on preemptive capacity investment under uncertainty, we discuss capital accumulation under uncertainty in a setting where firms' capital expansions act as a negative externality on each other's payoff. The model developed herein is solved with the help of techniques from the literature on stochastic differential games. Deriving the Markov perfect equilibrium is usually more involved than deriving the myopic equilibrium since an additional partial derivative obtains in the HJB equation. For a convex investment cost function, the investment rule on marginal Tobin's q still holds but with an additional constraint on the value function. If the investment costs are partly unrecoupable then there exists a hysteresis region in which the firm waits to invest or divest. In Markov perfect equilibrium, the intensity at which firms invest or divest outside the hysteresis band are affected by the presence of rivals. A natural hedge effect is shown to exist, whereby firms' capital stock evolve in a less erratic manner under competition. In other words, under competition, a firm invests with less intensity than if it were ignoring competition. Besides, the firm with the larger capital stock disciplines its smaller rival away from investing.

3. Future Research Directions

Research efforts can be directed toward applying and extending these results in several directions. Here a list of areas where further research would be fruitful to achieve enhanced practical and theoretical contributions:

- (1) *Impact of information structure on equilibrium.* Dynamic strategic interactions are sometimes ignored in the strategy space especially in models dealing with capacity expansion.
- (2) *Bayesian learning in investment games.* Information about certain characteristics of the underlying market (e.g., the growth rate) may not be readily available at the outset but learned over time through firms' exercise strategies.
- (3) *Focus on timing.* In the literature on option games, strategy choices are often defined in a very simple manner, namely in terms of timing for lumpy investments or for additional atomic capital units. The objective is to achieve tractability. Yet, real-life investment strategies encompass many other aspects, such as short-term and long-term competition.
- (4) *Narrow view on strategic interdependencies.* Most papers consider horizontal competition in a given market, disregarding favorable or detrimental strategic interdependencies with other value chain parties.
- (5) *Lack of empirical research and application.* Most papers have been concerned with deriving analytical results at the expense of applicability. To increase practical relevance, new decision-making tool kits should be developed. Extending stochastic programming to multiple players would be cumbersome as this approach might give rise to multiple equilibrium solutions.

Throughout our research program, we have highlighted how the choice of an appropriate solution approach (open-loop versus closed-loop) is critical to determining equilibria in real options models that involve competition under uncertainty. Beyond observing a gap in the existing literature, we proposed ways to tackle the informational problem and showed the value and welfare implications of each information structure choice. When firms can observe and react to rivals' actions, the appropriate solution involves closed-loop strategies and a Markov perfect equilibrium. However, deriving the MPE in closed-loop strategies is usually more involved than deriving the Nash equilibrium for the case of open-loop strategies. Since the distinction and the resulting implication were not always clear, researchers in the relevant literature stream often relied on approaches with an open-loop flavor, effectively discarding certain interesting dynamic strategic interactions. Our research will hopefully contribute both theoretically and practically to the literature on strategic investment under uncertainty by developing tools to tackle the informational problem, a problem recurrent in all related research endeavors.

Bibliography

- Abel, A. B. (1983). Optimal investment under uncertainty. *American Economic Review* 73(1), 228–233.
- Abel, A. B. and J. C. Eberly (1994). A unified model of investment under uncertainty. *American Economic Review* 84(5), 1369–1384.
- Abel, A. B. and J. C. Eberly (1996). Optimal investment with costly reversibility. *Review of Economic Studies* 63(4), 581–593.
- Abel, A. B. and J. C. Eberly (1997). An exact solution for the investment and value of a firm facing uncertainty, adjustment costs, and irreversibility. *Journal of Economic Dynamics and Control* 21(4-5), 831–852.
- Aguerrevere, F. L. (2003). Equilibrium investment strategies and output price behavior: A real-options approach. *Review of Financial Studies* 16(4), 1239–1272.
- Aguerrevere, F. L. (2009). Real options, product market competition, and asset returns. *Journal of Finance* 64(2), 957–983.
- Akerlof, G. A. (1970). The market for “lemons.” Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84(3), 488–500.
- Bachelier, L. (1900). La theorie de la spéculation. *Annales Scientifiques de l’Ecole Normale Supérieure* 17, 21–86.
- Back, K. and D. Paulsen (2009). Open-loop equilibria and perfect competition in option exercise games. *Review of Financial Studies* 22(11), 4531–4552.
- Başar, T. and G. J. Olsder (1999). *Dynamic Noncooperative Game Theory* (2 ed.). Society for Industrial and Applied Mathematics.
- Baldursson, F. M. (1998). Irreversible investment under uncertainty in oligopoly. *Journal of Economic Dynamics and Control* 22(4), 627–644.
- Baldursson, F. M. and I. Karatzas (1996). Irreversible investment and industry equilibrium. *Finance and Stochastics* 1(1), 69–89.
- Bar-Ilan, A. and W. Strange (1996). Investment lags. *American Economic Review* 86(3), 610–622.
- Barelli, P. and S. de Abreu Pessa (2003). Inada conditions imply that production function must be asymptotically Cobb-Douglas. *Economics Letters* 81(3), 361–363.
- Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.
- Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica* 52(4), 1007–1028.
- Bertola, G. (1998). Irreversible investment. *Research in Economics* 52(1), 3–37.
- Bertola, G. and R. Caballero (1994). Irreversibility and aggregate investment. *Review of Economic Studies* 61(2), 223–246.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3), 637–654.
- Borkar, V. S. (2005). Controlled diffusion processes. *Probability Surveys* 2, 213–244.
- Bouis, R., K. J. M. Huisman, and P. M. Kort (2009). Investment in oligopoly under uncertainty: The accordion effect. *International Journal of Industrial Organization* 27(2), 320–331.
- Boyer, M., E. Gravel, and P. Lasserre (2004). Real options and strategic competition: a survey. Working paper, University of Montreal.
- Boyer, M., P. Lasserre, T. Mariotti, and M. Moreaux (2004). Preemption and rent dissipation under price competition. *International Journal of Industrial Organization* 22(3), 309–328.

- Boyer, M., P. Lasserre, and M. Moreaux (2010). The dynamics of strategic capacity investments under uncertain market growth. IDEI Working Paper 617, Toulouse School of Economics.
- Brennan, M. J. and E. S. Schwartz (1977). The valuation of American put options. *Journal of Finance* 32(2), 449–462.
- Carlson, M. D., E. J. Dockner, A. J. Fisher, and R. Giammarino (2010). Leaders, followers, and risk dynamics in industry equilibrium. Working paper series, SSRN.
- Cellini, R. and L. Lambertini (2008). Weak and strong time consistency in a differential oligopoly game with capital accumulation. *Journal of Optimization Theory and Applications* 138(1), 17–26.
- Chevalier-Roignant, B., C. M. Flath, A. Huchzermeier, and L. Trigeorgis (2011). Strategic investment under uncertainty: A synthesis. *European Journal of Operational Research*. forthcoming.
- Chevalier-Roignant, B. and L. Trigeorgis (2011). *Competitive Strategy: Options and Games*. The MIT Press.
- Cottrell, T. and G. Sick (2002). Real options and follower strategies: The loss of real option value to first-mover advantage. *Engineering Economist* 47, 232–263.
- Cox, J. C. and S. A. Ross (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3(1-2), 145–166.
- Décamps, J.-P. and T. Mariotti (2004). Investment timing and learning externalities. *Journal of Economic Theory* 118(1), 80–102.
- del Sol, P. and P. Ghemawat (1999). Strategic valuation of investment under competition. *Interfaces* 29(6), 42–56.
- Dias, M. A. G. and J. P. Teixeira (2010). Continuous-time option games: review of models and extensions. *Multinational Finance Journal* 14(3/4), 219–254.
- Dixit, A. K. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy* 97(3), 620–638.
- Dixit, A. K. (1991). Irreversible investment with price ceilings. *Journal of Political Economy* 99(3), 541–557.
- Dixit, A. K. (2002). *The Art of Smooth Pasting*. Harwood Fundamentals of Applied Economics, Routledge.
- Dixit, A. K. and R. S. Pindyck (1994). *Investment under Uncertainty*. Princeton University Press.
- Dockner, E. J., S. Jorgensen, N. Van Long, and G. Sorger (2000). *Differential Games in Economics and Management Science*. Cambridge University Press.
- Doebelin, V. (1940). Sur l'équation de Kolmogoroff. *Comptes rendus de l'Académie des Sciences de Paris* 1(331), 1059–1102.
- Duffie, D. (2001). *Dynamic Asset Pricing Theory* (3 ed.). Princeton University Press.
- Dumas, B. (1991). Super contact and related optimality conditions. *Journal of Economic Dynamics and Control* 15, 675–685.
- Dyson, R. G. (2000). Strategy, performance and operational research. *Journal of the Operational Research Society* 51(1), 5–11.
- Eberly, J. C. and J. A. Van Mieghem (1997). Multi-factor dynamic investment under uncertainty. *Journal of Economic Theory* 75(2), 345–387.
- Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* 322(8), 549–560.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business* 38(1), pp. 34–105.
- Ferreira, N., J. Kar, and L. Trigeorgis (2009). Option games: The key to competing in capital-intensive industries. *Harvard Business Review* 87(3), 101–107.
- Fine, C. H. and L. Li (1989). Equilibrium exit in stochastically declining industries. *Games and Economic Behavior* 1(1), 40–59.

- Fudenberg, D. and D. Levine (1988). Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory* 44(1), 1–18.
- Fudenberg, D. and J. Tirole (1983). Capital as commitment: Strategic investment to deter mobility. *Journal of Economic Theory* 250, 227–250.
- Fudenberg, D. and J. Tirole (1984). The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *American Economic Review* 74(2), 361–366.
- Fudenberg, D. and J. Tirole (1985). Preemption and rent equilization in the adoption of new technology. *Review of Economic Studies* 52(3), 383–401.
- Fudenberg, D. and J. Tirole (1986). A theory of exit in duopoly. *Econometrica* 54(4), pp. 943–960.
- Fudenberg, D. and J. Tirole (1991). *Game Theory*. The MIT Press.
- Fudenberg, D. and J. Tirole (2002). *Dynamic models of oligopoly*. Harwood Fundamentals of Applied Economics, Routledge.
- Ghemawat, P. and B. Nalebuff (1985). Exit. *RAND Journal of Economics* 16(2), 184–194.
- Ghemawat, P. and B. Nalebuff (1990). The devolution of declining industries. *Quarterly Journal of Economics* 105(1), 167–186.
- Gilbert, R. and R. Harris (1984). Competition with lumpy investment. *RAND Journal of Economics* 15(2), 197–212.
- Grenadier, S. R. (1996). The strategic exercise of options: Development cascades and overbuilding in real estate markets. *Journal of Finance* 51(5), 1653–1679.
- Grenadier, S. R. (1999). Information revelation through option exercise. *Review of Financial Studies* 12(1), 95–129.
- Grenadier, S. R. (2000a). Equilibrium with time-to-build: A real options approach. In M. J. Brennan and L. Trigeorgis (Eds.), *Project Flexibility, Agency, and Competition: New Developments in the Theory and Application of Real Options*, pp. 275–296. Oxford University Press.
- Grenadier, S. R. (Ed.) (2000b). *Game Choices: The Intersection of Real Options and Game Theory*. Risk Books.
- Grenadier, S. R. (2002). Option exercise games: An application to the equilibrium investment strategies of firms. *Review of Financial Studies* 15(3), 691–721.
- Harrison, J. M. (1985). *Brownian Motion and Stochastic Flow Systems*. John Wiley & Sons.
- Harrison, J. M. and D. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20(3), 381–408.
- Hoppe, H. C. (2000). Second-mover advantages in the strategic adoption of new technology under uncertainty. *International Journal of Industrial Organization* 18(2), 315–338.
- Hsu, Y.-W. and B. Lambrecht (2007). Preemptive patenting under uncertainty and asymmetric information. *Annals of Operations Research* 151, 5–28.
- Huisman, K. J. M. (2001). *Technology Investment: A Game Theoretic Real Options Approach*. Kluwer Academic Publishers.
- Huisman, K. J. M. and P. M. Kort (1999). Effects of strategic interactions on the option value of waiting. Working paper, Tilburg University.
- Huisman, K. J. M. and P. M. Kort (2009). Strategic capacity investment under uncertainty. In *Proceedings of the 13th Annual International Real Options Conference, June 17-20, 2009*.
- Huisman, K. J. M., P. M. Kort, G. Pawlina, and J. J. J. Thijssen (2004). Strategic investment under uncertainty: Merging real options with game theory. *Zeitschrift für Betriebswirtschaft* 74, 97–123.
- Inada, K.-I. (1963). On a two-sector model of economic growth: Comments and a generalization. *Review of Economic Studies* 30(2), 119–127.
- Isaacs, R. (1965). Differential games i, ii, iii, iv. *The RAND Corporation*.
- Itô, K. (1942). Differential equations determining Markov processes. *Zenkoku Shijo Sugaku Danwakai* 1077, 1352–1400.

- Itô, K. (1951). On stochastic differential equations. *Memoirs of American Mathematical Society* 4, 1–51.
- Joaquin, D. C. and K. C. Butler (2000). Competitive investment decisions — a synthesis. In M. J. Brennan and L. Trigeorgis (Eds.), *Project Flexibility, Agency, and Competition: New Developments in the Theory and Application of Real Options*, pp. 324–339. Oxford University Press, New York.
- Jorgenson, D. (1963). Capital theory and investment behavior. *American Economic Review* 53(2), 247–259.
- Karatzas, I. and S. E. Shreve (1984). Connections between optimal stopping and singular stochastic control. I. Monotone follower problems. *SIAM Journal on Control and Optimization* 22, 856.
- Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (2nd ed.). Springer Verlag.
- Kong, J. J. and Y. K. Kwok (2007). Real options in strategic investment games between two asymmetric firms. *European Journal of Operational Research* 181(2), 967–985.
- Kreps, D. and A. M. Spence (1984). Modelling the role of history in industrial organization and competition. pp. 340–378. Macmillan.
- Kuhn, H. (1953). Extensive games and the problem of information. *Annals of Mathematics Studies* 28, 193–216.
- Kulatilaka, N. and E. C. Perotti (1998). Strategic growth options. *Management Science* 44(8), 1021–1031.
- Kunita, H. and S. Watanabe (1967). On square integrable martingales. *Nagoya Mathematical Journal* 30(209-245), 209–245.
- Lambrecht, B. (2000). Strategic sequential investments and sleeping patents. In M. J. Brennan and L. Trigeorgis (Eds.), *Project Flexibility, Agency, and Competition: New Developments in the Theory and Application of Real Options*, pp. 297–323. Oxford University Press.
- Lambrecht, B. and W. Perraudin (2003). Real options and preemption under incomplete information. *Journal of Economic Dynamics and Control* 27, 619–643.
- Leahy, J. V. (1993). Investment in competitive equilibrium: The optimality of myopic behavior. *Quarterly Journal of Economics* 108(4), 1105–1133.
- Lévy, P. (1965). *Processus stochastiques et mouvement Brownien*. Gauthier-Villars.
- Lucas, R. E. J. and E. Prescott (1971). Investment under uncertainty. *Econometrica* 39(5), 659–681.
- Luss, H. (1982). Operations research and capacity expansion problems: A survey. *Operations Research* 30(5), 907–947.
- Maskin, E. S. and J. Tirole (2001). Markov perfect equilibrium. i. Observable actions. *Journal of Economic Theory* 100(2), 191–219.
- Mason, R. and H. Weeds (2010). Investment, uncertainty, and pre-emption. *International Journal of Industrial Organization* 28(3), 278–287.
- McDonald, R. L. and D. R. Siegel (1985). Investment and the valuation of firms when there is an option to shut down. *International Economic Review* 26(2), 331–349.
- McDonald, R. L. and D. R. Siegel (1986). The value of waiting to invest. *Quarterly Journal of Economics* 101(4), 707–728.
- Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4(1), 141–183.
- Merton, R. C. (1992). *Continuous-Time Finance*. Wiley-Blackwell.
- Merton, R. C. (1998). Applications of option-pricing theory: Twenty-five years later. *American Economic Review* 88(3), 323–349.
- Miltersen, K. and E. S. Schwartz (2004). R&D investments with competitive interactions. *Review of Finance* 8(3), 355–401.
- Murto, P. (2004). Exit in duopoly under uncertainty. *RAND Journal of Economics* 35(1), 111–127.

- Murto, P., E. Näsäkkälä, and J. Keppo (2004). Timing of investments in oligopoly under uncertainty: A framework for numerical analysis. *European Journal of Operational Research* 157(2), 486–500.
- Myers, S. (1977). Determinants of corporate borrowing. *Journal of Financial Economics* 5(2), 147–175.
- Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Harvard University Press.
- Nash, J. F. (1951). Non-cooperative games. *The Annals of Mathematics* 54(2), 286–295.
- Nishihara, M. and M. Fukushima (2008). Evaluation of firm's loss due to incomplete information in real investment decision. *European Journal of Operational Research* 188(2), 569–585.
- Novy-Marx, R. (2007). An equilibrium model of investment under uncertainty. *Review of Financial Studies* 20(5), 1461–1502.
- Novy-Marx, R. (2009). Preempting Preemptive Investment. Working paper, University of Chicago.
- Øksendal, B. (2007). *Stochastic Differential Equations* (6th ed.). Springer Verlag.
- Pawlina, G. and P. M. Kort (2003). Strategic capital budgeting: asset replacement under market uncertainty. *OR Spectrum* 25, 443–479. 10.1007/s00291-003-0137-3.
- Pawlina, G. and P. M. Kort (2006). Real options in an asymmetric duopoly: Who benefits from your competitive disadvantage? *Journal of Economics & Management Strategy* 15(1), 1–35.
- Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52(4), 1029–1050.
- Pennings, E. (2004). Optimal pricing and quality choice when investment in quality is irreversible. *Journal of Industrial Economics* 52(4), 569–589.
- Peskir, G. and A. Shiryaev (2006). *Optimal Stopping and Free-Boundary Problems*. Birkhauser Verlag.
- Pindyck, R. S. (1988). Irreversible investment, capacity choice, and the value of the firm. *American Economic Review* 78(5), 969–985.
- Porter, M. E. (1980). *Competitive Strategy*. Free Press.
- Protter, P. E. (2003). *Stochastic Integration and Differential Equations* (2nd ed.). Springer Verlag.
- Ramey, V. and M. Shapiro (2001). Displaced capital: A study of aerospace plant closings. *Journal of Political Economy* 109(5), 958–992.
- Ramsey, F. (1928). A mathematical theory of saving. *Economic Journal* 38(152), 543–559.
- Reinganum, J. F. (1981a). Market structure and the diffusion of new technology. *Bell Journal of Economics* 12(2), 618–624.
- Reinganum, J. F. (1981b). On the diffusion of new technology: A game theoretic approach. *Review of Economic Studies* 48(3), 395–405.
- Reynolds, S. S. (1987). Capacity investment, preemption and commitment in an infinite horizon model. *International Economic Review* 28(1), 69–88.
- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6(2), 41–49.
- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4(1), 25–55.
- Sick, G. (1995). Real options. In V. M. R.A. Jarrow and W. Ziemba (Eds.), *Finance*, Volume 9 of *Handbooks in Operations Research and Management Science*, Chapter 21, pp. 631–691. Elsevier.
- Slade, M. E. (1994). What does an oligopoly maximize? *Journal of Industrial Economics* 42(1), 45–61.
- Smets, F. (1991). *Exporting Versus Foreign Direct Investment: The Effect of Uncertainty, Irreversibilities and Strategic Interaction*. Ph. D. thesis, Yale University, New Haven, CT.
- Smit, H. T. J. and L. A. Ankum (1993). A real options and game-theoretic approach to corporate investment strategy under competition. *Financial Management* 22(3), 241–250.
- Smit, H. T. J. and L. Trigeorgis (2001, 9). Flexibility and commitment in strategic investment. In E. S. Schwartz and L. Trigeorgis (Eds.), *Real options and investment under uncertainty: Classical readings and recent contributions*, Chapter 21, pp. 451–498. The MIT Press.

- Smit, H. T. J. and L. Trigeorgis (2004). *Strategic Investment: Real Options and Games*. Princeton University Press.
- Smit, H. T. J. and L. Trigeorgis (2007). Strategic options and games in analysing dynamic technology investments. *Long Range Planning* 40(1), 84–114.
- Smit, H. T. J. and L. Trigeorgis (2009). Valuing infrastructure investment: An option games approach. *California Management Review* 51(2), 79–90.
- Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1), 65–94.
- Sparla, T. (2004). Closure options in a duopoly with strong strategic externalities. *Zeitschrift für Betriebswirtschaft* 67, 125–155.
- Spence, A. M. (1979). Investment strategy and growth in a new market. *Bell Journal of Economics* 10(1), 1–19.
- Spencer, B. J. and J. A. Brander (1992). Pre-commitment and flexibility: Applications to oligopoly theory. *European Economic Review* 36(8), 1601–1626.
- Starr, A. and Y. Ho (1969). Nonzero-sum differential games. *Journal of Optimization Theory and Applications* 3(3), 184–206.
- Stokey, N. L. (2008). *The Economics of Inaction*. Princeton University Press.
- Swan, T. (1994). Economic growth and capital accumulation. *Knut Wicksell: critical assessments* 32, 153.
- Takashima, R., M. Goto, H. Kimura, and H. Madarame (2008). Entry into the electricity market: Uncertainty, competition, and mothballing options. *Energy Economics* 30(4), 1809–1830.
- Thijssen, J. J. J., K. J. M. Huisman, and P. M. Kort (2002). Symmetric equilibrium strategies in game theoretic real option models. CentER Working paper 81, Tilburg University.
- Thijssen, J. J. J., K. J. M. Huisman, and P. M. Kort (2006). The effects of information on strategic investment and welfare. *Economic Theory* 28(2), 399–424.
- Tirole, J. (1988). *The Theory of Industrial Organization*. The MIT Press.
- Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of money, credit and banking* 1(1), 15–29.
- Touzi, N. and N. Vieille (2002). Continuous-time Dynkin games with mixed strategies. *SIAM Journal on Control and Optimization* 41(4), 1073–1088.
- Trigeorgis, L. (1991). Anticipated competitive entry and early preemptive investment in deferrable projects. *Journal of Economics and Business* 43(2), 143–156.
- Trigeorgis, L. (1996). *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. The MIT Press.
- Tsekrekos, A. E. (2010). The effect of mean reversion on entry and exit decisions under uncertainty. *Journal of Economic Dynamics and Control* 34(4), 725–742.
- Van Mieghem, J. A. (2003). Capacity management, investment, and hedging: Review and recent developments. *Manufacturing & Service Operations Management* 5(4), 269–302.
- Van Mieghem, J. A. (2008). *Operations Strategy: Principles and Practice*. Dynamic Ideas.
- Weeds, H. (2000). Sleeping patents and compulsory licensing: An options analysis. Economic research paper 577, University of Warwick.
- Weeds, H. (2002). Strategic delay in a real options model of R&D competition. *Review of Economic Studies* 69, 729–747.
- Yong, J. and X. Y. Zhou (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer Verlag.
- Zhu, K. and J. Weyant (2003a). Strategic decisions of new technology adoption under asymmetric information: A game-theoretic model. *Decision Sciences* 34(4), 643–675.

- Zhu, K. and J. Weyant (2003b). Strategic exercise of real options: Investment decisions in technological systems. *Journal of Systems Science and Systems Engineering* 12(3), 257–278.
- Ziegler, A. C. (2010, 12). *A Game Theory Analysis of Options: Corporate Finance and Financial Intermediation in Continuous Time* (2 ed.). Springer Finance. Springer.

Last name: Chevalier-Roignant

First names: Benoît Damien

AFFIRMATION - STATUTORY DECLARATION

*According to §10 No. 7 of the Doctoral Studies' Guidelines
(As Amended on the 5th March 2008)*

I hereby declare, that the *doctoral dissertation* submitted to the Wissenschaftliche Hochschule für Unternehmensführung (WHU) – Otto-Beisheim-Hochschule– was produced independently and without the aid of sources other than those which have been indicated. All ideas and thoughts coming both directly and indirectly from outside sources have been noted as such.

This work has previously not been presented in any similar form to any other board of examiners.

Sentences or text phrases, taken out of other sources either literally or as regards contents, have been marked accordingly. Without notion of its origin, including sources which are available via internet, those phrases or sentences are to be considered as plagiarisms. It is the WHU's right to check submitted dissertations with the aid of software that is able to identify plagiarisms in order to make sure that those dissertations have been rightfully composed. I agree to that kind of checking my dissertation.

The following persons helped me gratuitous/non-gratuitous in the indicated way in selecting and evaluating the used materials:

- Arnd HUCHZERMEIER and Lenos TRIGEORGIS acted as PhD supervisor and provided gratuitous advise;
- Christoph FLATH co-authored Chapter 9 (gratuitous);
- Matt DARNELL acted a copyeditor for chapters 7 and 9 (non-gratuitous).

Further persons have not been involved in the preparation of the presented dissertation as regards contents or in substance. In particular, I have not drawn on the non-gratuitous help of placement or advisory services (doctoral counsels / PhD advisors or other persons). Nobody has received direct or indirect monetary benefits for services that are in connection with the contents of the presented dissertation.

The dissertation does not contain texts or (parts of) chapters that are subject of current or completed dissertation projects.

Place and date of issue:

Signature: