Toxvaerd, Flavio

(2002)

Studies in the dynamics of contracts and markets.

Doctoral thesis, University of London: London Business School.

DOI: https://doi.org/10.35065/PNFT3596

STUDIES IN THE DYNAMICS OF CONTRACTS AND MARKETS

FLAVIO MARTIN OBEDMAN TOXVAERD

June 2002

London Business School

Submitted to the University of London for the degree of Doctor of Philosophy

STUDIES IN THE DYNAMICS OF CONTRACTS AND MARKETS

FLAVIO MARTIN OBEDMAN TOXVAERD

June 2002

London Business School

Submitted to the University of London for the degree of Doctor of Philosophy

ABSTRACT. The present thesis studies two distinct issues, namely merger waves and optimal contracts for delivery in settings with time to build.

The first part of the thesis proposes a theoretical explanation for the occurrence of merger waves. Mergers and acquisitions (M&A) come in waves, a fact that has puzzled economists as far back as the 1950's. Accordingly, there exists a vast empirical literature studying the time-series properties of M&A activity, and numerous studies devoted to identifying correlation between M&A activity and economic and financial variables. A review of merger theory is presented in the first chapter. In the second chapter, a formal model is developed that incorporates both dependence of the merger decision on macroeconomic variables and strategic interaction between firms. Specifically, a model is set up in which a number of acquiring firms compete over time for a small number of target firms. In each period, an acquirer may either attempt a takeover, or postpone the takeover decision. By delaying a takeover attempt, the acquiring firm may gain from more favourable future market conditions. On the other hand, postponement of the takeover involves the risk of preemption from rivals. This tradeoff leads to equilibria in which all acquirers simultaneously seek to merge. An extension of the model into one of incomplete information (i.e. a setting in which there is strategic uncertainty) allows one to pin down a unique perfect Bayesian equilibrium, and thus the expected timing of the merger wave.

The second part of the thesis studies contractual relationships between economic agents in situations where there is time to build. Specifically, it seeks to analyse how delivery time is determined under asymmetric information. In order to do this, two different models are presented, each one focusing on a separate issue. The first is a continuous time adverse selection model in which a principal hires an agent to complete a project, but where the agent's ability is private information and unknown to the principal. Furthermore, the principal is unable to monitor the agent's rate of effort or progress on the project. In this setting the optimal contract is derived and characterised. The main finding is that the principal can use completion time to screen agents of different efficiency. The optimal contract thus specifies wages as a function of completion time in a way that optimally trades off efficiency with informational rents. It turns out that the optimal contract has the most efficient agent deliver at the efficient point in time, paying him large informational rents. For less efficient agents, the optimal contract stipulates inefficient delay in delivery time, with the most inefficient agent receiving no informational rents. The second model is one of dynamic moral hazard, in which the agent's effort is unobservable to the principal. In order to complete the project, the agent must successfully complete a sequence of distinct tasks in a fixed pre-specified order. Whether or not a task is completed depends on the agent's effort. In this setting, different contracts are analysed. Namely, the cases of observable effort and unobservable effort with spot contracting are characterised. The analysis shows that under both scenarios, project delays are most likely to occur in early stages of development, and are related directly to the power of the offered incentive scheme. Last, contracts with commitment on the part of the principal are discussed.

List of Figures

For my parents and Chryssi.

Problems worthy of attack

Prove their worth by hitting back

— Piet Hein

CHAPTER 1

Preface

The work contained in this thesis is characterised by one main feature, namely that it is dynamic in nature. While economic theory has gained profound insights from the study of static models, many situations are inherently dynamic and cannot reasonably be analysed within static frameworks. In my opinion, some of the more interesting puzzles in economic theory are those related to the timing of events. Static models may illuminate *why* agents act in certain ways (e.g. sell, buy, merge, consume, work on thesis etc.), but are by their very nature inadequate in predicting *when* agents act. The work in the present thesis considers two such puzzles, namely the fact that mergers and acquisitions happen in waves and that large-term projects are usually delayed.

While a researcher usually works on a topic in order to fully understand it, it is common that the process raises more questions than it answers. My work is no different in this respect. Time and again, I find myself working both forwards and backwards from my chosen topic. Forwards in the sense that I start wondering where else my models may lead. Backwards in the sense that I start wondering about which economic processes may lead to the setup I have chosen to analyse. I have thus often had to force myself to constrain my curiosity and limit the scope of my work. Fortunately, this may prove a blessing in disguise. Having chosen to continue in economic research, I hope in the future to answer some of the many questions I have asked myself and further explore some of the many avenues I have had to ignore.

While writing this thesis, I benefited from the generous hospitality of several departments and institutes. In particular, GREMAQ at the Universite de Toulouse, the Haas School of Business at Berkeley, Institut d'Analisi Economica in Barcelona, the Department of Economics at Universitat Autonoma de Barcelona and the Stern School of Business at NYU. Despite the obvious effort involved in so much traveling, I greatly benefited from all these stays and visits.

CHAPTER 2

Acknowledgments

Special gratitude is owed to Luis Cabral, who supervised my doctoral studies, even beyond the point of duty. I thank him for his hospitality in Berkeley and in New York. Last, I thank him for his continuous insistence that I take up challenging projects, and firmness when dissuading me from working on the impossible ones.

I wish to express gratitude to Andrew Scott, who provided moral support over the years, and selflessly took on duties that were not his.

Early in my studies, Birgit Grodal showed me how beautiful formal economic models can be. Her passion inspired me to pursue economic research, and for this I am grateful.

Needless to say, my parents have been instrumental in me reaching this point. Their loving support and keen interest (in what to them must have seemed incomprehensible esoteric abstractions) made me stronger. To them I owe everything.

Last, I thank my partner in life Chryssi Giannitsarou. Your support and help cannot be understated. Your unwavering belief in me kept me going through maddening times of intense pressure. If not for you, I may well have thrown in the towel on several occasions. Most of all, I thank you for sharing your life with me. Each day with you is like the first day of spring.

Part 1

Strategic Merger Waves

CHAPTER 3

Background and Review of the Literature

"To merge or not to merge, is that the question?"

- Mathewson and Quigley (1988)

1. Introduction

The study of mergers has long been a staple of industrial organisation theory. This is not only due to the intrinsic interest in mergers as a theoretical issue, but largely due to the potentially important welfare implications of mergers. Regulators have long been suspicious of mergers, fearing that they can have adverse effects on competition and consumer welfare. Spurred by these concerns, economists have built models of mergers, traditionally cast in the framework of well known oligopoly models such as quantity setting or price setting with imperfect substitutes.

In this part of the thesis, I will be presenting a theory, not of mergers per se, but of merger waves. That is, my starting point will be that for some unspecified reason, mergers are potentially profitable, and show how this may lead to wave patterns. The following section will give a brief overview of theoretical work on mergers and merger waves. In the subsequent chapter, a formal model of strategic merger waves will be presented.

2. Merger Theory

2.1. Static Merger Theory. The early literature studied what has become known as exogenous mergers, i.e. it started the analysis by assuming that mergers take place, and then analysed the effects of these on prices, market concentration, profits and consumer surplus. Salant, Switzer and Reynolds (1983) were the first to point out the dangers of studying the effects of exogenous mergers. Using Cournot's original model with symmetric, constant marginal costs, linear demand and no fixed costs, they show that two or more firms who merge will find it optimal to behave as a single firm, and not like a multiplant firm. They find that merging in this model can lead to losses for the merging firms, while increasing profits for non-merging rival firms. In a sense, merging is a public good, as it increases industry profits at the cost of the merging firms. This paper led other researchers to study endogenous mergers more closely in order to rectify this apparent paradox.

A comprehensive analysis of endogenous mergers appeared first in Farrell and Shapiro (1990). They consider a homogenous good Cournot model with possibly non-linear prices and asymmetric cost functions. They show that the Salant et al. model provides a lower bound on merger profitability. Specifically, they show that if synergies from mergers are large enough, mergers can indeed be profitable. Perry and Porter (1985) reverse Salant et al.'s (un)profitability result by assuming that production employs an asset in limited supply. A merged firm will not behave as a single firm after a merger, but combine the assets of the two merging parties. They further show that, in this setting, a merger can indeed be profitable for the merging firms and that merging no longer constitutes a public good. Similar results are obtained by Davidson and Deneckere (1983) in a differentiated goods Bertrand model.

Summing up, the theoretical literature on static mergers predicts that in most settings, mergers can indeed be profitable, although some contributions point to the contrary. Last, it is worth noting that most papers ignore the takeover process itself, and assume that if a merger is profitable, the parties will find a way to split the gains. Notable exceptions are Saloner (1987), Hviid and Prendergast (1993) and Harris (1994). Importantly, most contributions on static mergers focus on either purely strategic issues, or on synergies created through the mergers. In other words, they do not consider any dependence on exogenous macroeconomic or industry-wide factors, although the empirical literature strongly suggests that such factors play an important role on merger profitability.

2.2. Mergers and Managerial Concerns. Fauli-Oller and Motta (1996) consider a model where some firms are run by professional managers, *managerial* firms, and other firms are run by owners or profit maximisers, *entrepreneurial* firms. They assume that the managers' incentives are not only based on profits, but also on sales or market share. If managers are also given discretion over merger decisions, they establish that, in equilibrium, managers will compete aggressively on the product market in order for rival firms' profits to decrease, thereby forcing them to sell out at lower prices. As a possible negative side effect, managers might undertake unprofitable mergers.

Teall (1992) employs concepts from cooperative game theory to study the interaction between managerial compensation and merger activity. He focuses on the bargaining between the manager and the owner over the division of the benefits of a merger, and shows that only if the manager has a significant stake in the firm, will he voluntarily initiate takeover bids.

Last, Roll (1986) advances a hubris hypothesis where managers consistently overestimate the gains from merger. He provides some empirical evidence consistent with that hypothesis.

Although some of this work makes some interesting points, common to most studies that emphasise managerial concerns in the merger decision is the absence of any link between managerial incentives and the timing of mergers.

2.3. Dynamic Merger Theory. The literature reviewed so far has emphasised that firms merge in order to reap the benefits of market power or synergies or that mergers were the effect of managers seeking private benefit. The first contribution to address the endogenous intertemporal linkage between mergers is Nilssen and Sørgard (1998). They consider a perfect information sequential Cournot model with fixed costs and constant marginal costs à la Salant et al. (1983). A crucial feature of their analysis is that firms that are about to merge, correctly anticipate the effect of their merger decision on subsequent firms' decisions. They show that the optimal merger policies are very sensitive to parameter values. Specifically, they allow for synergies following a merger by letting both variable and fixed costs change. They show that for different parameterisations, the model predicts the following outcomes: (i) a merger can lead to subsequent mergers, (ii) a merger can cause other firms not to merge, (iii) not merging will lead other firms to merge, and (iv) not merging will prompt other firms not to merge either. In short, anything can happen, depending on model specifics. An obvious conclusion is that standard oligopoly theory does not provide us with robust predictions.

In contrast to the previous model, Fauli-Oller (2000) considers a sequential setting, where efficient firms bid competitively for inefficient firms in the same industry. He assumes that if a bid is successful and leads to a merger, the previously inefficient firm becomes efficient. An important feature of his model is that the acquisition stage is modeled explicitly. This allows for a richer analysis, while still assuming perfect information. His main results are that a merger increases the profitability of future mergers and that early acquirers pay lower prices for their targets than do later acquirers.

Some interesting work on dynamic endogenous mergers is that by Fridolfsson and Stennek (2000). They study mergers in a preemption type game, where a merger is a way of barring one's partners from merging with other firms. An important assumption driving their results is that although merging does not increase profits per se, being alone when other firms have merged is even worse (i.e. there are negative externalities through the product market). Importantly, their analysis (carried out within a three firm Cournot model) assumes that significant synergies are created through mergers.

A couple of papers raise the interesting question of whether an industry where mergers take place will eventually end with full monopolisation. Kamien and Zang (1990) develop a model where firms competitively bid for each other. They show that only in relatively small industries can equilibrium bidding lead to total monopolisation. However, they further establish that this possibility is ruled out, if competition laws bar any firm from owning more than half of the industry's firms. In Kamien and Zang (1993) they extend the analysis to a sequential game and again find limited scope for full monopolisation. Recognising the inherent difficulties in explicitly solving models of dynamic endogenous mergers, Gowrisankaran (1999) and Gowrisankaran and Holmes (2000) offer simulated results on the evolution of an industry with infinite horizon. They largely confirm the stylised fact that full monopolisation is very rare.

3. Merger Waves

Mergers and acquisitions (M&A) come in waves, both economy-wide and industry-wide. During these waves, billions worth of assets change hands. In 1995 alone, the value of M&A equaled 5% of United States GDP.[1] It is thus hard to ignore the economic importance of mergers and acquisitions. Accordingly, a vast empirical literature has sought to uncover the forces leading to mergers.[2] The evidence suggests that macroeconomic variables play an important role in determining the timing of mergers. Specifically, merger activity is found to be highly procyclical, slightly leading the business cycle. Other research has documented a relation between merger activity and factors such as economy-wide

[1]See Andrade and Stafford (1999).

[2]For extensive reviews of the empirical literature on merger activity see Golbe and White (1988), Weston, Chung and Hoag (1990) and Blair and Schary (1993). For time series analyses, see also Nelson (1959), Shughart et al. (1984), Town (1993) and Barkoulas et al. (2001).

dispersion in Tobin's q (Jovanovic and Rousseau, 2002), industrial production (Gort, 1969 and Mitchell and Mulherin, 1996) and stock price indices.

On the other hand, the business and popular press often stress that managers take other managers' actions into account when deciding on if and when to merge. The industrial organisation literature on mergers has had little to say about this issue. It has largely ignored any dependence on factors relating to aggregate activity, and focused almost exclusively on issues of synergies, or purely strategic issues such as market power. The theoretical corporate finance literature has mainly studied the effects of financial variables on the merger decision, such as "free cash flow", the real cost of capital (i.e. absent any strategic considerations) and gone to great lengths at studying the transaction process itself.

Merger wave theories can be categorised according to whether or not they incorporate strategic elements. I will refer to them as *strategic* and *non-strategic* theories, respectively. Strategic theories of merger waves explicitly account for the mechanism through which one merger is related to the other. For example, the industrial organisation literature has mainly focused on strategic interaction through the product market. Nevertheless, standard oligopoly theory is not entirely satisfactory in explaining merger waves. In fact, the simplest model predicts that mergers should *not* happen in waves. To see this, consider an n-firm homogeneous product Cournot game with constant marginal costs, linear demand and (duplicated) fixed costs. The gain from merger, $\pi(n-1) - 2\pi(n)$, is *increasing* in n. In other words, the incentive for any two firms to merge decreases as the number of mergers in the industry increases. In general, it should be noted that there is an inherent weakness in all models that rely on product market interaction, in that they can at most explain a subset of observed mergers. Specifically, a model of horizontal mergers does not yield much insight on merger waves that consist of vertical or conglomerate mergers.

At the other end, non-strategic theories of merger waves emphasise the effects of exogenous factors such as deregulation, globalisation or the introduction of new technologies. In this context, merger waves are characterised by the fact that it is not the merger activity of other firms per se that induces a firm to merge, but rather an exogenous shift in the economic environment that simultaneously makes all mergers attractive. For example, Gort (1969) and Mitchell and Mulherin (1996) report evidence that M&A activity is significantly correlated with technological shocks and generally with disturbance to the economy or a specific industry. In line with these findings, Jovanovic and Rousseau (2002) show that bursts in merger activity may follow from technological shocks as assets are reallocated from less efficient targets to more efficient acquirers. In this view, a merger wave is the effect of inefficiencies caused by exogenous shifts in the economic environment.[3]

In practice, both strategic and non-strategic elements seem to play an important role in creating merger waves. This calls for new theory that encompasses both features.

According to Weston, Hoag and Chung (1990),

[3]See also Faria (2002) for interesting work building on Jovanovic and Rosseau (2002).

> A complete theory of mergers should have implications on the timing of aggregate merger activity. As the matter stands, there does not exist an accepted theory which simultaneously explains motivations behind mergers, characteristics of acquiring and acquired firms, and the determinants of the levels of aggregate merger activity.

In other words, any merger theory should be tested by its implications for the time series properties of M&A activity. Therefore, a satisfactory merger theory should be able both to explain the motivation behind mergers and to generate the observed wave patterns, as a result of exogenous factors and strategic interaction. This is an ambitious programme, and will not be pursued here. Instead, I will present an abstract model of the timing of mergers that broadly corresponds to with empirical findings.

Last, I should like to echo the call of Weston et al. (1990) for further research on mergers to be carried out, especially on the interaction of macroeconomic variables and industry and firm specific factors in creating value through mergers.

CHAPTER 4

A Theory of Strategic Merger Waves

1. Introduction

In this chapter, I propose a stochastic preemption model of merger activity in which waves occur as an equilibrium phenomenon. The underlying economic fundamental determining merger profitability is modeled as an exogenous process, but merger waves occur as a result of strategic interaction.

The model builds on three simple blocks. First, I pose that there is relative *scarcity* of potential desirable targets. This is a plausible assumption, given that there are often multiple suitors for specific targets. As a practical matter, there is usually no problem in distinguishing between potential targets and acquirers, where the identities of the acquirer and the target are determined by some notion of size, e.g. capacity, market share or market capitalisation.[1] For example, in the world airline industry, there is a natural distinction between European and North American airlines. There is also a sense among the latter that potential European targets for takeovers or strategic alliances are scarce. Note that the interpretation of scarcity of targets need not be literal. An alternative interpretation is that the targets own or control scarce resources or

[1] There is an extensive finance literature that distinguishes acquirers from targets along values of Tobin's q. See e.g. Chappell and Cheng (1984), Lang, Stulz and Walkling (1989), Servaes (1991), Bittlingmayer (1996), Shelton (2000) and Andrade, Mitchell and Stafford (2001). Moerck, Shleifer and Vishny (1988) furthermore study characteristics of targets of hostile and friendly takeovers.

assets. Such assets could be access to restricted (geographical) markets, existing costumer bases, patents, business practices or as in the airlines example, landing slots in key European hubs. Last, one may consider a target population ranked according to some quality index, such as the ease with which the target can be successfully merged with an acquiring firm. Competition would then start for the set of high quality targets, with lower quality targets being competed over in the future.

The second feature driving my model is that there is a *value of delay*, i.e. there is an *options value* in waiting to acquire a target, at least over a non-trivial period of time. Several circumstances can give rise to such a value of delay, as explained by the real options literature.[2] For example, time may allow firms to look for the best fit; or it may be that the returns from the merger are realised in the future (when new markets are created), whereas implementation costs are borne immediately after the merger. Also, technological progress or convergence of hitherto separate industries may make it optimal not to merge straight away. Last, waiting may be valuable in resolving uncertainty.[3]

The third crucial assumption is that competition for targets is imperfect. Specifically, what is ruled out is competition à la Bertrand with homogeneous

[2]The view of mergers as an investments is not new, and can be traced back to Mueller (1969) and Bittlingmayer (1996).

[3]Generally, an options value of delay can arise both because of growth in the variable determining profitability, or because the evolution of the variable is stochastic. While growth is the appropriate assumption in some cases, it is somewhat restrictive since one then has to explain the sources of this growth. Assuming uncertainty about the economic fundamental seems more palatable in general. In order to use the real options framework, the investment to be undertaken has to be at least partially irreversible. In the case of a merger, some degree of irreversibility is clearly present, as substantial resources are required both for implementing the merger and for later divestiture. For a more thorough exposition, see e.g. McDonald and Siegel (1986), Dixit and Pindyck (1994) and reference therein.

products, where all rents are dissipated. If there was a perfectly functioning price mechanism, it would "punish" a surge in demand by increasing the price level accordingly. In general, imperfections in the price mechanism can arise because of private information, target management idiosyncrasies or agency problems.

The trade-off between a value of delay and competitive considerations has previously been identified in the literature. For example, Smith and Triantis (1995) point out that

> In the case of acquisitions in an environment characterized by an absence of competition, a firm may delay its decision to acquire while waiting for more resolution of uncertainty regarding market conditions and other economic factors such as interest rates. However, since competition for specific targets is often significant, firms in practice may not be able to wait indefinitely to acquire a target, but must instead react quickly at the right time.

I first consider a complete information stochastic model where a measure of raiders compete over time for a smaller measure of targets. I show that there exists a continuum of subgame perfect equilibria. In all equilibria, *all potential acquiring firms raid the target firms simultaneously,* a feature that may be interpreted as a merger wave. The intuition for this type of equilibrium is simple. While waiting is optimal when all other firms wait, fear of being stranded without a firm to merge with can lead firms to attempt a preemptive takeover. This in turn vindicates the belief that there will be a merger wave, thus leading all firms to raid.

Although all equilibria share the same qualitative features, multiplicity is problematic. To resolve the multiplicity, I extend the model to a dynamic global game by introducing incomplete information. This is achieved by letting acquirers receive slightly imperfect private information about the realisations of the randomly evolving economic fundamental variable. In this setting, it is shown that there exists a *unique perfect Markovian Bayesian equilibrium in monotone strategies.*

Methodologically, the present model has features in common with several strands of literature. First, the technique employed to solve for equilibrium in the incomplete information extension of the model is the *global games* approach, first introduced by Carlsson and van Damme (1993) and subsequently developed by Morris and Shin (1998, 2002). This literature has revisited a large body of the theory of coordination games, such as models of currency attacks and bank runs, and has shown that multiplicity of equilibria may not be robust to the introduction of incomplete information. While there has been some work on dynamic global games in the literature (see e.g. Chamley, 1999, 2002, Frankel and Pauzner, 2000, Burdzy et al., 2001, Levin, 2001 and Oyama, 2001), it takes a different approach to the one adopted here. Specifically, and in contrast to existing work, I will deal with infinitely lived decision makers who are not restricted by randomly arriving revision opportunities or by a pre-specified order of moves.

Second, my model has elements in common with the literature on real options games. These are games in which decision makers choose the timing of their actions as some random variable evolves (as in the real options literature), but there is also an element of competition between players (see Grenadier, 2000).

Finally, there is a large literature on temporal clustering as equilibrium phenomena. I will briefly contrast my findings with the main models of that literature. First, much work has been devoted to theories of rational herding, following Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992) and Gul and Lundholm (1995). In these models, agents make decisions based on private information and on the observed behaviour of other agents, and thus herding is the effect of informational externalities. In contrast, in my model there are no informational externalities, as observing other players' past behaviour has no useful informational content.

Second, my model differs from Bulow and Klemperer's (1994) model of rational frenzies in several respects. In their private values setting, both informational externalities and strategic considerations are present. What creates frenzies in their model is that players' actions are informative about their private values, and thus about aggregate demand at a given price. Since information is released unevenly over time, demand displays radical shifts. My model resembles theirs in that demand externalities play an important role in creating merger waves. However, in my model all bidders value the targets equally.

More importantly, merger waves are shown to happen in a decentralised market, whereas Bulow and Klemperer (1994) consider a monopolist whose optimal pricing rule creates excess demand. In a sense, the equilibria of my model are better characterised as a *stampede* ("sudden mass movement") rather than as a *frenzy* ("brief delirium that is almost insanity").

Third, the results in this paper differ from those of Fudenberg and Tirole (1985). They consider a perfect information setting in which agents contemplate when to act. In their setting, it is not possible to rule out multiple equilibria. Furthermore, they show existence of diffusion equilibria, in which agents' actions are dispersed over time. In the complete information version of my model, no such diffusion equilibria exist. This is partly due to the symmetry of agents, and the fact that late movers receive zero payoffs. In the limit of the incomplete information version of my model, as noise becomes negligible, there is no dispersion either. With non-negligible noise, however, it is well possible to have equilibria in which not all acquirers raid simultaneously. Finally, the predictions of my model rely on the assumption that there is no rent equalisation in equilibrium, a result that follows from the assumption that the price mechanism is imperfect. In Fudenberg and Tirole's analysis of a duopoly game, each player can react immediately to the rival's action, an assumption that in effect implies the outcome of perfect competition.

2. The Model

Time is discrete and indexed by the non-negative integers $t = 0, 1, 2, \ldots.$ There is a continuum of *targets* and a continuum of *acquirers* with unit demand for a target. All acquirers are risk neutral, and discount the future with the common factor $\delta \in]0, 1[$. In every period, each acquirer faces the choice between *raiding* and *waiting.* An acquirer who waits remains inactive until the next period.

For every period $t = 0, 1, 2, \ldots$, let x_t and y_t denote the measures of remaining targets and acquirers respectively, and $z_t \in [0, y_t]$ the measure of acquirers who choose to raid (*raiders*). Denote by X_t, Y_t and Z_t the sets of targets, acquirers and raiders respectively. Once an acquirer decides to raid, he participates in an *allocation game* $A_t : Z_t \times X_t \times \mathbb{R} \rightarrow \mathbb{R}$ with von Neumann-Morgenstern expected payoff $R(z_t, x_t, \theta_t)$.[4] The single dimensional variable θ_t represents some economic fundamental that determines merger profitability.

The expected payoff $R(z_t, x_t, \theta_t)$ from participating in the allocation game is called the *raiding value.* It is assumed to be a bounded function of θ_t and is to be thought of as the expected value of obtaining, through some bidding process, an infinite flow of future profits. The expected *waiting value* is given by the option to raid in future periods, and thus given by the recursive expression

$$W(z_t, x_t, \theta_t) = \delta E_t \max \left[R(z_{t+1}, x_{t+1}, \theta_{t+1}), W(z_{t+1}, x_{t+1}, \theta_{t+1}) \right]$$

[4] It is implicitly assumed that there exists a unique equilibrium in the allocation game.

Note that since an acquirer always has the option of waiting indefinitely, it follows that $W(z_t, x_t, \theta_t) \geq 0$ for all t. Finally, the *net waiting value* $\Delta(z_t, x_t, \theta_t)$ is defined as

$$\Delta(z_t, x_t, \theta_t) \equiv W(z_t, x_t, \theta_t) - R(z_t, x_t, \theta_t)$$

If for any values of the variables, $\Delta(z_t, x_t, \theta_t) < 0$, raiding is the dominant strategy, while waiting is dominant for $\Delta(z_t, x_t, \theta_t) > 0$.

Next, I make the following assumptions. First, I assume that the initial measure of acquirers is larger than that of targets, so $y_0 > x_0$. This captures the notion that targets are scarce.[5] Second, it is assumed that an acquisition is irreversible and that the economic fundamental evolves randomly over time. The process $\{\theta_t\}_{t=0}^{\infty}$ is assumed to be Markov and such that the random variable θ_t is stochastically increasing, that is a higher realisation of θ_t today shifts the distribution of future realisations according to first-order stochastic dominance. Denote by $G(\theta_t \,|\theta_{t-1})$ the cumulative distribution of θ_t. Stochastic dominance amounts to the requirement that for $\theta_{t-1} > \theta'_{t-1}$, $G(\theta_t \,|\theta'_{t-1}) > G(\theta_t \,|\theta_{t-1})$. Irreversibility and persistence in the evolution of the fundamental yield a value of delay, as is standard in the real options literature. Third, it is assumed that the competition for targets is imperfect. This allows me to impose the condition that the raiding value $R(z_t, x_t, \theta_t)$ is strictly increasing in θ_t. Fourth, the raiding

[5] The assumption that targets are scarce is also made by Klemperer (1997), Bulow, Huang and Klemperer (1999), Norbak and Persson (2001, 2002), Rhodes-Kropf and Viswanathan (2002) and many others and is virtually standard in the corporate finance literature on "takeovers as auctions". Conceivably, this assumption could be dispensed with by also dispensing with the assumption of unit demand.

value $R(z_t, x_t, \theta_t)$ is assumed to be weakly decreasing in z_t and weakly increasing in x_t, with $R(z_t, 0, \theta_t) = 0$ for all θ_t. The variable z_t measures the degree of competition for targets, and thus the expected payoff from participating in the allocation game is reasonably assumed to be weakly decreasing in this variable. The variable x_t measures the stock of remaining targets. Thus higher x_t relaxes competition, which weakly increases the value of raiding. This specification of payoffs is quite general and allows us to work with several different explicit allocation games.[6] I do not require that the expected value be strictly decreasing in the measure of raiders (although it will be so for some range in the allocation game presented later), only when the measure of raiders is above some critical measure. When no targets remain, the game is over and raiding is no longer an option. Finally, it is assumed that no raider gets rationed in the allocation game as long as $z_t < x_t$. This implies that the laws of motion for the endogenous state variables are

$$x_t = \max\left\{x_0 - \sum_{r=0}^{t-1} z_r, 0\right\}$$

$$y_t = \max\left\{y_0 - \sum_{r=0}^{t-1} z_r, y_0 - x_0\right\}$$

since a raider who obtains a target leaves the game.[7] This assumption is without loss of generality. All that is needed is that the expected measure of successful raiders, say $\widetilde{z}_t$, is strictly increasing in the measure of raiders z_t. This couple

[6]One such model will be shown in Section 6.

[7]It is also implicitly assumed that a raider who is not rationed in the allocation game actually goes through with the takeover and leaves the game.

of identities captures the main strategic element in the interaction between acquiring firms. The higher the measure of raiders in any given period, the scarcer targets become in future periods, thereby eroding the options value of waiting.[8]

Assume throughout that $x_t > 0$ and denote by $z^t \equiv \{z_s\}_{s=t}^{\infty}$ a sequence of current and future measures of raiders. Also, let $z^t \geq \widehat{z}^t$ if $z_s^t \geq \widehat{z}_s^t$ for all $s \geq t$. Define the following:

DEFINITION 1. *(Merger Triggers)*

(i) ***Marshallian Trigger:*** $\underline{\theta} \equiv \inf\{\theta_t : R(z_t, x_t, \theta_t) \geq 0\}$

(ii) ***Strategic Trigger.*** $\widetilde{\theta}(z^t) \equiv \inf\{\theta_t : \Delta(z_t, x_t, \theta_t) \leq 0\}$

(iii) ***First-Best Trigger:*** $\overline{\theta}(z^t) \equiv \inf\{\theta_t : \Delta(z_t, x_t, \theta_t) \leq 0 \textit{ for } z_s = 0, s \geq t\}$

That is, $\underline{\theta}$ is the lowest value of θ_t at which the raiding value is non-negative. The term Marshallian trigger is borrowed from the real options literature. Next, the strategic trigger $\widetilde{\theta}(z^t)$ is the lowest value of θ_t such that raiding in the current period dominates waiting, given a sequence of current and future measures of raiders. Last, the first-best trigger $\overline{\theta}(z^t)$ is the lowest value of θ_t such that, even in the absence of competitive pressure, delaying a takeover one period further

[8]Thus the model has strategic complementarities, as defined by Bulow, Geanakoplos and Klemperer (1985) in the sense that increased aggregate merger activity increases the incentive to merge. Gale (1995) studies a dynamic coordination game with features similar to this model (but where players' interests are perfectly aligned). In his terminology, the present model has intertemporal strategic complementarities, through negative spill-over effects.

is not optimal.[9] Of course, $\overline{\theta}(z^t)$ is just the strategic trigger $\widetilde{\theta}(z^t)$ evaluated at z^t with $z_s = 0,\ s \geq t$. Trivially, $\widetilde{\theta}(z^t) \in [\underline{\theta}, \overline{\theta}(z^t)]$ and $\widetilde{\theta}(z^t) \rightarrow \overline{\theta}(z^t)$ as $x_t \rightarrow y_t$.

The following lemmata simply state that the above merger triggers exist and are well defined.

LEMMA 1. *(Single Crossing in θ_t). For any sequence z^t there exists a unique $\widetilde{\theta}(z^t) \in [\underline{\theta}, \infty[$ such that $\Delta(z_t, x_t, \theta_t) > 0$ for $\theta_t < \widetilde{\theta}(z^t)$, $\Delta(z_t, x_t, \theta_t) = 0$ for $\theta_t = \widetilde{\theta}(z^t)$ and $\Delta(z_t, x_t, \theta_t) < 0$ for $\theta_t > \widetilde{\theta}(z^t)$. Furthermore, $\widetilde{\theta}(z^t)$ is weakly increasing in x_t and weakly decreasing in z_t and z^t.*

PROOF. See appendix A.1 ∎

LEMMA 2. *(Single Crossing in z_t). For any sequence z^{t+1} and $\theta_t \in [\underline{\theta}, \widetilde{\theta}(z^{t+1})]$ there exists a unique $z_t^* \in [x_t, y_t]$ such that $\Delta(z_t, x_t, \theta_t) > 0$ for $z_t < z_t^*$, $\Delta(z_t, x_t, \theta_t) = 0$ for $z_t = z_t^*$ and $\Delta(z_t, x_t, \theta_t) < 0$ for $z_t > z_t^*$. Furthermore, z_t^* is weakly increasing in x_t and weakly decreasing in θ_t.*[10]

PROOF. See appendix A.1 ∎

These results have a straightforward interpretation. First, given any level of future and present competition, the expected value for an acquirer contemplating wether to raid or wait increases in the economic fundamental. In fact, both the value of raiding and waiting increase. But as the economic fundamental

[9]It is assumed that both $\underline{\theta}$ and $\overline{\theta}$ are independent of x_t and y_t as long as $x_t > 0$. This is esentially without loss of generality and simplifies the subsequent analysis.

[10]*If $[\underline{\theta}, \widetilde{\theta}(z^{t+1})] = \emptyset$ for all z^{t+1} then for $\theta_t \geq \underline{\theta}$, $\Delta(z_t, x_t, \theta_t) \leq 0$ for all $z_t \in [0, y_t]$. It is thus assumed throughout that $[0, \widetilde{\theta}(z^{t+1})] \neq \emptyset$ for some z^{t+1}.*

increases, the options value of delay is eroded, and ultimately the raiding value overtakes the value of waiting. Similarly, given a level of the economic fundamental, an increase in the measure of raiders reduces the measure of future targets, thereby increasing the opportunity cost of postponing a merger.

In the analysis that follows, the fundamental variable θ_t will play a prominent role, and thus deserves some further comments. In keeping with the empirical observation that exogenous variables significantly influence the merger decision, θ_t is taken to be any variable determined at aggregate level, such as technological progress, interest rates, a stock market index or other macroeconomic variable.

It should be noted that in the current analysis, the value of raiding is interpreted as a flow, which is a function not only of the current realisation, but also of the future evolution of the economic fundamental. In other words, the value of being merged remains subject to random fluctuations in the economic environment. With this interpretation, the assumption of persistence is crucial, as is the assumption of irreversibility. With irreversibility, the acquirer must be confident that the value of being merged is not likely to disappear. This is assured by assuming persistence, which in effect makes the value function monotone in the economic fundamental. Many different models can lead to a reduced form like the one employed here. For example, θ_t can parameterise a flow which the winner of the allocation process obtains at a further sunk nonstochastic cost, or the flow profit can be constant but maintenance costs evolve stochastically. Last, both the profit flow and the costs can evolve according to different processes, in which case the relative evolutions of the two processes determine merger profitability. Such frameworks are discussed in McDonald and Siegel (1986). What is important for the present model is that θ_t is a sufficient statistic for the value of raiding, given a specific level of competition. One may think of the raiding value as a composite function such that $R(z_t, x_t, \theta_t) = \zeta(z_t, x_t, \theta_t, r(\theta_t))$ where $r(\theta_t)$ is a suitably discounted infinite stream of profits, and the function ζ is a function describing the outcome of the allocation process, i.e. who gets the targets and what they pay.

There is another interpretation of the raiding value which fits into the current framework and requires somewhat weaker assumptions. Namely, that in which the value of a merger is a prize to be collected once and for all, which does not depend on the future evolution of the economic fundamental. In this setup, the assumption of persistence is not needed. For example, shocks that are identically and independently distributed over time can be handled and is actually easier to analyse. Also, the requirement that the underlying process be Markovian can be relaxed. For an excellent treatment of optimal stopping problems of this type see Chow, Robbins and Siegmund (1991).

Last, and in contrast to the scenario with stock interpretation, a more general model would have the flow value of being merged be a function of the future merger activity of rivals. For simplicity, and to be consistent with the fact that product market interaction is not considered, it is assumed in the present analysis that after a takeover, a raider's payoffs are independent of future takeover activity. The more general setting could easily be analysed along the same lines as the model considered here. This would involve assuming that non-raiding acquirers receive flow payoffs that are lost when they raid, but further monotonicity assumptions would need to be imposed. This added generality would not add much to the analysis and the qualitative features of the current model would remain unchanged. Specifically, it would just have the effect of raising the opportunity cost of a merger, thus delaying the merger wave.

Throughout the paper, it will be assumed that past realisations of the fundamental variable are commonly known. This is thus public information, and serves in forecasting the future evolution of the variable. In the two scenarios to be analysed, the difference lies in what is assumed about current information about the fundamental. In the complete information benchmark, common knowledge is assumed. That is, all players are assumed to have only public information at their disposal, and fully know the current realisation of θ_t. In the incomplete information extension, the assumption of common knowledge is relaxed. Specifically, it is assumed that the players have both public information (i.e. history) and imperfect private information about the current state of the fundamental.

3. The Complete Information Game

For the complete information game[11] it is assumed that both past and current realisations of the economic fundamental θ are common knowledge. Let

[11]Note that the analysis of the complete information game does not in any crucial way depend on the continuum player assumption or on symmetry between the acquirers.

$h_t = (\theta_0, ..., \theta_{t-1}; z_0, ..., z_{t-1})$ denote history at time t and H_t the set of all possible histories at time t. In this setting, a monotone strategy is defined as follows:

DEFINITION 2. *A monotone Markovian strategy is a mapping* $a : \mathbb{R}^{2t} \times \mathbb{R} \longrightarrow \mathbb{R}$ *such that* $(h_t, \theta_t) \longmapsto a(h_t, \theta_t) \equiv k_t$

Thus, for a history h_t and current fundamental θ_t, a strategy picks a real number k_t with the interpretation that an agent raids whenever $\theta_t \geq k_t$ and waits whenever $\theta_t < k_t$. Given a strategy k_t, the chosen action will thus be

$$I_{k_t}(\theta_t) = \begin{cases} 1 & \text{for} \quad \theta_t \geq k_t \\ 0 & \text{for} \quad \theta_t < k_t \end{cases}$$

where 1 stands for raid and 0 stands for wait and $I_{k_t}(\theta_t)$ denotes the indicator function. With this definition in place, the following result can be stated:

PROPOSITION 1. *(Merger Waves) For any sequence* z^{t+1} *and history* $h_t \in H_t$*, any cutoff* $k_t \in [\underline{\theta}, \widetilde{\theta}(z^{t+1})]$ *constitutes an equilibrium strategy. Furthermore, there exist no equilibria in asymmetric strategies.*

PROOF. The first part of the proposition follows immediately from the definitions. To see the second part, fix a sequence of future cutoff strategies $\{k_s\}_{s=t+1}^{\infty}$ (and thus z^{t+1}) and consider period t. Let a measure m^i, $i = a, b$ use strategies with cutoff k_t^i with $k_t^a < k_t^b$ and $m^a + m^b = y_t$. Recall that a cutoff k_t is only an equilibrium strategy if it is optimal for *any* realisation of the economic fundamental θ_t. For $\theta_t \in [\underline{\theta}, k_t^a]$ all acquirers wait and the asymmetric strategies can co-exist in equilibrium. Similarly, for $\theta_t \in [k_t^b, \overline{\theta}]$ all acquirers raid, which is also an equilibrium. Now consider the case where $\theta_t \in [k_t^a, k_t^b]$. Given the considered strategies, a realisation in this range prompts a measure m^a to raid and a measure m^b to wait. If $m^a \geq z_t^*$, the cutoff k_t^b cannot be an equilibrium strategy. Similarly, if $m^b < z_t^*$, the cutoff k_t^b cannot be an equilibrium strategy. Thus equilibria in asymmetric strategies do not exist ∎

This continuum of equilibrium cutoffs is Pareto ranked, with higher cutoffs dominating lower ones. It follows that even the Pareto efficient subgame perfect equilibrium is strictly (socially) inefficient.[12]

COROLLARY 1. *(Indeterminacy) A merger wave may happen in any period t where $\theta_t \in [\underline{\theta}(x_t, 0), \widetilde{\theta}(z^{t+1})]$.*

PROOF. Follows immediately from the proposition ∎

As usual in this type of models, there is no clear way in which to determine the equilibrium outcome. The only thing that can be said at this point is the earliest and latest time at which a rush can occur. It has been argued (see

[12]The game also has an equilibrium in mixed strategies. To characterise this equilibrium, recall that by definition, a raid of measure z_t^* leaves a bidder exactly indifferent between raiding and waiting. This yields a straightforward expression for the mixed strategy equilibrium. Assume that all bidders at time t raid with probability p_t. The measure of raiding bidders is thus simply given by $p_t y_t$. This in turn implies that, in the mixed strategy equilibrium, bidders raid with probability $p_t^* = z_t^*/y_t$. It follows trivially that when z_t^* is undefined, the mixed strategy equilibrium is degenerate, i.e. the unique pure strategy equilibrium is to raid.

e.g. Fudenberg and Tirole, 1985 and a related discussion in Carlsson and van Damme, 1993) that Pareto optimal equilibria should take precedence over other equilibria on the grounds that players "should be" able to coordinate on good outcomes, and that these thus be adequate "predictions" of equilibrium play. Be that as it may, the fact remains that without some explicit theory to guide the selection of equilibrium, model-based predictions are a delicate matter. As is well known, this type of model leaves ample room for self-fulfilling beliefs and payoff-irrelevant sunspots, and is thus an inadequate vehicle for doing comparative statics/dynamics exercises. The problem is that without any knowledge of how equilibrium is reached, comparative analysis becomes very dependent on which equilibrium one takes as the benchmark, thereby inviting questions about the robustness of its predictions. In essence, what creates multiplicity is the assumptions of complete information and common knowledge. These assumptions imply that acquirers are perfectly able to predict rivals' behaviour in equilibrium. In practice, these assumptions seem hard to justify. In general one should expect at least some degree of informational differentiation. Therefore, I will now enrich the model by assuming incomplete information. This assumption will have radical implications for the equilibrium set. Namely, it will be shown that once incomplete information is introduced, a unique Markovian perfect Bayesian equilibrium in monotone strategies exists.

4. The Incomplete Information Game

The model is now extended to a dynamic global game by assuming that the realisation of the fundamental variable θ_t is no longer common knowledge. Let G_t denote the distribution of θ_t, conditional on past information, and denote by g_t the corresponding probability density function.

The next step is to specify acquirers' information. Each acquirer i receives a private signal $s_{it} = \theta_t + \sigma\varepsilon_{it}$, where the noise $\varepsilon_{it} \sim F$ (with probability density function f) is identically and independently distributed over time and across acquirers, and $\sigma > 0$ measures the precision of the signal. The distribution of noise F is assumed to be such that the variables θ_t and $\{s_{it}\}_{i \in Y_t}$ are affiliated, and thus satisfy the monotone likelihood ratio property.[13] That is, an increase in either variable shifts the distribution of the other in the sense of first-order stochastic dominance. This fact is important, as signals are not only used to estimate θ_t, but also to make inferences about other acquirers' signals.

In this setting, a monotone Markovian strategy is a mapping $(h_t, s_t) \longmapsto a(h_t, s_t) \equiv k_t$, and actions given by

$$I_{k_t}(s_t) = \begin{cases} 1 & \text{for} \quad s_t \geq k_t \\ 0 & \text{for} \quad s_t < k_t \end{cases}$$

where 1 denotes *raid* and 0 denotes *wait* and $I_{k_t}(s_t)$ denotes the indicator function. Thus, the choice variable is the cutoff level k_t. Given these strategies,

[13] For a formal definition and further results on affiliated variables, see Milgrom and Weber (1982).

the measure of raiders for given cutoff k_t is determined by the distribution of signals. But given a realisation of θ_t, the distribution of signals is given by F. One can use this to express the measure of raiders as $z_t = y_t \left[1 - F(k_t \,|\theta_t)\right]$.

Since the game is dynamic, the optimal action at any given point in time will depend on competitors' play in subsequent periods. Thus, players must forecast other players' actions, which in turn implies that players must forecast other players' forecasts.

It is important to realise that history only serves to the extent that it yields a prior belief G_t on θ_t. Past actions have no useful informational content, and only feed through to the current decisions through their influence on the state variables x_t and y_t.

As already discussed, a drawback of the complete information model is the multiplicity of equilibria. This feature of the model is eliminated once incomplete information is introduced. Specifically, it will now be shown that under the maintained assumptions, there exists a unique Markovian perfect Bayesian equilibrium in monotone strategies.

To prove this, the following four step procedure will be followed. First, the infinite horizon game is truncated to obtain a finite horizon game, denoted by $\Gamma(T) \equiv \{\Gamma_t\}_{t=0}^{T}$. Second, due to the recursive structure of $\Gamma(T)$, it is possible, for all t, to associate Γ_t with a simplified (*associated*) static game Γ_t^*, for which uniqueness can be shown by using the techniques developed for static games by Frankel, Morris and Pauzner (2000), and Morris and Shin (2002). This association is achieved by showing that in each period t, the function constituting the waiting value in Γ_t is well defined as a function of current information and actions. By solving the game backwards, the players are faced with an essentially static problem in each period. The third step is to show that the truncated (*underlying)* game $\Gamma(T)$ (where payoffs depend on the realisation of the state and where history is informative about the distribution of this period's realisation) converges uniformly to the sequence of associated games $\Gamma^*(T) \equiv \{\Gamma_t^*\}_{t=0}^T$, as noise becomes negligible. This implies that the underlying truncated game $\Gamma(T)$ has a unique perfect Bayesian equilibrium in cutoff strategies. The last step is to show that equilibria of the infinite horizon game $\Gamma(\infty)$ are well approximated by equilibria of the truncated game $\Gamma(T)$, as the horizon T tends to infinity.

Using the terminology of Morris and Shin (2002), the problem to be solved in each period satisfies *action single crossing* (established in Lemma ??), *state monotonicity* (established in Lemma ??), *limit dominance* (follows from the existence of $\underline{\theta}$ and $\overline{\theta}(z^t)$) and a *monotone likelihood ratio property* of the signal distribution (follows from the assumption of affiliation). Under these and some further continuity conditions, Morris and Shin (2002) show that there exists a unique Bayesian equilibrium in monotone strategies. The key to Morris and Shin's proof is to realise that as noise becomes negligible, rival's actions are believed to be uniformly distributed. But given a uniform distribution of actions, there is generically a unique signal at which indifference obtains. Once the existence of an indifferent acquirer has been established, the last step is to verify that for signals below the cutoff, waiting is optimal while for signals above the cutoff, raiding is optimal. In the complete information game, it was proved that there are no equilibria in asymmetric strategies. This result carries over to incomplete information game and thus one can restrict attention to symmetric cutoffs k_t.

To gain intuition, a heuristic proof of uniqueness in the static setting due to Morris and Shin (2002) is now provided. Consider a simple static (one-shot) game in which the fundamental is drawn from a uniform distribution on the real line, and where acquirers care only about the signal and not about the fundamental. In choosing the optimal strategy based on available information, an acquirer must estimate the measure of rivals who choose to raid. To perform this estimate, the acquirer has to make inferences about the distribution of rivals' signals. Denote by $\Psi_{\sigma}^{*}(z; s, k)$ the probability that an acquirer observing signal s, assigns to a measure less than z of other agents observing signals above k. This probability is important since the acquirer's optimal action depends on the expected measure of rivals choosing each action. Given a realisation of the fundamental θ, the measure observing signals above k is less than z exactly when

$$\theta \leq k - \sigma F^{-1}(\frac{y-z}{y})$$

It follows that

$$\begin{aligned}\Psi_\sigma^*(z;s,k) &= \int_{-\infty}^{k-\sigma F^{-1}(\frac{y-z}{y})} \frac{1}{\sigma} f(\frac{s-\theta}{\sigma}) d\theta \\ &= \int_{s-k+\sigma F^{-1}(\frac{y-z}{y})}^{\infty} f(m)dm \\ &= 1 - F(s-k+F^{-1}(\frac{y-z}{y}))\end{aligned}$$

where $m = (s-\theta)/\sigma$. Last, setting $s = k$ yields

$$\Psi_\sigma^*(z;s,k) = \frac{z}{y}$$

This is the cumulative distribution function of z when distributed uniformly on $[0, y]$, with corresponding probability density function $1/y$. Now, if an acquirer believes his opponents' actions to be uniformly distributed, there is generically a unique signal at which indifference obtains. In other words, in this setting there is a unique cutoff such that an acquirer receiving a signal below the cutoff waits, while an acquirer receiving a signal above the cutoff raids.

The next step is to show that with the general priors assumption, beliefs converge to those held under a uniform prior, as signals become very precise. To this end, consider the case where the past is informative about today's realisation. Let the prior distribution of θ be given by the probability density function g. The probability $\Psi_\sigma(z;s,k)$ that an acquirer observing signal s assigns to a measure smaller than z observing signals higher than k is given by

Bayes' rule as

$$\Psi_\sigma(z; s, k) = \frac{\int_{-\infty}^{k-\sigma F^{-1}(\frac{y-z}{y})} g(\theta) f(\frac{s-\theta}{\sigma}) d\theta}{\int_{-\infty}^{\infty} g(\theta) f(\frac{s-\theta}{\sigma}) d\theta}$$

Setting $m = (s - \theta)/\sigma$ yields

$$\Psi_\sigma(z; s, k) = \frac{\int_{\frac{s-k}{\sigma}+F^{-1}(\frac{y-z}{y})}^{\infty} g(s - m\sigma) f(m) dm}{\int_{-\infty}^{\infty} g(s - m\sigma) f(m) dm}$$

Now, let $u = (s - k)/\sigma$ to obtain

$$\Psi_\sigma(z; s, s + \sigma u) = \frac{\int_{u+F^{-1}(\frac{y-z}{y})}^{\infty} g(s - m\sigma) f(m) dm}{\int_{-\infty}^{\infty} g(s - m\sigma) f(m) dm}$$

As $u \to 0$,

$$\Psi_\sigma(z; s, s + \sigma u) \to 1 - F\left(u + F^{-1}\left(\frac{y-z}{y}\right)\right) = \frac{z}{y}$$

This is just the cumulative distribution function of z uniformly distributed on $[0, y]$. In effect, when signals become very precise, it does not matter what the prior is; nor does it matter if the players' payoffs depend directly on the signals or on the realisation of the fundamental. Summing up, the beliefs held under general priors converges to those held under a uniform prior as noise vanishes. But since the uniform prior game is known to have a unique equilibrium, one can conclude that so does the game with general priors.

The proof of existence of a unique perfect Bayesian equilibrium in monotone strategies proceeds in a similar fashion to that of the heuristic proof. Begin by assuming that the waiting value is well defined as a function of current information and strategies. That this is indeed the case will be verified shortly. Consider the associated game $\Gamma_t^*(T)$, where it is assumed that the received signal is a sufficient statistic of the state and θ is drawn from a uniform distribution on the real line. The posterior of θ will be centered on the signal, and distributed according to F (the distribution of noise in the signal). Letting $\Delta_\sigma^*(s,k)$ denote the expected payoff gain to "waiting", after having received signal s and believing that all other players use strategies with cutoffs k, it follows that

$$\Delta_\sigma^*(s,k) \equiv \int_{-\infty}^{\infty} \Delta\left(y\left[1 - F(\frac{k-\theta}{\sigma})\right], x, s\right) \frac{1}{\sigma} f(\frac{s-\theta}{\sigma}) d\theta \tag{4.1}$$

This function is just the expectation of the net "waiting" value, conditional on the received signal. For comparison, consider the underlying game at time t, and denote by $\Delta_\sigma(s_t, k_t)$ the expected payoff gain to waiting when signal s_t has been observed and all other players use cutoffs k_t. This is given by

$$\Delta_\sigma(s_t, k_t) \equiv \frac{\int_{-\infty}^{\infty} g(\theta_t) f(\frac{s_t-\theta_t}{\sigma}) \Delta\left(y_t\left(1 - F\left(\frac{k_t-\theta_t}{\sigma}\right), x_t, \theta_t\right)\right) d\theta_t}{\int_{-\infty}^{\infty} g(\theta_t) f(\frac{s_t-\theta_t}{\sigma}) d\theta_t} \tag{4.2}$$

where g is the prior density on θ. This is essentially the expected net waiting value where expectations are taken over possible realisations of θ_t by using Bayes' formula to generate the posterior beliefs conditional on the received signal s_t. The differences between (4.1) and (4.2) are twofold. First, in (4.1), the signal s replaces the economic fundamental θ. Second, the posterior distributions over the economic fundamental are generated by different prior beliefs. All other properties are shared.

Assume for now that the functions $\Delta_\sigma(s_t, k_t)$ and $\Delta^*_\sigma(s_t, k_t)$ are well defined, and that all players receiving identical signals would have identical beliefs about the exact shapes of the functions.

With these definitions in place, the following lemmata needed for the proof of Proposition 2 can be established:

LEMMA 3. *For any history $h_t \in H_t$, there exists a unique cutoff signal k^*_t in the associated static game Γ^*_t such that: $\Delta^*_\sigma(k^*_t, k^*_t) = 0$, $\Delta^*_\sigma(s_t, k^*_t) > 0$ for $s_t < k^*_t$ and $\Delta^*_\sigma(s_t, k^*_t) < 0$ for $s_t > s^*_t$.*

PROOF. See Appendix A.2 ∎

LEMMA 4. *For any history $h_t \in H_t$, as $\sigma \to 0$, $\Delta_\sigma(s_t, s_t - \sigma u_t) \to \Delta^*_\sigma(s_t, s_t - \sigma u_t)$ uniformly, where $u_t = (s_t - k_t)/\sigma$.*

PROOF. See Appendix A.3 ∎

Lemma 3 states that any static associated game has a unique Bayesian equilibrium. Lemma 4 shows that when private information is very precise, any finite version of the underlying game becomes arbitrarily close to a sequence of simplified static associated games.

Having established these results, the following proposition can be stated:

PROPOSITION 2. *As $\sigma \to 0$, for each sequence of realisations $\{\theta_t\}_{t=0}^{\infty}$, there exists a unique sequence of cutoffs signals $\{k^*_t\}_{t=0}^{\infty}$ such that for any history*

$h_t \in H_t$: $\Delta_\sigma(k_t^*, k_t^*) = 0$, $\Delta_\sigma(s_t, k_t^*) > 0$ *for* $s_t < k_t^*$ *and* $\Delta_\sigma(s_t, k_t^*) < 0$ *for* $s_t > s_t^*$.

PROOF. It is first established that $\Delta_\sigma(s_t, k_t)$ and $\Delta_\sigma^*(s_t, k_t)$ are well defined. Consider the truncated game, where play is exogenously terminated after some period T. At time T, optimality dictates that remaining acquirers raid for all signals that convince them of receiving a non-negative payoff. Note that this is irrespective of what other players do (i.e. independent of z_T). Now consider the (possibly trivial) decision at time $T-1$. Because equilibrium actions are well defined (and unique) at time T, the expected waiting value at time $T-1$ is well defined. But so then is the expected net waiting value $\Delta_\sigma(s_{T-1}, k_{T-1})$. The problem to be solved at time $T-1$ is essentially a static game as the one considered in the lemmata, and thus there exists a unique equilibrium with cutoff k_{T-1}^*. Having established uniqueness at time $T-1$, assume that at time $\tau < T-1$ there exists a unique sequence $\{k_t^*\}_{t=\tau+1}^T$ of equilibrium cutoffs. Our inductive hypothesis is that there then exists a unique equilibrium with cutoffs k_τ^*. The expected payoff gain from waiting is given by $\Delta_\sigma(s_\tau, k_\tau)$. This function shares all the properties of the function $\Delta_\sigma(s_{T-1}, k_{T-1})$ and thus there exists a unique equilibrium in monotone strategies with cutoff k_τ^*. Having shown uniqueness for arbitrary finite horizon version of the model, the infinite horizon game is considered. First note that the optimal strategy at any point in time optimally trades off the value of waiting with the value of raiding, i.e. the function $\Delta_\sigma(s_t, k_t)$. Clearly, $\Delta_\sigma(s_t, k_t)$ converges to a unique limit as $T \to \infty$, since both the value of raiding and that of waiting are bounded monotone functions of T. But then $k_t^*(T) \to k_t^*(\infty)$ as $T \to \infty$, where $k_t^*(T)$ is the equilibrium cutoff in period t in the game truncated after period T and $k_t^*(\infty)$ is the equilibrium cutoff at time t in the infinite horizon game[14] ∎

Recall that under complete information, there was a continuum of realisations of the economic fundamental which could constitute equilibrium cutoffs. The striking result of Proposition 2 is that under incomplete but very precise information, there exists only one equilibrium cutoff θ^* (which is of course a function of the state variables).[15] In 2×2 games, the equilibrium selected by the global perturbations approach coincides with Harsanyi and Selten's notion of risk-dominant equilibrium (see Carlsson and van Damme, 1993). For this reason, θ^* will be referred to as the *risk-dominant trigger.* When an acquirer observes a signal equal to the risk-dominant trigger, he is exactly indifferent between raiding and waiting. For higher signals, waiting is too risky; for lower signals, waiting is expected to yield higher payoffs.

It should be noted that while the analysis of the complete information game is robust to asymmetric acquirers and a relaxation of the continuum acquirer assumption, the analysis of the incomplete information game makes use of both of these assumptions. Morris and Shin (2002) show that in models with linear payoffs, the continuum player analysis carries over to the finite player setting.

[14]Since the raiding value is a discounted infinite flow, monotonicity of $\Delta_\sigma(s_t, k_t)$ in T is not immediate. If the value of raiding was a stock, monotonicity of $\Delta_\sigma(s_t, k_t)$ in T would be trivial, since the raiding value would be independent of the horizon while the waiting value is weakly increasing in T.

[15]Of course $\theta^* = \underline{\theta}$ cannot be excluded, in which case the equilibrium is degenerate.

It is still an open question if similar results hold for more general specifications of payoffs. As for asymmetries beyond the informational differentiation considered here, it is still not know if limit uniqueness holds for the family of quasisupermodular games to which the present one belongs. Frankel, Morris and Pauzner (2000) allow for heterogeneity in supermodular games and show limit uniqueness. Possibly, results along the lines of their work may extend to quasisupermodular games as well.

5. Comparative Analysis

Note that the proof of Proposition 2 exclusively considers monotone Markovian strategies in which acquirers choose cutoff levels k_t. It is still an open question if there exist equilibria in other strategies with general distribution functions.[16] Yet, monotone strategies are clearly the most natural class of strategies in this type of games.

Although the model presented here is quite general, it still allows for an identification of factors influencing the timing of mergers. Figure 1 illustrates how the different triggers are ordered and a sample path of the economic fundamental. The first variable of interest is the stock of target firms x_t. Ceteris paribus, a smaller measure of targets increases scarcity, thereby eroding the options value of delaying a takeover. The effect of lower x_t is thus to shift both the strategic and the risk-dominant triggers downwards. In the extreme case

[16]Goldstein and Pauzner (2000) are able to rule out existence of other types of equilibria assuming that the fundamental and noise are uniformly distributed. See also Morris and Shin (2002) for a discussion.

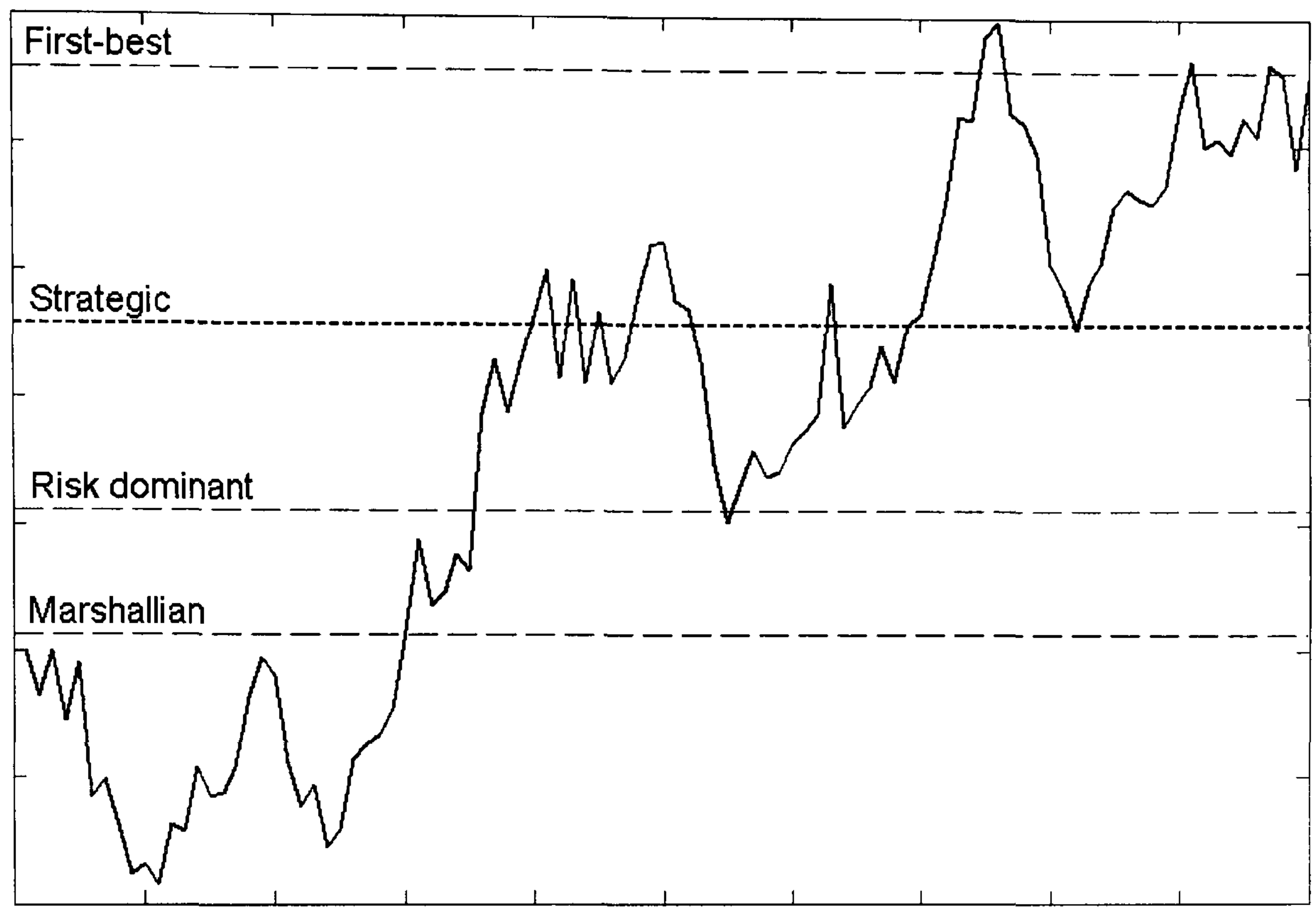

FIGURE 1. Triggers and sample path of θ_t.

where there are very few targets, one should expect an almost immediate rush, although this may not be identified empirically as a merger wave, since it involves very few takeovers. An increase in the measure of potential acquirers y_t has the exact opposite effect as a decrease in x_t.

Second, the evolution of the economic fundamental determining merger profitability has a direct implication for the expected timing of the wave. Higher growth or realisations closer to the strategic trigger shifts the risk-dominant trigger downwards. The model predicts that the higher the growth rate of the economic fundamental, the earlier the wave occurs. This is consistent with the stylised fact that M&A activity is particularly intense in industries with fast technological progress, such as those of software, hardware and pharmaceuticals.

Third, the exact way in which the expected payoffs $R(z_t, x_t, \theta_t)$ in the allocation game are determined, i.e. the way in which raiders and targets are matched and how the created surplus is divided, has implications for the timing of mergers. Many authors point to auctions theory when modeling takeover behaviour. While some argue that mergers seeking to exploit synergies are best viewed as private value auctions (see Fishman, 1988 and Burkart, 1995), there seems to be a consensus that most takeover contests, especially those that seek to replace inefficient management, resemble common-value auctions.[17] The comparative dynamics implications of different auction mechanisms is left for future research, but conceivably, the degree of competitiveness in some bidding games may be more sensitive than others to small changes in the measures of targets and acquirers.[18]

One comparative dynamics result that follows from the real options literature (see e.g. Dixit and Pindyck, 1994) is that the first-best trigger $\overline{\theta}(z^t)$ is increasing in the volatility of $\{\theta_t\}_{t=0}^{\infty}$. This is so because increased volatility increases the

[17]See discussion in Klemperer (1997), Cramton and Schwartz (1991) and Cramton (1998).

[18]The setup of the current model would fit naturally with a sequence of common-value auctions with endogenous participation and a reserve price. Harstad (1990) and Hausch and Li (1993) study common-value auctions with endogenous participation. See also related work by Engelbrecht-Wiggans and Weber (1979). In general, it is no mean feat to do comparative statics analysis with general distributional assumptions and affiliated values. Gordy (1998) shows numerically that under specific priors and signal distributions, it is indeed the case that expected payoffs to participating in a common-values auction is increasing in the value of the object and decreasing in the number of participants. His analysis also holds when reserve prices are introduced. In the present model, these are exactly the characteristics needed for uniqueness. Last, note that the equilibrium cutoff signal plays the same role as a reserve price.

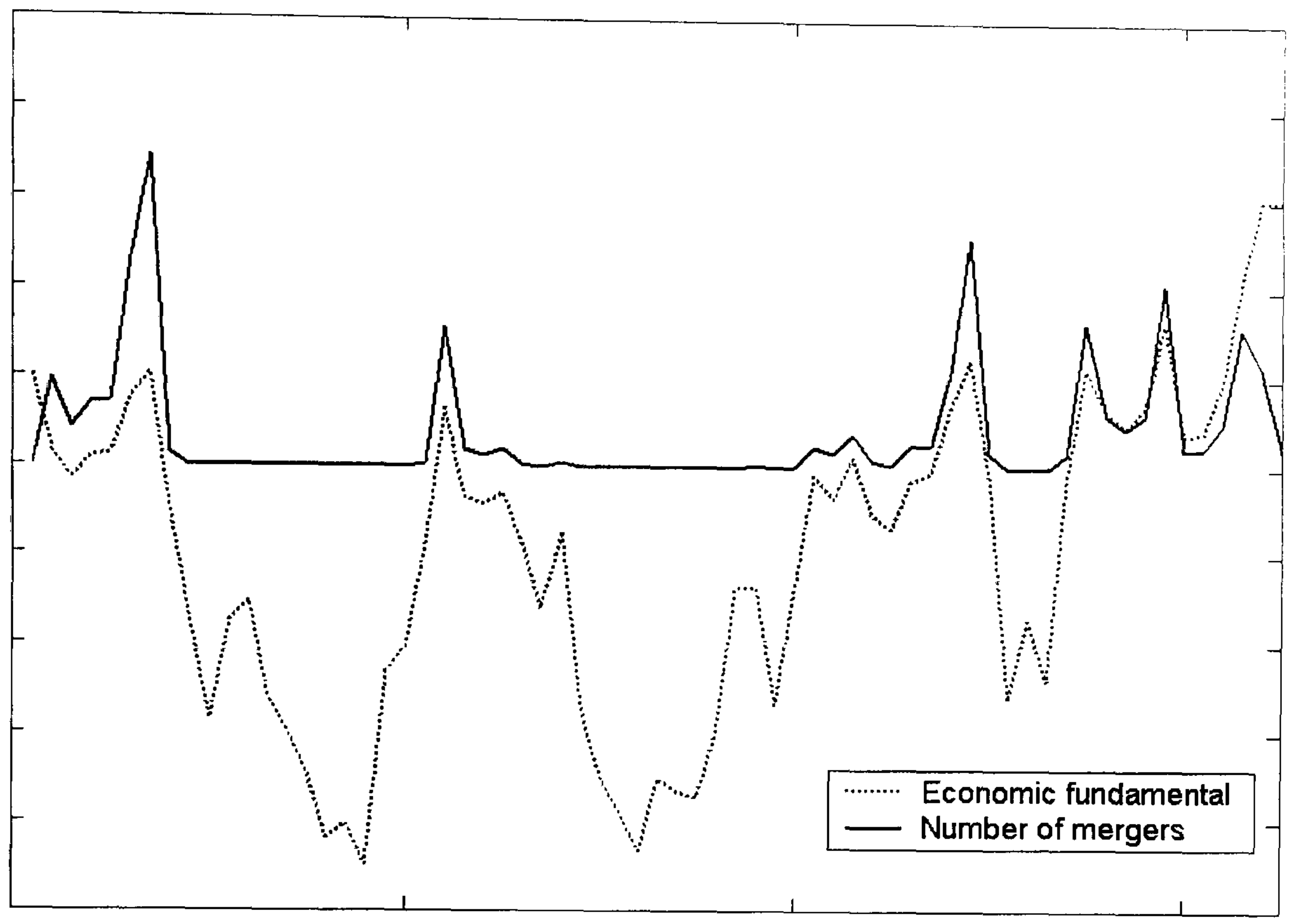

FIGURE 2. Simulated merger activity.

upwards potential, i.e. shifts probability mass towards higher realisations of θ_t while the distribution remains truncated below. This has the effect of increasing the options value of delay, or alternatively increasing the opportunity cost of an immediate raid. In turn, this has a direct effect on the other triggers of interest, the strategic trigger and the risk dominant trigger, which both tend to shift upwards. But there is also a second (strategic) effect of increased volatility. Ceteris paribus, increased volatility decreases the probability of an imminent merger wave by shifting the triggers upwards. However, increased volatility has the effect of making extreme realisations more likely. Thus for any given level of the trigger, there is higher probability that it will be hit. Which of these two effects is dominating is ambiguous and will depend on the specifics of the model.[19]

There is a last and somewhat subtler effect of increased volatility. Recall that knowledge of the process $\{\theta_t\}_{t=0}^{\infty}$ has two uses, namely forecasting the future evolution and generating a prior distribution over θ_t. The latter is influenced by the volatility of the process $\{\theta_t\}_{t=0}^{\infty}$ in that the volatility determines how informative the prior distribution is. Specifically, the more volatile the process $\{\theta_t\}_{t=0}^{\infty}$ is, the less precise is public information. This has the effect of weakening the requirement of signal precision needed for uniqueness. Morris and Shin (2002) show that with general Lipschitz continuous payoff functions, normal prior and normal noise, there exists a threshold of the relative informativeness of private and public information such that uniqueness obtains whenever the relative informativeness is lower than the threshold. Specifically, uniqueness obtains when new information is much more informative than history. Their results can be directly applied to the present model under the assumption that $\theta_t|\theta_{t-1} \sim N(\theta_{t-1} + \mu_\theta, \sigma_\theta^2)$ and $s_{it} = \theta_t + \varepsilon_{it}$ with $\varepsilon_{it} \sim N(\mu_\varepsilon, \sigma_\varepsilon^2)$. This result seems to be robust. Thus, in a very volatile environment, uniqueness of a perfect Bayesian equilibrium obtains even if there is significant noise in private information. In turn, increased noise in private information increases the probability of dispersed equilibria, in which not all acquirers raid simultaneously. If private information is very precise, a merger wave will involve all potential acquirers. Figure 2 illustrates simulated merger activity and a sample path of the economic

[19]See Grenadier (1996) for a related discussion.

fundamental with non-zero noise in private information. As is apparent from the graph, equilibria of this model generate distinct peaks in merger activity that resemble those observed in practice. For very low realisations of the fundamental, there is no merger activity. As the fundamental increases, merger activity picks up.

The predictions of the model presented here are broadly consistent with empirical studies. Gort (1969) and Mitchell and Mulherin (1996) find evidence of considerable cross industry variation in the rate of takeover activity as a response to economic shocks.[20] Mitchell and Mulherin (1996) cite their findings as support for the hypothesis that the shocks causing merger waves are industry-specific. This conclusion is consistent with my model. However, the previous discussion shows that industry-specific shocks need not be the sole cause of merger waves. That follows since economy-wide shocks could have different impact on different industries if one allows for differentiation of industries by the scarcity of target firms or in the way in which raiders compete for targets. Also, even if all industries are affected qualitatively in the same way by an economy-wide shock, industry-specific factors may play an important role in how much these shocks feed through to the future prospects of that particular industry. Blair and Schary (1993) discuss these issues at length, and conclude that

[20]Note however, that although Gort (1969) emphasises economic shocks, in his theory mergers are caused by systematic valuation differences. Thus his theory relies on shocks only to the extent that rapid change in the economic environment creates informational differences between market participants.

> ...[the evidence] suggests a formal model of restructuring activity as a function of a set of macroeconomic and industry-specific conditions [...]. Restructuring is triggered when those conditions reach some threshold point.

This observation closely resembles the nature of the equilibrium of the present model. Another implication of this model is that there can be merger waves that happen "for no apparent reason". That is, discontinuities in the rate of M&A activity can be triggered by very small increases in the economic fundamental. If this is indeed the case, empirical studies that ignore competition between acquirers and focus exclusively on the effects of changes in the economic fundamental would have great difficulty in explaining these waves.

It is a stylised fact that merger activity is highly procyclical. One explanation often proposed is that the cost of capital, proxied by some measure of real interest rates, decreases when the economy is expanding. Such an explanation is fully consistent with the present model. Empirical evidence also suggests that merger waves lead the business cycle, a finding which is also predicted by my model. Weston, Chung and Hoag (1990) note that

> The fact that mergers peak before overall economic activity may reflect that there is at any one time a pool of firms suitable for acquisitions and, as they are acquired in a period of high merger activity, the pool is diminished and merger activity returns to a low level.

In the current model, the merger wave is not triggered at the first-best level of the fundamental, i.e. when the fundamental is very high. Rather, mergers occur at the lower risk-dominant trigger, due to competitive pressure. The model assumes the existence of a scarce set of target firms, consistent with Weston et al.'s observation. Thus, although the economic fundamental may continue to rise, the merger wave peaks when all targets have been acquired.

6. An Example

In order to make the results less abstract, an explicit model is now presented that fits in the general framework presented so far. Consider a setting where two separate industries believe that at some uncertain point in the future, there will be demand of products whose production requires the participation of both industries. As an example, think of providers of media and content (e.g. AOL and Time-Warner). Let T denote the point in time where this new market opens, and let V denote the profits accruing to a supplier in this market. Suppose that upon merging, the parties incur a fixed implementation cost $c > 0$. Discounted profits from a merger are thus given by $\delta^T V - c$. For $T \leq 0$ an immediate merger is optimal, while for T sufficiently large, a merger yields negative profits. The surplus created through the merger is thus strictly increasing and bounded in $\theta \equiv -T$. Last, assume that there is uncertainty about the date at which the market will open.[21]

[21]In this particular model, the value of a takeover (absent competitive pressure) is expected to increase strictly over time, although the fundamental may be expected to be constant.

Turning to the allocation game, assume that a target facing a single bidder engages in some bargaining over the terms of the takeover. A target facing multiple bidders picks a single bidder with probability α and engages in bilateral bargaining. With probability $1-\alpha$, the target conducts an auction. This setup is consistent with the empirical observation that both bilateral negotiations and competitive bidding take place. In the auction, assume that target i receives N_i bids. The value for the target of the offer from bidder j is given by

$$U_{ij} = b_j + v_{ij}$$

where the idiosyncratic component is random and identically and independently distributed over i, j and satisfies standard assumptions of the probit model, and b_j is bidder j's bid.[22] Underlying these preferences lie non-modeled factors such as the tastes of target management, differences in corporate culture etc. Bidder j seeks to maximise

$$P_{ij}(b_j)\left[\pi(\theta) - b_j\right]$$

where

$$P_{ij}(b_j) = \frac{\exp[b_j/\mu]}{\Sigma_{k=1}^{N_i} \exp[b_k/\mu]}$$

[22] See e.g. Anderson, de Palma and Thisse (1992) for a thorough exposition.

is the probability that bidder j wins the target. In symmetric equilibrium all bids are equal, leading to

$$b^* = \pi(\theta) - \frac{\mu N_i}{1 - N_i}$$

This in turn yields equilibrium payoffs given by

$$\frac{\mu}{N_i - 1}$$

The payoff from the auction is independent of the exact value of the target, an instance of the Bertrand trap. Denote by $\pi(z_t, x_t, \theta_t)$ the expected share of the surplus obtained by a raider in the bargaining game. I explicitly let this share depend on z_t and x_t as these may influence the relative bargaining powers. The expected payoff to a raider is $\pi(z_t, x_t, \theta_t)$ for $z_t \leq x_t$. For $z_t > x_t$, expected payoffs are given by

$$\frac{x_t}{z_t}\left[\alpha\pi(z_t, x_t, \theta_t) + (1 - \alpha)\left(\frac{\mu}{N_i - 1}\right)\right]$$

Summing up, the raiding value is given by

$$R(z_t, x_t, \theta_t) = I_{[0,x_t]}(z_t)\pi(z_t, x_t, \theta_t) + I_{[x_t,y_t]}(z_t)\frac{x_t}{z_t}\left[\alpha\pi(z_t, x_t, \theta_t) + (1 - \alpha)\left(\frac{\mu}{N_i - 1}\right)\right]$$

where $I_{[a,b]}(z_t)$ is the indicator function. This game is simple yet realistic, and satisfies the conditions imposed on $R(z_t, x_t, \theta_t)$.

7. Conclusion

This chapter set forth a theory to explain the occurrence of merger waves. Merger waves were derived as an equilibrium phenomenon, in a simple timing game. It has been argued elsewhere that there are as many reasons for mergers as there are mergers. Indeed, the number of explanations set forth in the business literature are daunting, ranging from managerial empire building to such fuzzy notions as "consolidation" and globalisation. The model studied here has several advantages. First, it derives from simple and intuitive assumptions, namely scarcity of targets, imperfect competition and a value of delay. Second, the analysis builds on quite general assumptions on payoff functions and the stochastic process determining the evolution of the economic fundamental. This has the advantage of flexibility, and the promise of encompassing diverse explanations at once. Another advantage of the present model is that it reconciles empirical findings with casual observation. Specifically, it encompasses both dependence of the merger decision on (exogenous) macroeconomic variables and (endogenous) strategic considerations.

There are two assumptions of the current model that it would be interesting to relax. In this model, it is assumed that the identities and measures of targets and acquirers are determined exogenously. While these assumptions seem justified in some circumstances, they may not be so in general. Although some empirical work has been devoted to uncovering specific characteristics of targets and acquirers, the matter seems far from settled. Finally, although the present analysis has mainly focused on the interaction between acquirers, the model did not exclude the possibility that target firms play an active role. Modeling the takeover process more explicitly seems a worthwhile exercise. Intuition suggests that targets would have an interest in delaying the takeover, thereby increasing the created surplus. This might go some way in avoiding very inefficiently timed takeovers.

The analysis of this model has been carried through without any specific mention of target characteristics. Conceivably, one may extend the model by introducing target heterogeneity. In such a setting, acquiring firms would rank targets, and compete for the most desirable ones first. This would show that the merger wave phenomenon is not due to the fact that all targets are identical, but due to competition on the acquiring side of the market. In a setting with heterogeneous targets but without competition between the acquirers, a combination of the economic fundamental and a specific firm's characteristic will determine when this target is taken over. Thus the outcome would be "cherry-picking" rather than merger waves. That is, for a given state of the economic fundamental, only some firms would be desirable takeover targets and without any competitive pressure from rival acquirers, there would be no pressure to rush to prevent preemption.

Last, it should be pointed out that although the present analysis has focused on a preemptive motive for mergers (a result of the assumption of target scarcity), the adopted modeling approach is flexible enough to encompass other motives for mergers. That is, the analysis would not change substantially if one assumes that mergers confer negative externalities on non-merging firms.

8. Epilogue

In this epilogue, I will review some of the literature on merger waves, and relate it to the model presented above. I will first survey some of the existing industrial organisation and finance literature on merger waves in some more depth, and then draw out a bit further some of the observable implications and advantages of the presented model.

As already mentioned earlier, only few papers deal explicitly with the issue of merger waves. Perhaps, this should come as no surprise, as the standard oligopoly framework within which mergers are usually analysed seem somewhat ill-suited for explaining the observed wave pattern. Under fairly general conditions, a merger will not increase other firm's incentives for merger. In fact, the opposite is true. Namely, while a merger may be profitable for the merging parties, the resulting market concentration and higher equilibrium prices benefit non-merging firms as well. Merging is thus a public good, which can create an "after you" situation in which all firms want other firms to merge. In light of this problem, industrial economists have augmented the standard Cournot framework with explicit bidding games in order to show existence of bandwagon equilibria.

Fauli-Oller (2000) presents a simple perfect information model of four firms in which two efficient firms sequentially make bids for two inefficient firms. After the bidding stage (and after the inefficient firms have either rejected or accepted the bids), remaining firms compete in quantities under linear demand. If an inefficient firms is taken over, the merged unit becomes efficient too. In this model, the profitability of a merger is inversely related to the level of demand, and the equilibria are parametrised by the demand intercept. Fauli-Oller identifies four regions of interest. For very low demand, the first firm does not buy any firms while the second firm absorbs both inefficient firms. For low to intermediate demand, the efficient firms take over one inefficient firm each. For intermediate to high demand, the first firm takes over one target, and so does the second firm. Last, for high demand no mergers occur. Of all these equilibria, the one corresponding to low to intermediate demand is of special interest. In this equilibrium, firm two only goes through with a merger if the first firm has done so. Anticipating this, the first firm finds it optimal to do so, and thus triggers what Fauli-Oller terms a merger bandwagon. The basic insight is as follows. Whenever a target receives an offer, it will compare this with the alternative profits earned by declining a merger. Naturally, this reservation price is determined by subsequent play, i.e. on what the other target and the remaining bidder will do (or have done). Ultimately, these decisions rely on the profits earned at the market stage, given the market structure resulting from the two rounds of bidding. In the interesting range of the demand intercept, the first buyer pays a lower price than does the second buyer. This is because he faces two inefficient firms, and thus benefits from competition between them. The second buyer on the other hand, faces a single remaining target, which has a higher reservation price (because market concentration increases more than proportionately going from three to two firms than from four to three firms).

While the results are suggestive, some reservations are in order. First, it is not clear whether the results generalise beyond the four firm case. Second, and more importantly, is the prediction that merger waves should occur in periods of low demand. While consolidating mergers in declining markets are indeed observed, merger waves happen in expanding markets, and are positively correlated with increases in industrial output. Hence, while the model may explain some mergers, it does not seem to fit observed wave patterns. A last issue is that the model assumes an exogenous order of moves. It is not clear if the model's predictions are robust to an extension in which both efficient firms bid simultaneously for the targets.

Rodrigues (2002) builds on the analysis of Fauli-Oller but assumes that all firms are ex ante identical, and furthermore allows any firm to bid for other firms. He largely confirms Fauli-Ollers' findings, but finds that the equilibria are parametrised by the ratio of fixed costs to market size. When the ratio is very high, no mergers occur. When the ratio is very low, the only equilibrium market structure is monopoly. For intermediate to high ratio, the first firm remains independent while the second firm absorbs both remaining firms. For low to intermediate ratio, the first firm takes over one other firm while the second firm takes over the remaining one. While Rodrigues relaxes some of the assumptions of Fauli-Oller, some restrictive features remain. Firms are only allowed to merge sequentially, and the analysis is conducted in a four firm setting. Last, the model confirms Fauli-Oller's finding that merger waves should happen in conditions of low demand.

One of the more interesting contributions is that of Fridolffson and Stennek (2000). They analyse a three firm model in which firms submit bid and ask prices for each other, and subsequently compete in quantities given the resulting market structure. The crux of their analysis is that for particular parameter values, being an outsider (i.e. staying independent while the other two firms merge) may be worse than participating in an unprofitable merger. This result hinges on there being very strong negative externalities through the product market. Specifically, it depends on the assumption that synergies arising from the merger are substantial. When synergies are large and firms compete in quantities, the non-merging firm may lose on two counts. First, the merging parties may expand their joint output, forcing the remaining independent firm to contract his output (since quantities are strategic substitutes). Furthermore, if the resulting total industry output is increased compared to the pre-merger equilibrium output, equilibrium prices will fall. Thus there is a preemptive motive for mergers. If a firm expects its rivals to merge, it can do no better than to participate in the merger itself.

While the analysis is quite elegant, the synergy assumption seems strong, particularly considering empirical evidence pointing to the contrary. Synergies are not generally found to be very substantial. In any case, even under the assumed parameter constellation, the equilibrium where no firms merge remains. It is thus hard to explain what triggers the wave. Last, it should be noted that merger waves happen both within industries and across industries. The

Fridolffson and Stennek model is by construction only suited for mergers within a specific industry.

Boeckem (2001) considers an industry of n heterogeneous firms competing in quantities. She assumes that merging firms retain their respective pre-merger technologies (i.e. there are no synergies), but that the merging firms may benefit from reallocating production between them. The setup is similar to that of Fridolffson and Stennek (2000) but importantly, it is assumed that firms have increasing marginal costs, which has important implications for the equilibrium merger activity. Specifically, she shows the following interesting result: A merger between the two lowest cost firms increases the profitability of subsequent mergers. In other words, there exist distributions of firm efficiencies such that if a first merger becomes profitable for exogenous reasons (e.g. due to shifting demand or decrease in merger costs), the next merger becomes profitable, in turn triggering subsequent mergers. The basic mechanism is the following. Consider a setting in which the firms are heterogeneous with respect to their marginal costs. In general, firms would want to delay a merger in order to free-ride on other firm's mergers (with resulting increase in prices). Also note that the gain from merger is decreasing in the marginal costs of the merging firms. In other words, low cost firms gain more from merging than do high cost firms, ceteris paribus. Now assume that only a merger between the two lowest cost firms is profitable (or has just become so for exogenous reasons). These firms have no incentive to delay the merger, since higher cost firms would find a merger under current market conditions unprofitable. Thus the two lowest cost firms merge immediately. By merging, the firms contract output, prompting non-merging firms to expand theirs. Now consider the next couple of firms. These are now the lowest cost independent firms remaining in the industry, and a merger between them has similar effects to that of the first merger. But there are important differences. First, two aggressive low cost competitors have been replaced by one larger firm. Second, and more importantly, all competing firms are now higher on their marginal cost curves (recall that marginal costs are increasing). Therefore the competing firms will increase outputs in response to the next merger, but at a decreasing rate. In other words, prices become less sensitive to increased merger activity, thus raising the profitability of subsequent mergers. For a carefully chosen distribution of marginal costs across the industry, one can thus construct equilibria in which initially, only one merger is profitable, but where this merger in turn makes the next merger profitable etc., and a bandwagon follows.

Molnar (2000) identifies a preemptive motive for mergers, similar in spirit to that of Fridolffson and Stennek (2000). He considers an industry of three firms, facing a market with linear demand. Prior to the market stage, two of the firms participate in a sealed bid second price common values auction for the third firm. The equilibria of the bidding game are parametrised by the degree of synergies resulting from a merger. He shows that for low synergies, no merger occurs. For intermediate synergies, only one bidder participates and bids the reservation price. For high synergies, both bidders participate, and bid exactly the increased profits due to the ensuing synergies. Interestingly, Molnar shows that for very high synergies, equilibrium bidding involves the winner actually paying so much for the target that the post-merger profits minus the price paid is below the pre-merger profits of the firm. In other words, the negative externality on the losing bidder when synergies are substantial are so severe that a bidder rationally overpays and is worse off than if no-one had merged.

Cabral (2000) considers a standard herding model in which firms sequentially decide wether or not to merge, after observing past firm's decisions and a private signal. She assumed that there are two possible states of nature, one in which mergers are profitable and one in which they are not. As usual in herding type models, this is an environment with common values, and thus the firms learn about which regime they are in from observing other firms' merger decisions. Under certain circumstances, the standard herding result obtains. That is, there may be informational cascades in which public information becomes so strong that firms ignore their private information. This can lead firms to merge, even after observing private information that merging is unprofitable. Unfortunately, the restriction to sequential moves seems hard to justify, as does an exogenous pre-specified order of moves. The herding result seems to some extent to hinge on this structure. Last, the merger decision seems wholly disconnected from either strategic interaction between firms contemplating to merge, or exogenous factors influencing the profitability of mergers. The waves are wholly driven by expectations which may be triggered by "mistakes" in early stages of the game. This may be an overly pessimistic view of why mergers happen in waves.

Nitsche (2001) considers a dynamic model of merger races in multi-market industries in which two large firms compete in taking over control of a large pool of smaller firms each serving separate geographical markets. In his model, the value of being large stems from the possibility of credibly predating smaller rivals. The takeover phase is modeled as a three stage game. In the first stage, the two large firms simultaneously submit sealed bids for targets. At the second stage, a target who has received an offer either accepts or rejects. In case of rejection, a second price auction is conducted. If the target rejects the outcome of the auction, the firm that submitted the highest bid decides wether to enter the market or not. The three stage process is designed to capture the following tradeoff: The target, by rejecting all offers, risks entry by a large rival, which is able to drive the target out of his local market. A bidder on the other hand, does not know if there are rival bids for the specific target. By offering the target a poor deal, the target may refuse and sell to another rival bidder. Nitsche assumes that credible predation is only possible when a firm has reached a critical size, and thus two symmetric large firms will rationally race in order to get first past the post. If one firm takes over a target, the rival firm will have more incentives to succeed in subsequent takeover attempts.

One advantage of Nitsche's model is that the incentive to grow is disconnected from product market competition. Given the empirical evidence, this is a desirable property. On the other hand, the model predicts significant market concentration, and may even lead to duopoly or near complete monopolisation.

This prediction seems at odds with what is observed in practice. Next, symmetry seems to play a very crucial role in the race result. Specifically, if any large firm gains an early advantage (i.e. controls two or more subsidiaries than the rival), the race is essentially lost for the laggard. Perfect symmetry is perhaps a strong assumption. Last, the incentive to merge it entirely independent on any macroeconomic variables. Thus the model suffers from some of the same deficiencies as other reviewed contributions.

Perhaps recognising the difficulties in explaining merger waves within an oligopoly framework, financial economists have built several different types of models that leave out product market interaction altogether.

An interesting approach to merger waves is that of Rhodes-Kropf and Viswanathan (2002). They consider a static setup in which n risk-neutral bidders compete over a single target through a second price sealed bid private values auction. Their model ties merger waves to different sorts of misvaluation, firm-specific as well as industry-wide. The basic setup is as follows. Each bidder perfectly knows its own value, which is private information. The less informed market (and the other bidders and the target) only knows the market value, which may differ from the true value due both to firm specific and industry wide misvaluation. Similarly, only the target perfectly knows his own true value, while the market has knowledge only of the market value (which may differ from the true value due to misvaluations similar to those of the bidders). Importantly, the model assumes that bidders pay with fractions of their stock. In turn, this implies that the target, in deciding which bid (if any) to accept, must estimate the true value of the bidders' stocks. In doing so, the target faces an inference problem. A high bid may signal either that merging with the specific bidder will create large synergies, that the market in general (and thus the bidder's stock) is overvalued, or that the target's stock is undervalued. The target cannot infer from the bids if the market's misvaluation of the target is specific to the target, or a result of market-wide misvaluation. The authors show that the more the market is overvalued, the larger is the target's expectation of his own misvaluation. In turn, this means that the target's assessment of the combined misvaluation puts too little weight on market-wide misvaluation. Thus the bids tend to look more generous during periods of high market-wide misvaluations, and the target accordingly more likely to accept an offer.

While the model is elegantly constructed and has interesting implications, some reservations are in order. First, the model is static and it is unclear how the analysis could be extended to a fully dynamic setting. In particular, a bidder's decision of when to participate in a takeover auction seems far from trivial. Second, a merger wave is identified with an increased probability of accepting a takeover bid. The bidders participate in the auction by assumption (i.e. there is an exogenous number of bidders), and the takeover is triggered solely due to the target's acceptance. The opposite seems to hold in practice. While targets surely have some say in the merger process, takeovers are usually triggered by the acquirer actively deciding to acquire the target. In other words, the model seems to leave the bidder with too little to decide (only the price, not wether or not to participate in the auction). An interesting feature of the model is that it does not need to assume that synergies are actually increased during boom times. The mergers are driven purely by increased misvaluation. As the authors point out, this does not preclude any synergy gains resulting from mergers. Rather, they show that even without increased synergies, mergers may be the rational outcome of misvaluation and private information. A last issue is that the auction is one of independent private values. Fishman (1988) assumes private values while Cramton (1998), Cramton and Schwartz (1991) and Bulow, Huang and Klemperer (1999) assume common values settings. Ultimately, it is a empirical question whether common values or private values are most important in a merger context. The empirical literature has shown that M&A activity is highly correlated with such notions as industry-wide technological progress, pointing to common values as the adequate model.

Gorton et al. (1999) consider a setting in which managers derive private benefits from managing firms, but also have stakes in the firm's profits. The model assumes that three firms, ranked by size, are randomly chosen one at the time to make a bid for a smaller firm. A manager, if given the opportunity, may decide to take over a smaller firm and thereby keep his job with probability one. That is, it is assumed that if a firm has taken over another firm, the merged entity is effectively shielded from takeover. The model assumes that at a future date, mergers are possibly profitable, with mergers between larger firms being more profitable than merger between smaller firms. The largest firm can by assumption not be taken over, and will thus only engage in profitable mergers. The manager of the middle firm on the other hand, may have a defensive motive for mergers. Too see this, consider a state of nature in which mergers are profitable. In this setting, if given the opportunity, the largest firm will take over the middle firm (because profitability is increasing in the size of the merging firms). If such a merger goes through, the manager of the middle firm loses his job certainly, but benefits from his share of the now more profitable merged firm. It is thus easy to see that if the private benefit for staying in charge is large enough compared to the stake in the new firm, the manager of the middle firm will try to defend himself by acquiring a smaller firm. The authors extend this analysis to a six firm setting where, in each of two periods, two firms are randomly given the opportunity to take over smaller rivals. They show that the defensive motive identifies in the three firm setting can lead to more than one firm at the time trying to take over smaller rivals in order not to be taken over themselves. They term this a merger wave. While the model is nice, the authors make assumption that practically yield the desired results almost by themselves. For example, it would be interesting to let the probability of a manager keeping his job be a nondegenerate function of the relative sizes of the firms, and extend the analysis to multiple periods. By also doing away with the assumption that large mergers are more profitable than small ones, and allowing merged firms to be taken over subsequently, the model would generate truly interesting dynamics. For example, it would no longer necessarily be true that a firm would seek a merger with the largest firm smaller than itself, since this would involve a high probability of the manager losing his job. Conceivably, it could be optimal to do a series of small acquisitions where survival probability is large, and build up comfortable size this way.

Shleifer and Vishny (2001) consider a behavioural model of merger activity in which acquisitions are driven purely by stock market valuations. They consider two firms with different stocks of capital. For unspecified reasons, the market has valuations of these capital stocks which may or may not reflect fundamentals. By merging, the two firms pool their capital stocks. Three crucial assumptions are made. First, the market values the merged entity as some linear (not necessarily convex) combination of the two firms' pre-merger valuations. These "perceived" synergies are not based on fundamentals, but rather on "beliefs" derived from "stories". Next, it is assumed that the perceived gain in stock market value is divided between target and acquirer through some bilateral bargaining process. Last, it is assumed that managers are myopic, and only seek to maximise short run stock market value. Within this setup, Shleifer and Vishny assert that during periods when the stock market perceives mergers to create synergies, managers will oblige and increase merger activity. Unfortunately, it is not explained why the market should perceive synergies when these are in fact absent, or why these deficiencies in managerial incentives are not corrected for through manager's compensation packages.

Building on a model of Dixit and Pindyck (1994), Lambrecht (2001) considers a continuous time model of two firms in a competitive industry contemplating when to merge. The firms have Cobb-Douglas production functions with decreasing returns to scale, and their output price follows a geometric Brownian motion. By merging, the firms may reduce marginal costs or increase the returns to scale (or a combination of both), but incur a fixed cost of merging. The setup is that of a standard real options model and the results accordingly standard. Specifically, he shows the existence of a trigger level of the output price such that the merger is implemented for prices above the trigger. He asserts that within this setup, mergers should happen during periods of expanding demand, in accordance with observed M&A patterns. Two main objections to the model can be raised. First, the model has no strategic element whatsoever. In that sense, it only tells part of the story. Second, merger waves in this setting must be interpreted as situations in which the economic fundamental experiences radical shifts. To see this, consider an extension of the model into one with a number of firms deciding when to merge, and assume (realistically) that firms are heterogeneous in some respect. With this assumption, it is straightforward to find a one-to-one mapping between each pair's idiosyncratic characteristic and a critical level of the economic fundamental such that this specific merger is implemented whenever the economic fundamental reaches a critical level. Thus, a smooth evolution of the economic fundamental will not create waves, since only the marginal pair of firms will find it optimal to merge. In other words, only sizeable shifts in the fundamental would trigger increased merger activity. In contrast, the model presented here is immune to this criticism. Due to competitive pressure, the level of the fundamental at which a given acquirer is indifferent between raiding and waiting jumps downwards when other acquirers raid, and this is true also with heterogenous acquirers. One advantage of the

Lambrecht model is however, that it is possible to get closed form solutions for the triggers. This facilitates comparative analysis considerably.

The model of merger waves presented here has several satisfactory features. First, as already mentioned, it predicts patterns of M&A activity broadly consistent with what is in fact observed. Furthermore, it contains both strategic and non-strategic elements, also consistent with experience. Third, the model is fully dynamic. While interesting comparative statics results may be obtained from static models, a fuller understanding of merger dynamics is obtained only through the study of full-blown dynamic models. The adopted framework of analysis is very general in nature, allowing the model to be expanded in several respects while maintaining the basic structure. For example, a second stage (after the completion of the allocation game) may be introduced in which product market interaction takes place, and one could introduce an element of private values (to be discovered by the raiders at a cost, as in Hausch and Li, 1993). Since the model does not assume a specific type of product market interaction, it is not necessarily restricted to horizontal mergers, but may be used in the analysis of vertical, conglomerate or cross-border mergers (although, of course, special attentions should be given to other effects resulting from the interaction in question).

Arguably, the strongest assumption is that of relative target scarcity. This assumption drives the model, and it would be interesting to see if the predictions of the model change if this assumption is relaxed. One possible avenue for such an extension is to assume that acquirers only have unit demand during a given period, and thus do no have to make a once-and-for-all decision.

While the real options framework is sensible from a methodological perspective, it is notoriously hard to evaluate empirically (but see Moel and Tufano, 2002, the discussion in Dixit and Pindyck, 1994 and references therein). The merger wave model developed in Part I is no exception. The first hurdle is determining the correct interpretation of the economic fundamental. Leaving aside this issue and assuming that the appropriate interpretation is found, some general things can be said. First, as the volatility of the economic fundamental increases, one should expect a decrease in the intensity of M&A activity. To my knowledge, no empirical study has analysed the relationship between the level of economic volatility and the rate of M&A activity. Next, because the measure of raiders in any given period is endogenous, the model predicts a clear pattern of single-bidder versus multiple-bidder mergers over the merger wave. Specifically, single-bidder mergers should be prevalent during periods of low merger activity, when competition for targets is relatively low. During the peak of the merger wave, competition for targets is significant, and thus multiple-bidder contests should be the norm. In existing data on mergers and acquisitions, activity during a given period of time (quarter, year etc.) is given as stocks where single-bidder and multiple-bidder takeovers are lumped together. Conceivably, it should be possible to create two separate series to study which type of contest is prevalent at a given stage of the wave.

One possible drawback of the model's generality is that the comparative statics tend to be somewhat vague. In comparison, papers such as the bank-runs models of Goldstein and Pauzner (2000) and Rochet and Vives (2001) are relatively rich, having more parameters with respect to which one may do comparative analysis (in particular, the specifics of the debt contract, the constraints and policies of the central bank etc.). Sharper predictions may be obtained from the presented model by sacrificing generality and assuming specific forms for the evolution of the economic fundamental and of the allocation game.

Another issue is the robustness of the results to model specifics. In particular, a possible concern is that the merger wave result of the presented analysis depends crucially on the symmetry between acquiring firms. While symmetry certainly streamlines the analysis, it seems not to be of great importance. In fact, the analysis of the complete information game carries over in a straightforward way to asymmetric acquirers (and for that matter, to a setting with a finite number of acquirers). For the incomplete information game, nothing definite can be said. Frankel, Morris and Pauzner (2000) analyse equilibrium selection in very general global game settings, allowing the players to belong to distinct types with different payoff functions and action spaces. In such a setting, they show a limit uniqueness result (i.e. when signals become very precise, a unique equilibrium survives). But the game considered here does not fit into their framework (in technical terms, they study supermodular games while the present model belongs to the more general class of quasisupermodular games).

Advances in the global games literature will show whether or not the present model could be extended along the lines of their paper.

From a methodological perspective, the presented model is interesting in its own right. It is a sequential bidding game in which the measure of participants in each period is fully endogenous, taking place in a fully decentralised setting. To my knowledge, only one other paper has dealt with sequential bidding games of this nature, namely Bulow and Klemperer (1994)'s paper on rational frenzies and crashes. Importantly though, and in contrast to my model, they study an independent private values setting in which a monopolist sells to a number of bidders. In their setup, the monopolist finds it optimal to set a decreasing sequence of prices over time, thus giving the buyers an incentive to delay their purchase. In a setting with more than one seller, their results would break down, and only the equilibrium with immediate frenzy would remain. In my model, the options value of delay is somewhat disconnected to the seller's prices.

Next, my model shows that the global games analysis carries over naturally to dynamic settings which has a simple recursive structure. In the present work, the intertemporal link is the stock of remaining targets, which in effect endows the payoff function with a single crossing property in each period. Possibly, a similar analysis would carry over to settings where other intertemporal constraints are present.

Last, it is an interesting contribution to the literature on games of timing. In such games, it is usually hard to pin down equilibrium behaviour. In the present framework, a unique equilibrium in pure strategies exists. This may be useful for the study of other timing games, e.g. dynamic games of speculative attacks where the central bank's reserved are eroded over time.

Part 2

Contracting and Deadlines

CHAPTER 5

Background and Review of the Literature

"Most attempts to explain why deadlines are missed and budgets overrun go no farther than Murphy's often-quoted aphorism" - Musgrove (1985)

This part of the thesis studies contractual relationships between economic agents in situations where there is time to build. In particular, I study the structure of optimal contracts which are made contingent on delivery time under two different economic scenarios, namely adverse selection and moral hazard.

Explicit stipulation of deadlines are often observed in practice. Examples include labour contracts within firms, sub-contracting of parts of a larger project to other firms, and procurement contracts for large-scale projects such as weapons systems and infrastructure. In some of these cases, a deadline is determined exogenously. For example, if a production process involves the use of a perishable input, failing to meet a production deadline might mean the loss of the input in question. Another example is a situation in which one of the parties is bound by a contract with a third party. But in many situations, the deadline is imposed endogenously by one of the parties. It seems natural that a principal might be impatient and discounts the value of a project from the date of its completion. As an example, one can consider a firm that out-sources the construction of a plant. The longer the time until completion of the plant by the sub-contractor, the larger are the foregone profits of production from the plant.

In the empirical literature, as well as in the popular press, the notion of delays and time overruns has received considerable attention. In a study of information systems development, Jenkins, Naumann and Wetherbe (1984) found that 90% of projects were delayed. Research at the Rand corporation by Marshall and Meckling (1962) showed that the ratio of actual completion time to estimated completion time for a sample of large scale weapons systems was 1.5, while Peck and Scherer (1962) found a ratio of 1.36 for another sample of weapons development programmes. Mansfield, Schnee and Wagner (1971) report similar ratios for drugs development, ranging from 1.61 to 2.95. Casual observation suggests that the occurrence of project delays are pervasive in all areas of construction and development where time plays an important role. Despite these observations, economic theory has been curiously silent about possible explanations. At first glance, one may be tempted to reduce the problem of project delays to irrational decision making or overly optimistic forecasts. But once it is recognised that delays occur consistently, one is faced with the question of why this information is not properly taken into account when performing estimates of project duration. The models presented here will not seek to settle the matter definitely. Rather, they illustrate the determination of development time in the context of incentives under asymmetric information. To the extend that the present work sheds new light on the subject, it is in recognising that incentives are crucial in understanding how fast work is performed, and that delays may be the inevitable consequence of not adequately incorporating incentive considerations in contract design.

The next two chapters of this part propose two separate models to study the determination of delivery time in settings of asymmetric information. The first is one of adverse selection and the second one of moral hazard.

In the adverse selection model of chapter 6, the agent can perfectly control when to complete the project, but has private information about his own ability (efficiency). Not being able to monitor progress, the principal writes contracts in which wage payments depend on delivery time. In this setting, it is shown that inefficient delay is willfully induced by the principal in order to separate agents of different efficiency types.

In the moral hazard model of chapter 7, the principal cannot observe how much effort the agent exerts. Progress in the work on the project depends stochastically on exerted effort, and thus the agent can only control delivery time probabilistically. The main finding of this analysis is that delays are closely connected to the degree of commitment on the part of the principal.

1. The Literature on Optimal Deadlines

Although the design of contracts that involve deadlines for completion of projects is of great practical interest, the theoretical literature on the subject is surprisingly small. Only one paper has dealt explicitly with contracting for delivery in long-term projects, namely Cukierman and Shiffer (1976). They consider the effects of commonly used payment schedules on an agent's incentive to deliver at the efficient point in time. In particular, they show that if the payment is exogenously determined, and thus not contingent on the date of delivery, the agent will benefit from delaying. However, their model has no uncertainty or asymmetry of information.

The problem of delays and its connection to asymmetric information has been treated in another context by O'Donoghue and Rabin (1999). They explain delays as the effect of the agent's propensity to procrastinate. In contrast, the models developed here assume that the agent's preferences are fully time consistent. Finally, another related contribution is that of Musgrove (1985). Based on empirical evidence, he presents a model to explain why project delays are so common.

2. Other Related Literature

Methodologically, the adverse selection model bears resemblance to three separate strands of literature, namely that on optimal regulation of monopolies, the literature on contracting in long-term projects, and the literature on optimal R&D programmes. The literature on optimal R&D programmes, exemplified by Kamien and Schwartz (1971), Lucas (1971) and Grossman and Shapiro (1986) can be viewed as a special case of the model presented here, as it features no asymmetric information or incentive considerations. I can thus use the results obtained by them as a benchmark for the adverse selection model.

When finding the optimal contract in the present setup, I draw heavily on techniques developed in the optimal regulation literature. Specifically, I use a mechanism design approach by letting the principal offer a menu of contracts, and let the agent self-select. Formally, this model belongs to the class studied by Guesnerie and Laffont (1984).

Last, there is a small literature on contracting for completion of long-term projects. This literature focuses on explanations of cost over-runs. As it turns out, the difficulties in explaining cost over-runs are similar to those encountered when explaining delays. Lewis (1986) presents a model of cost over-runs in which a principal hires an agent to complete a number of tasks sequentially. The value of the project is realised upon completion of the last task, and neither the principal nor the agent can commit to contracts specifying more than the completion of a single task. He assumes that the principal can abandon the project altogether if cost realisations on initial tasks prove too high. In this setting, he shows that the cost distribution at later stages of the project stochastically dominate that of earlier tasks. Unfortunately, he works with an exogenously determined incentive schedule, and is thereby unable to characterise the optimal contract. Picking up this point, Arvan and Leite (1990) build a similar model in which they fully endogenise the contract. They confirm Lewis' stochastic dominance result, and find that the optimal incentive scheme becomes steeper as progress on the project is realised.

The moral hazard setup is a type of sequential development model in the spirit of Grossman and Shapiro's (1986) work on R&D programmes. That is, progress is measured through success on a number of distinct tasks performed in sequence. In contrast to existing work, progress depends on effort exerted by the agent, which is subject to moral hazard. The model has a real options interpretation, in the sense that successfully completing a task allows the agent to commence work on subsequent tasks (and receive associated wages).

CHAPTER 6

A Theory of Optimal Deadlines I: Adverse Selection

In this chapter, I propose a principal-agent model of adverse selection where it is assumed that the agent's efficiency is unknown to the principal, and that the agent's effort over time is unobservable (and thus non-contractible). I also assume that the principal can perfectly commit to any contract signed at the beginning of the game. In this setting, the optimal contract will be derived and characterised.

1. The Model

A principal hires an agent to complete some project. The principal enjoys a known constant utility $S > 0$ on completion, and discounts it at rate r. I assume a perfect capital market in which the riskless return on assets is also r.[1] Upon completion, the agent receives a wage payment $w \geq 0$. Let $V(t) = e^{-rt}S$ denote the principal's discounted present value of the completed project, and $w(t) = e^{-rt}w$ the discounted wage payment to the agent. The agent exerts effort over time at rate $x(t) \geq 0$, which is transformed to cumulated effective units of effort (or progress) $z(t)$ according to the production function

$$z'(t) = \theta x(t)$$

[1]Thus the timing of transfers is immaterial. See appendix 5.

where $\theta \in [\underline{\theta}, \overline{\theta}]$ is the agent's efficiency parameter, and $\overline{\theta} > \underline{\theta} > 0$. To complete the project, a cumulated amount of effective effort L must be exerted by the agent. That is, termination time T is implicitly defined by the boundary condition $z(T) = L$, or equivalently by

$$\int_0^T \theta x(t)dt = L$$

The agent has convex disutility of the rate of effort $\psi(x(t))$ with $\psi' > 0$, $\psi'' > 0$ and $\psi''' \geq 0$. Thus total discounted disutility for termination time T and efficiency θ is given by

$$c(\theta, T) = \int_0^T e^{-rt}\psi(x(t))dt$$

The principal wishes to maximise the following separable utility function:

$$\max_{w(T),T} \int_0^T [V(T) - w(T)]dt$$

subject to $w(T) \geq c(\theta, T)$. Letting this participation constraint be binding, the principal's problem is thus

$$\max_{x(x),T} \left\{ e^{-rT}S - \int_0^T e^{-rt}\psi(x(t))dt \right\}$$

with $z'(t) = \theta x(t)$ and $z(0) = 0$ and $z(T) = L$. Letting $\lambda(t)$ be the Pontryagin multiplier of the constraint, the Hamiltonian is given by

$$H = -e^{-rt}\psi(x(t)) + \lambda(t)\theta x(t)$$

It follows that the co-state variable obeys the differential equation $\lambda_t(t) = 0$ and thus $\lambda(t)$ is constant over time. Maximising the Hamiltonian with respect to the control variable $x(t)$ yields the optimality condition characterising the optimal path implicitly. This is given by

$$\psi'(x(t)) = \lambda(t)\theta e^{rt}$$

Since ψ is strictly increasing, effort is strictly increasing over time along the optimal path.[2] Last, the transversality condition characterising the optimal termination time T is given by

$$rS + \psi(x(t)) = e^{rT}\lambda(t)\theta x(T)$$

Combining the transversality condition and the optimality condition yields

$$rS + \psi(x(t)) = x(T)\psi'(x(t)) \tag{1.1}$$

In other words, the principal chooses T such that marginal benefit from completing the project an instant sooner is equal to the marginal cost of doing so.

[2]The second order condition with respect to the effort rate is $-e^{-rt}\psi''(x(t)) < 0$ which holds given the assumption $\psi'' > 0$.

Next, substituting the optimal effort path in the isoperimetric constraint defines the multiplier $\lambda(T, S, L, \theta)$ implicitly as a function of the parameters of the problem:

$$\int_0^T \theta(\psi')^{-1}\left[e^{rt}\lambda(T, S, L, \theta)\theta\right] dt - L = 0$$

By the implicit function theorem, $\lambda(T, S, L, \theta)$ is unique and has continuous derivatives in all arguments (when all other arguments are held constant). Thus the following derivatives:

$$\lambda_T < 0, \quad \lambda_L > 0, \quad \lambda_\theta < 0$$

It will now be determined how the optimal termination time and effort path vary with the parameters of the model. First note that for given T, the optimal path of effective effort has to integrate to L. Thus, an increase in L will raise the rate of effort $x(t)$ for all t. Equivalently, for given L, an increase in T will decrease $x(t)$ for all t. Totally differentiating the optimality condition (1.1) with respect to S and rearranging yields

$$\frac{dT}{dS} = \frac{r}{x(T)x'(T)\psi''(x(T))} < 0$$

Next, totally differentiating the transversality condition with respect to L and rearranging yields

$$\frac{dT}{sL} = \frac{-e^{rT}\theta x(T)\lambda_T}{r^2 S + r\psi(x(T))} > 0$$

Last, note that the isoperimetric constraint can be rewritten as

$$\int_0^T x(t)dt \geq \theta^{-1}L$$

Defining $\widetilde{L} \equiv \theta^{-1}L$, clearly $dT/d\widetilde{L} > 0$ follows from the transversality condition. It follows immediately that $dT/d\theta < 0$. Summing up, the efficient completion time is a decreasing function of the value of the project S, an increasing function of the scale of the project L and a decreasing function of the agent´s efficiency θ.

Figure 3 illustrates the (gross) value to the principal, costs, (gross) marginal value and marginal costs as functions of completion time.

2. Perfect Information Case

In the perfect information case, θ is a known technology parameter. As shown in the previous section, the profit maximising termination date schedule is given implicitly by the transversality condition

(2.1) $$V'(T^*(\theta)) = c'(\theta, T^*(\theta))$$

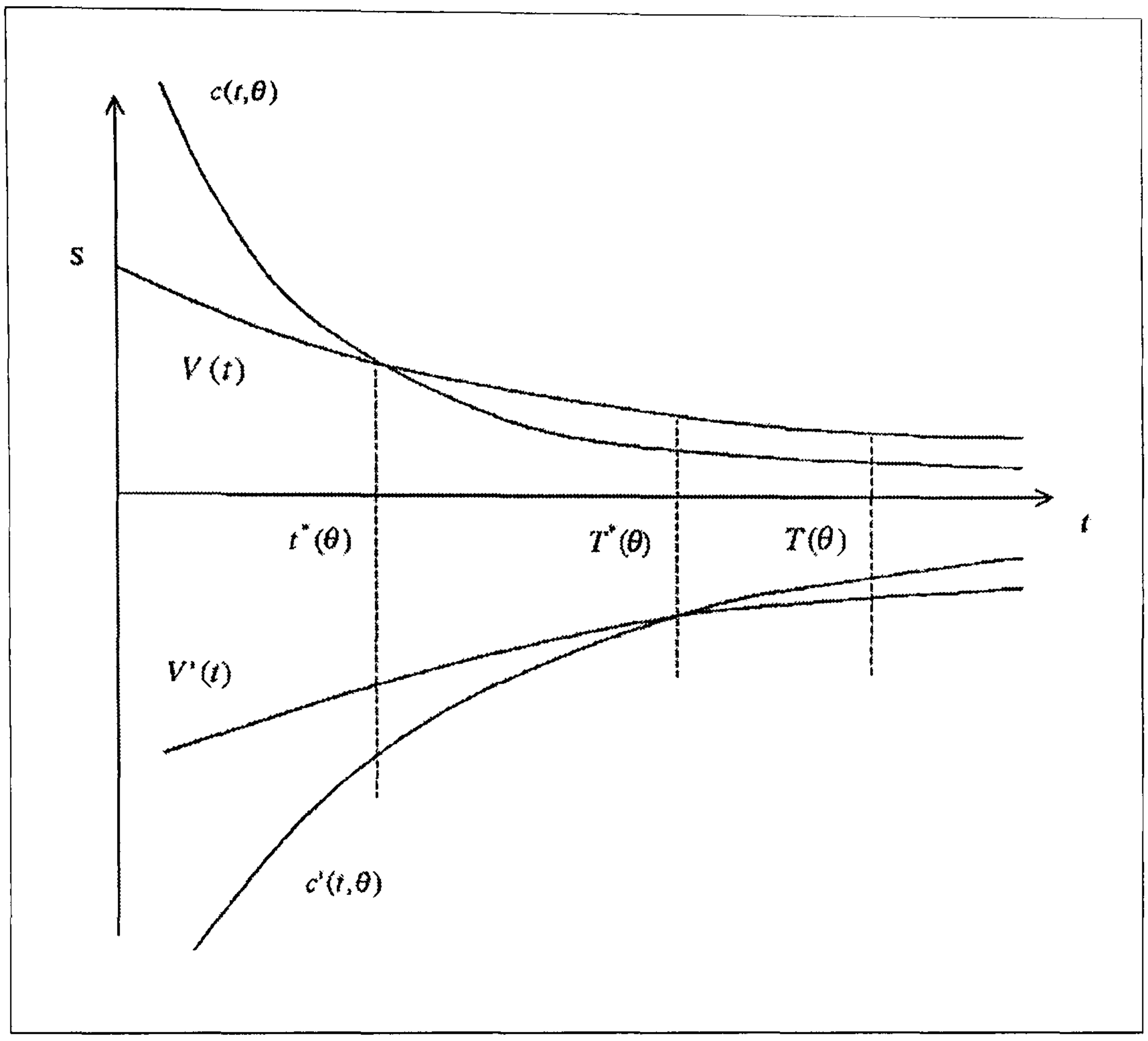

FIGURE 3. Value, costs, marginal value and marginal costs.

where derivatives are taken with respect to termination time T. To ensure that this condition is both necessary and sufficient for optimality, the following second order condition, which is satisfied given the maintained assumptions, needs to hold:

$$V''(T(\theta)) - c''(\theta, T(\theta)) < 0 \tag{2.2}$$

The optimal contract should thus seek to implement a termination schedule that for each type of efficiency θ, equalises discounted marginal benefit and discounted marginal cost.

A host of different contracts yield the desired outcome. A natural way of implementing such a termination schedule is to announce for each type θ a deadline $T(\theta)$. If the agent delivers at or before the deadline, $t \leq T^*(\theta)$, he is paid $w^*(\theta)$. For termination at any $t > T(\theta)$, he is paid nothing. In case the agent meets the deadline, the wage is determined such that his individual rationality constraint is binding. Alternatively, the contract may include a specification of liquidated damages such that the agent fully internalises delay costs for delivery past the deadline. I.e. the principal can set the wage $w(\theta^*)$ as above for all t but require that the agent pay damages at rate $d(t) = V'(t)$ for $t > T^*$. Last, the principal can simply sell the entire project to the agent for a fixed fee $P(\theta) = V(T^*(\theta)) = c(\theta, T^*(\theta))$.

In the special case of quadratic disutility of the rate of effort, $\psi(x(t)) = x(t)^2/2$, one can derive the optimal delivery time schedule and associated wages explicitly:

PROPOSITION 3. *When the agent has quadratic disutility of the rate of effort, the optimal perfect information contract takes the form*

$$T^*(\theta) = r^{-1} \log \left[\frac{\theta\sqrt{2rS}}{\theta\sqrt{2rS} - rL} \right]$$
$$w^*(\theta) = \frac{rL^2}{2\theta^2(e^{rT^*(\theta)} - 1)}$$

PROOF. See Appendix B.1 ∎

Before continuing with the asymmetric information case, I will briefly discuss some comparative statics of the perfect information case. As noted in the previous chapter, this model is similar to that of Kamien and Schwartz (1971), Lucas (1971) and Grossman and Shapiro (1986). They consider a setup with a concave production function and linear cost function, and profits are thus concave. In the present setup, costs are strictly convex and production is linear, thus also leading to a concave profit function. I can thus compare my results to theirs.

Their main conclusions are the following: For interest rates $r > 0$, effort $x(t)$ is strictly increasing over time along the optimal path. Next, they conclude that an increase in the value of the completed project S decreases the optimal time to completion, while an increase in the difficulty of the project L increases the optimal time to completion. As will be shown shortly, these features will still hold under the optimal contract when there is asymmetric information.

3. Asymmetric Information Case

Let θ be distributed according to the differentiable, continuous, cumulative distribution function F on $[\underline{\theta}, \overline{\theta}]$ with density f, and assume that the monotone likelihood ratio property holds, i.e.

$$\frac{d}{d\theta}\left(\frac{F(\theta)-1}{f(\theta)}\right) \geq 0 \tag{3.1}$$

This assumption is necessary for the objective function of the principal to be concave. I use the revelation principle to design a direct revelation mechanism.

Note that the single crossing condition is satisfied. The agent reports type $\widehat{\theta}$ and is offered contract $\{w(\widehat{\theta}), T(\widehat{\theta})\}$. Incentive compatibility (truth telling) requires that

$$\theta \in \arg\max_{\widehat{\theta}} \left[w(\widehat{\theta}) - c(\theta, T(\widehat{\theta})) \right]$$

Denote by $U(\theta, \widehat{\theta}) = w(\widehat{\theta}) - c(\theta, T(\widehat{\theta}))$ type θ's rent when reporting to be of type $\widehat{\theta}$. The first order condition for truth telling (with respect to reported type $\widehat{\theta}$) is

$$U_2(\theta, \widehat{\theta}) = w'(\widehat{\theta}) - c'(\theta, T(\widehat{\theta}))(dT(\widehat{\theta})/d\widehat{\theta}) = 0 \tag{3.2}$$

Denote the rent when truth telling is induced by

$$U(\theta) = U(\theta, \theta) \tag{3.3}$$

The individual rationality constraints are then

$$U(\theta) \geqslant 0 \quad \forall \theta \in [\underline{\theta}, \overline{\theta}]$$

Totally differentiating (3.3), yields

$$\begin{aligned} dU(\theta) &= U_1(\theta, \theta) + U_2(\theta, \theta) \\ &= -c'(\theta, T(\theta)) + \left[w'(\theta) - c'(\theta, T(\widehat{\theta}))(dT(\widehat{\theta})/d\widehat{\theta}) \right] \\ &= -c'(\theta, T(\theta)) \end{aligned} \tag{3.4}$$

where the envelope theorem has been used to eliminate the bracketed expression. The second order condition for truth telling is thus

$$U_{22}(\theta,\theta) < 0 \tag{3.5}$$

Totally differentiating (3.2) evaluated at the equilibrium yields

$$dU_2(\theta,\theta) = U_{21}(\theta,\theta) + U_{22}(\theta,\theta) = 0 \Leftrightarrow$$

$$U_{22}(\theta,\theta) = -U_{21}(\theta,\theta) \tag{3.6}$$

From (3.5) and (3.6) it follows that

$$U_{21}(\theta,\theta) > 0$$

From expression (3.2) it follows that

$$U_{21}(\theta,\theta) = -c''(\theta,T(\theta))(dT(\theta)/d\theta) > 0 \Leftrightarrow$$

$$dT(\theta)/d\theta < 0$$

since $c''(\theta,T(\theta)) > 0$. This is the local second order condition. In appendix B.2, it is shown that incentive compatibility and the local second order condition are indeed necessary and sufficient conditions for truth telling to be a global optimum. I will temporarily assume that the second order condition holds and later verify that this is indeed the case. Since $U(\theta)$ is non-decreasing from incentive compatibility ($U_1(\theta,\theta) = -c'(\theta,T(\theta)) > 0$), one can replace the individual rationality constraints with $U(\underline{\theta}) \geqslant 0$. The principal thus solves the following programme:

$$\max_{\{T(\cdot),U(\cdot)\}} \int_{\underline{\theta}}^{\overline{\theta}} [V(T(\theta)) - c(\theta, T(\theta)) - U(\theta)]\, dF(\theta) \tag{3.7}$$

subject to

$$dU(\theta) = -c'(\theta, T(\theta))$$

$$U(\underline{\theta}) \geq 0$$

Now, integrating the incentive compatibility constraints (3.4) yields

$$U(\theta) = -\int_{\underline{\theta}}^{\theta} c'(\theta, T(\widetilde{\theta}))d\widetilde{\theta} - U(\underline{\theta})$$

Since the principal does not want to leave any rent for the most inefficient type, he sets $U(\underline{\theta}) = 0$.

Integrating by parts, it follows that

$$\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} U(\theta)dF(\theta) &= -\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} c'(\theta, T(\widetilde{\theta}))d\widetilde{\theta}dF(\theta) \\
&= -\left[[F(\theta) - 1] \int_{\underline{\theta}}^{\theta} c'(\theta, T(\widetilde{\theta}))d\widetilde{\theta} \right]_{\underline{\theta}}^{\overline{\theta}} + \int_{\underline{\theta}}^{\overline{\theta}} [F(\theta) - 1]\, c'(\theta, T(\theta))d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}} [F(\theta) - 1]\, c'(\theta, T(\theta))d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{F(\theta) - 1}{f(\theta)} \right) c'(\theta, T(\theta))dF(\theta)
\end{aligned}$$

Substituting this back into the principal's objective (3.7), one obtains the following maximisation problem:

$$\max_{T(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} \left[V(T(\theta)) - c(\theta, T(\theta)) - \left(\frac{F(\theta) - 1}{f(\theta)} \right) c'(\theta, T(\theta)) \right] dF(\theta)$$

Using point-wise maximisation yields the following optimality condition:

$$V'(T^*(\theta)) - c'(\theta, T^*(\theta)) = \left(\frac{F(\theta) - 1}{f(\theta)} \right) c''(\theta, T^*(\theta)) \tag{3.8}$$

This yields a familiar result. For $\theta = \overline{\theta}$,

$$V'(T^*(\overline{\theta})) = c'(\overline{\theta}, T^*(\overline{\theta}))$$

That is, the discounted marginal reduction in utility of waiting one instant, is exactly offset by the discounted marginal reduction in the cost resulting from waiting one instant. This is identical to equation (2.1). In other words, there is no inefficiency for the highest type. For all types $\theta < \overline{\theta}$,

$$V'(T^*(\theta)) < c'(\theta, T^*(\theta))$$

That is, for $\theta < \overline{\theta}$, the marginal cost function is steeper than the marginal benefit function. But this happens only for $t > T^*(\theta)$ (see figure 3). In other words, the optimal contract induces inefficient delay. Thus

PROPOSITION 4. *For all types $\theta < \overline{\theta}$, the optimal incentive compatible contract induces inefficient delay in project completion.*

Now consider how the optimal termination date function $T^*(\theta)$ varies with efficiency θ. Implicitly differentiating equation (3.8) with respect to θ and rearranging, it follows that

$$\frac{dT^*(\theta)}{d\theta} = \frac{\frac{d}{d\theta}\left(\frac{F(\theta)-1}{f(\theta)}\right) c''(\theta, T^*(\theta))}{V''(T^*(\theta)) - c''(\theta, T^*(\theta)) + \left(\frac{1-F(\theta)}{f(\theta)}\right) c'''(\theta, T^*(\theta))}$$

For the first order condition (3.2) to be sufficient for optimality, i.e. for the second order condition for truth telling (3.5) to be satisfied, it has to hold that $dT^*(\theta)/d\theta < 0$. Given the symmetric information second order condition (2.2) and the monotone likelihood ratio property (3.1), $T^*(\theta)$ is decreasing in θ.

An implication of the monotonicity of $T^*(\theta)$ is that it is invertible. Thus one may define the inverse function $\theta^*(T)$. From the function expressing the agent's rent (3.3), one may obtain an explicit expression for the optimal contract's wage schedule as a function of delivery time.

We know that $w(\theta) = U(\theta) + c(\theta, T(\theta))$. Substituting for the inverse $\theta^*(T)$, it follows that

$$w(\theta^*(T)) = U(\theta^*(T)) + c(\theta^*(T), T^*(\theta^*(T)))$$

It can also be determined how the optimal incentive compatible termination schedule changes with an increase in S, the value for the principal of the completed project. This can be done, since the completion schedule $T^*(\theta)$ depends on the principal's discounted value $V(T(\theta))$, which in turn depends on

S. Implicitly differentiating (3.8) with respect to S and rearranging yields

$$\frac{dT^*(\theta)}{dS} = \frac{re^{-rT^*(\theta)}}{V''(T^*(\theta)) - c''(\theta, T^*(\theta)) + \left(\frac{1-F(\theta)}{f(\theta)}\right) c'''(\theta, T^*(\theta))} < 0$$

Thus,

PROPOSITION 5. *Under the second order condition (2.2), an increase in the value of the completed project S strictly decreases the optimal time to completion.*

In the special case of two types of efficiency, $\theta \in \{\underline{\theta}, \overline{\theta}\}$ and quadratic disutility of effort, one can derive the optimal contract explicitly. For a detailed derivation, see appendices B.3 and B.4.

Instead of assuming unobservable efficiency θ, one could have assumed common knowledge of the production technology but asymmetric information about the difficulty of the project L. This would change nothing in the optimal contract. Grossman and Shapiro (1986) consider such a setup (without a principal-agent relationship). They assume that the probability of project completion increases in cumulated effort and conclude that the optimal rate of effort is still strictly increasing over time.

4. Discussion

The model presented here is a first attempt to understanding the choice of deadlines in long-term projects with asymmetric information. It is worth mentioning that an important assumption on which the analysis rests is that the principal can perfectly commit to any contract specified at the beginning of the game. Assuming perfect commitment amounts to disregarding any question of time consistency. An effect of relaxing this assumption is that immediately after the date when the most efficient type of agent is expected to deliver, i.e. at time $T^*(\overline{\theta})$, the principal would have incentives to renegotiate the contract. But this would be correctly anticipated by the agent, who would supply effort accordingly. In other words, the contract I have derived is not negotiation proof.

A second assumption is that effort (and progress) is unobservable to the principal. If one allows the principal to observe effort and renegotiate the initial contract, he would have strong incentives to use his acquired information on the agent's efficiency to expropriate informational rents. As this would also be correctly anticipated, agents of different types would pool, and exert low effort in order not to reveal their true type. This effect is known in the literature on dynamic adverse selection as the ratchet effect. Both these assumptions could be relaxed in principle, but since the analysis becomes extremely complicated even in discrete time models, doing so in this continuous time setup seems virtually impossible.

Nevertheless, the main result, i.e. that the choice of deadline can be used to screen agents, contributes to a deeper understanding of the effects of deadlines in settings of asymmetric information, and opens channels for further exploration of the issue[3].

[3] For examples of how a monopolist can use delivery time to screen consumers with unknown valuations, see Wilson (1993).

CHAPTER 7

A Theory of Optimal Deadlines II: Moral Hazard

In this chapter I propose a moral hazard model of project completion with time to build. In contrast to the model set out in chapter 6, the agent cannot perfectly control the date of completion, but can influence completion time probabilitically by exerting costly effort. Furthermore, the assumption of perfect commitment on the part of the principal is relaxed.

1. The Model

A principal hires an agent to complete a project involving n distinct phases or tasks to be performed in a fixed pre-specified order. Work on task i cannot be started before task $(i-1)$ has been completed. Attempting to complete a task takes one period. For each task, the agent chooses the probability $q \in [0,1]$ of successfully completing it, at personal disutility $\psi(q) > 0$ for $q > 0$, with $\psi' > 0$, $\psi'' > 0$ and $\psi''' \geq 0$. If task i fails, it has to be attempted again until success occurs[1]. Then the agent can attempt to complete the next task. It is

[1]Similar models have been considered by Sobel (1992) and Kremer (1993) who both assume that if task i fails, all previously completed tasks $j \leq i-1$ must be repeated. Importantly, they both abstract from incentive considerations.

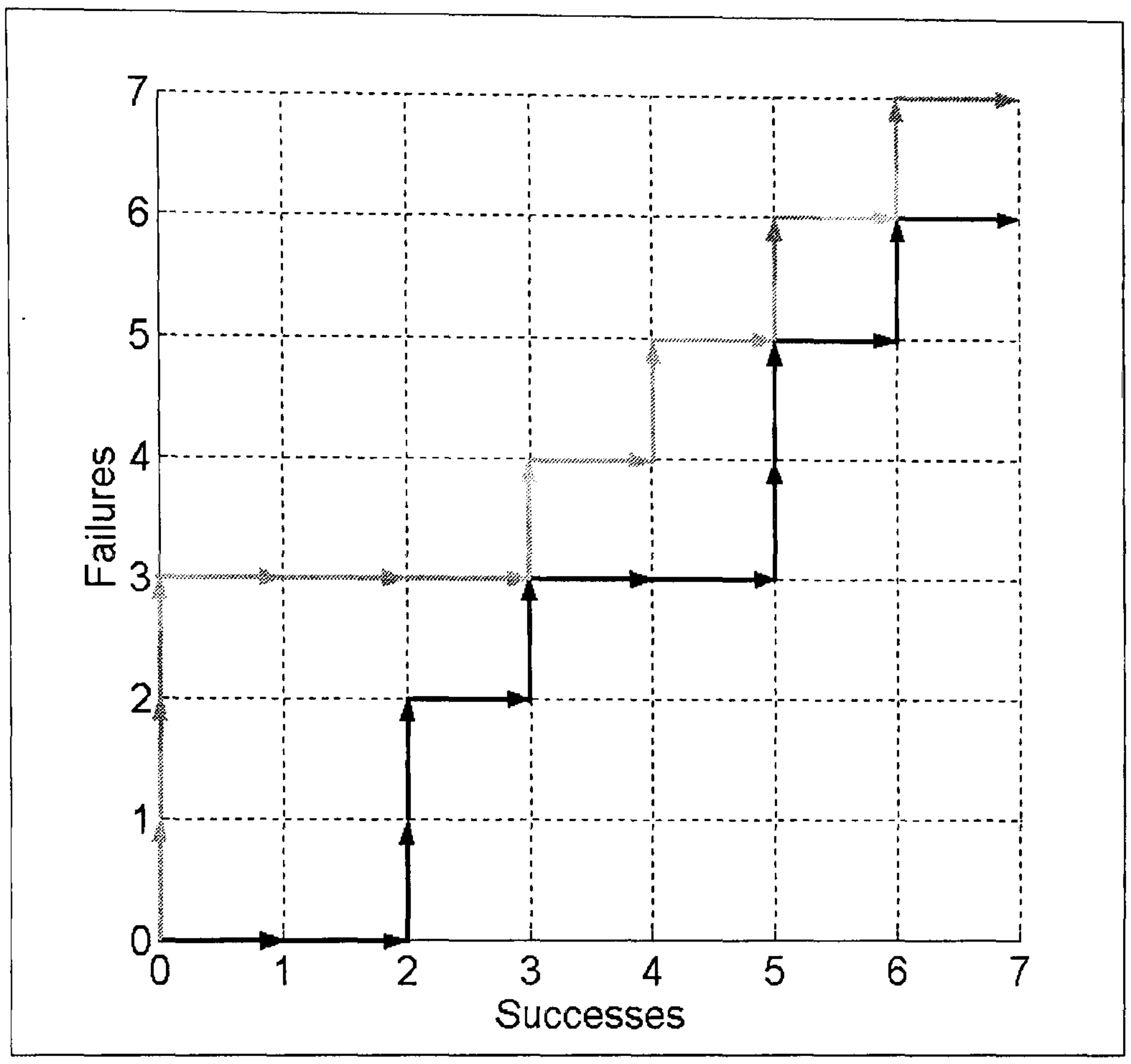

FIGURE 4. Progress and delay, $n = 7$.

assumed that

$$\lim_{q \to 1} \psi(q) = \infty$$

$$\psi(0) = 0$$

so that it will never be optimal to complete any task with certainty, and that remaining inactive is costless for the agent. On completion of task n, the principal receives value $S > 0$, which is discounted with factor $\delta \in [0, 1]$.

Assume that both the principal and the agent are risk neutral, but that the agent is protected by limited liability. Both principal and agent have separable utility functions, and the agent's outside utility is normalised to zero. Naturally, the date of completion of any task (and thus, also the completion of the project) is a random variable. Ceteris paribus, the principal wishes the agent to finish the project as soon as possible.

The number of tasks n to be completed will be referred to as the *scale* of the project, and the curvature of ψ will be referred to as the *difficulty* of the project. Figure 4 shows a *progress grid*, and two possible paths for the case $n = 7$. Movement in the horizontal direction denotes success while movement in the vertical direction denotes failure. Each vertex corresponds to one period of time. Hence, the path shown in black achieves project completion in 13 periods while the path in grey achieves completion in 14 periods. Of course, this does not imply that the black path displays smaller delay on all tasks.

2. Perfect Information Case

As a benchmark, consider the optimal contract when effort is perfectly observable and contractible[2]. Let V_i^{fb} denote the value of having successfully completed $(i-1)$ tasks, and define $V_{n+1}^{fb} \equiv S$, where superscript fb denotes first-best values. Therefore at each point in time, if $(i-1)$ tasks have been successfully completed, the principal wishes to maximise

$$V_i^{fb} = \max_{q_i} \left\{ -w + \delta \left[q_i V_{i+1}^{fb} + (1-q_i) V_i^{fb} \right] \right\} \tag{2.1}$$

subject to $w \geq \psi(q_i)$. Although this is not a finite horizon problem, it can still be solved using the Euler conditions due to the stationarity of the optimal

[2]A similar model is considered in Grossman and Shapiro (1986), but without moral hazard.

policy function[3]. This is done by using the boundary condition $V_{n+1}^{fb} \equiv S$ and solving backwards.

Taking the first order condition of (2.1), the optimal choice of effort is given implicitly by

$$\psi'(q_i^{fb}) = \delta\left[V_{i+1}^{fb} - V_i^{fb}\right] \tag{2.2}$$

The interpretation of (2.2) is that the agent is instructed to choose q such that the marginal cost is equal to discounted marginal benefit.

It is easily established that $V_{i+1}^{fb} > V_i^{fb}$ for all i. To see this, it is noted that if $V_i^{fb} \geq V_{i+1}^{fb}$ for any i, equilibrium effort would be $q_i^{fb} = 0$. But since the value of the project accrues only when task n is completed, $q_i^{fb} = 0$ for any $i < n$ implies that $q_1^{fb} = 0$ and the project is simply not undertaken.

Next, the equilibrium effort q_i^{fb} is increasing in V_{i+1}^{fb}. To see this, solve (2.1) for V_i^{fb} to get

$$V_i^{fb} = \frac{\delta q_i V_{i+1}^{fb} - \psi(q_i)}{1 - \delta(1 - q_i)}$$

Let the optimal policy function be given by

$$q_i^* \in \arg\max_{q_i}\left[\frac{\delta q_i V_{i+1}^{fb} - \psi(q_i)}{1 - \delta(1 - q_i)}\right] \tag{2.3}$$

[3] Recall that a policy is *stationary* if the map from states into actions is independent of time.

The optimal effort q_i^* is now characterised. Differentiating (2.3) with respect to q_i and rearranging, yields

$$V_{i+1}^{fb} = \frac{\psi'(q_i^*) - \delta\psi'(q_i^*) + \delta q_i^* \psi'(q_i^*) - \delta\psi(q_i^*)}{\delta(1-\delta)} \tag{2.4}$$

From (2.4) it follows that an increase in V_{i+1}^{fb} must correspond to an increase in the right hand side of the equation. This quantity is increasing in q_i^* since its derivative is[4]

$$\frac{\psi''(q_i^*)\left[1-\delta(1-q_i^*)\right]}{\delta(1-\delta)} > 0$$

Since $V_{i+1}^{fb} > V_i^{fb}$ for all i, it follows that for all i

$$q_{i+1}^{fb} > q_i^{fb}$$

This means that an increase in the value of the completion of task i will, ceteris paribus, increase the cost of delay. Consequently, this will lead to an increase in the chosen (equilibrium) probability of success. Therefore, using (2.2) it follows that

$$\psi'(q_i^{fb}) = \delta\left[V_{i+1}^{fb} - V_i^{fb}\right] > \delta\left[V_i^{fb} - V_{i-1}^{fb}\right] = \psi'(q_{i-1}^{fb})$$

[4]In effect, we have an implicit function of the type $y = g(q)$. Since $g' > 0$, the function g is monotone in q and thus invertible.

and thus for all i it holds that

$$V_{i+1}^{fb} - V_i^{fb} > V_i^{fb} - V_{i-1}^{fb}$$

Hence, the value of the project is convex in the amount of progress (number of completed tasks). In turn, this implies that effort is not only increasing as the project progresses, but also at an increasing rate. An immediate observation is that delays are more likely to occur at the early stages of development, where the value of the project is relatively low (since few costs have been sunk). Since effort is observable, there is no incentive compatibility constraint to consider. The wages are just set such that individual rationality is binding for each attempt at completion of a task, i.e. for all i

$$w_i^{fb} = \psi(q_i^{fb})$$

Last, expected completion time can be derived explicitly, for each task as well as for the overall project. First note that the stationarity of the policy function means that each time task i is attempted, the same effort level q_i will be implemented. But then work on task i constitutes a sequence of independent Bernoulli trials with probability of success q_i, constant over trials. Denote by X_i be the number of trials on task i until the first success. Then the probability function is given by

$$f(x_i) = P(X_i = x_i) = q_i(1 - q_i)^{x_i - 1}, \text{ for } x_i = 1, 2, ...$$

and the cumulative distribution function is

$$F(x_i) = \begin{cases} 0, \text{ for } -\infty < x_i < 1 \\ 1 - (1 - q_i)^{[x_i]}, \text{ for } 1 \leq x_i < \infty \end{cases}$$

where $[x_i]$ is the integer part of x_i. The distribution of X_i is then geometric with parameter q_i.

The mean and variance of termination time of task i are given by

$$E(X_i) = q_i^{-1}$$

$$V(X_i) = (1 - q_i)q_i^{-2}$$

The project as a whole is thus characterised by a sequence of random variables $\{X_i\}_{i=1}^{n}$ that follow the geometric distribution with parameter q_i which is increasing in i. The sum of the random variables X_i represents the total number of trials in all stages until the project is successfully completed. Hence the expected number of trials until the whole project is completed is

$$T_n = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} q_i^{-1} \tag{2.5}$$

Since it is assumed that the project is viable, and thus that $q_i > 0$ for all i, it follows that T_n is strictly increasing and convex in n, bounded below by n. In other words, the expected time of project completion is an increasing convex function of the scale of the project.

Let $r(q) \equiv -\psi(q)/\psi'(q)$ denote the difficulty of the project. This measure is the standard Arrow-Pratt coefficient, and measures the curvature of the disutility function ψ. The more sharply $\psi(q)$ increases in effort q, the more expensive it becomes to implement effort. In other words, as the difficulty $r(q)$ of the project increases, the induced effort q will decrease. But expected termination time is a decreasing convex function of effort. Thus an increase in the difficulty of the project induces a more than proportionate increase in the expected completion time. Last, the variance of T_n is a decreasing convex function of effort levels q_i.

Given gross project value S, the net value of the project decreases in n, which in turn decreases the implemented efforts q_i. Thus the larger the scale of the project n, the more variable is completion time. Last, since effort q_i is increasing in progress, the estimate of remaining development duration becomes less variable the closer the project is to completion.

3. Asymmetric Information Case

Next, I turn to the case where effort is unobservable to the principal. Since he cannot contract upon the q's, he must contract upon successful completion of tasks. As usual in dynamic contracting problems, one needs to specify the kind of contracts that the principal can commit to. I will start at one extreme by assuming that the principal has no commitment power at all, and that the relationship is governed by a sequence of spot contracts. Next, I will discuss the cases where the principal has more commitment. Specifically, I will discuss short-term commitment, i.e. contracts for the successful completion of a single task, and perfect commitment, i.e. where the principal can commit to a contract for the completion of any number of tasks $m \leq n$.

3.1. The No-Commitment Case: Spot Contracting. Consider the case where the principal has no commitment at all, and thus offers a new contract in each period. Note that if it had been assumed that contracts were written for each task as opposed to each period, it would implicitly have been assumed that the principal had some degree of commitment. Since the agent is protected by limited liability, the principal cannot pay him less than zero in any state of the world, i.e. even if the task fails. Keeping a similar notation as in the previous section, the principal has to solve

$$V_i^{sc} = \max\left\{-q_i w + \delta\left[q_i V_{i+1}^{sc} + (1-q_i)V_i^{sc}\right]\right\} \tag{3.1}$$

subject to

$$q_i^{sc} \in \arg\max_{q_i}\{-\psi(q_i) + q_i w\} \tag{3.2}$$

where superscript sc denotes spot contracting. Using the first-order approach, the incentive compatibility constraint (3.2) reduces to

$$w = \psi'(q_i) \tag{3.3}$$

Substituting (3.3) in (3.1) yields

$$V_i^{sc} = \max_{q_i} \left\{ -q_i\psi'(q_i) + \delta \left[q_i V_{i+1}^{sc} + (1-q_i) V_i^{sc} \right] \right\}$$

Differentiating this with respect to q gives the optimal policy ${q_i}^{sc}$:

$$\psi'(q_i^{sc}) = \delta \left[V_{i+1}^{sc} - V_i^{sc} \right] - q_i^{sc}\psi''(q_i^{sc}) \tag{3.4}$$

Note the difference to the optimal policy under observable effort. Under spot contracting with unobservable effort, the implemented effort does not equalise marginal disutility with discounted marginal increment in project value.

Next, $q_{i+1}^{sc} \geq q_i^{sc}$ for all i. This is established using arguments similar to those of the observable effort case. It is straightforward to show that equilibrium effort is increasing over time if

$$\left(2\psi''(q_i^{sc}) + \psi'''(q_i^{sc})\right)\left(1 - \delta(1 - q_i^{sc})\right) \geq 0$$

which is always satisfied given our assumptions. In turn, this implies that for all i

$$V_{i+1}^{sc} - V_i^{sc} > V_i^{sc} - V_{i-1}^{sc}$$

and the value of the project is thus also strictly increasing and convex in progress under spot contracting.

It is interesting to know if equilibrium effort is indeed larger under observable effort than under unobservable effort with spot contracting. The following holds:

PROPOSITION 6. *Equilibrium probabilities of success under spot contracting are lower than the first best probabilities.*

PROOF. Follows directly from revealed preferences ∎

Thus, $q_i^{sc} < q_i^{fb}$ for all i. This means that under spot contracting, the principal is inducing inefficiently low probability of success, thereby increasing the likelihood of task (and project) delay.

In the first best case, the principal could observe effort and thus pay the agent a wage equal to the disutility of effort. With spot contracting however, the agent only gets paid if there is success. Since the agent is risk neutral, the principal sets the wage such that the individual rationality constraint binds. That is, wages are given by

$$q_i^{sc} w_i^{sc} = \psi(q_i^{sc})$$

Note that $w_i^{sc} = \psi(q_i^{sc})/q_i^{sc} > \psi(q_i^{sc})$ so the agent earns a positive rent in case of success. The agent has no incentive to increase effort at this wage because $w_i^{sc} = \psi'(q_i^{sc})$. Last, recall that implemented effort increases in progress. This in turn implies that in the agent's rent in case of success increases in the progress of the project.

4. Discussion

The assumption of this chapter has so far has been that the principal has no commitment at all, and thus is unable to write long-term contracts. This may be an overly pessimistic assumption, given that long-term contracts are in fact observed in practice. It would thus be interesting to analyse what can be obtained by relaxing the no-commitment assumption. While a comprehensive treatment will be left for future research, the following sections contain some preliminary results and considerations.

4.1. Short-Term Commitment. To make progress payment feasible, one needs to impose at least some degree of commitment on the part of the principal. As a natural starting point, assume that the principal can commit to a contract that specifies payment after completion of a single task. Technically, this poses a potential problem in that the optimal policy function is no longer stationary, and standard dynamic programming techniques cannot be used. This is because the agent's wage dependents on the number of periods of delay d. To see this, consider the choice of effort at the beginning of the first period. The agent's expected future utility at the beginning of the game is given by

$$U_0 = -\psi(q_{11}) + q_{11}w(0) + \delta(1 - q_{11})\left[-\psi(q_{12}) + q_{12}w(1) + \delta(1 - q_{12})\left[...\right]\right]$$

where q_{ij} is effort exerted on jth attempt on task i. Since $w(d)$ is clearly not constant in d in the optimal contract, the optimal policy will not be stationary ($q_{ij} \neq q_{ik}$ for $j \neq k$). However, the problem can be transformed into an equivalent stationary one, by introducing some appropriate auxiliary variables, as follows:

$$U_0 = \sum_{t=0}^{\infty} \delta^t \left[-\psi(q_t) + w(d_t) I_{[0,q_t]}(s_t)\right] z_t$$

$$z_t = I_{[q_t,1]}(s_t) z_{t-1}, \quad z_{-1} = 1$$

$$d_t = d_{t-1} + I_{[q_t,1]}(s_t), \quad d_{-1} = 0$$

$$s_t \sim U[0,1]$$

where

$$I_{[a,b]}(x) = \begin{cases} 1 & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

The function z_t indicates when the game is still in play, and thus switches to zero after a successful attempt. Since the game "ends" after a success, $q_t = w(d_t) = 0$ for all $\tau \geq t$ if $z_t = 0$. The state variable d_t keeps track of the number of periods of delay until and including period t, and is a simple counting process. To solve the problem, the functional equation corresponding to the sequence problem needs to be determined. To proceed, start by defining the following mapping:

$$M(u)(1,d,s) = \max\left[-\psi(q)w(d) + \delta(1-q)u(1,d+1,s)\right]$$

The mapping M is a contraction on the space of bounded and continuous functions with the sup-norm, and this is verified by checking that Blackwell's sufficiency conditions hold. Let u, ν be bounded functions such that $u(x) \geq \nu(x)$ for all x. Thus

$$\begin{aligned} M(u)(1,d,s) &= \max\left[-\psi(q)w(d) + \delta(1-q)u(1,d+1,s)\right] \\ &\geq \max\left[-\psi(q)w(d) + \delta(1-q)\nu(1,d+1,s)\right] \\ &= M(\nu)(1,d,s) \end{aligned}$$

and therefore monotonicity is satisfied. Also, since

$$\begin{aligned} M(u+a)(1,d,s) &= \max\left[-\psi(q)w(d) + \delta(1-q)\left[u(1,d+1,s)+a\right]\right] \\ &= \max\left[-\psi(q)w(d) + \delta(1-q)u(1,d+1,s) + \delta(1-q)a\right] \\ &\leq \max\left[-\psi(q)w(d) + \delta(1-q)u(1,d+1,s)\right] + \max\left[\delta(1-q)a\right] \\ &= \max\left[-\psi(q)w(d) + \delta(1-q)u(1,d+1,s)\right] + \delta a \\ &= M(u)(1,d,s) + \delta a \end{aligned}$$

the mapping M also satisfies discounting. From the contraction mapping theorem it follows that M has a unique fixed point u^* such that $M(u^*) = u^*$. This fixed point will be the sought functional equation.

Note that in the first-best and spot contracting cases, one could solve the model using the Euler conditions. This is essentially because the problem is stationary, in the sense that once failure on a given task has occurred, the decision-maker faces the same problem again. The optimal choices are thus identical to last period's choices, just with a period's lag. One can thus work backwards using value at completion as the relevant boundary condition. In the case of commitment, there is no last period to work backwards from, and thus the Euler conditions fail to pin down the optimal policy function. I thus need to employ an alternative method when looking for the optimal policy function. Specifically, I need to use an iterative procedure whereby I start with a candidate value function, and show that one can improve upon the outcome by changing the value function. This procedure is repeated until one arrives at the unique maximum.

Having established that an optimal contract exists, I will informally discuss how this contract may look like. Consider the gross costs of delay. Assume that the agent's policy is stationary, and consider the principal's cost of delay, gross of any wage transfers to the agent. Define for task i and number of failures k

$$\Lambda_{i,k} \equiv (1 - \delta^k)V_i$$

as the total cost of $k \geq 0$ periods of delay in the termination of task i (reduction in value of the project due to discounting). Naturally, $\Lambda_{i,k} \geq 0$ for all i and k. Also, since $V_{i+1} > V_i$ for all i, it follows that

$$\Lambda_{i+1,k} > \Lambda_{i,k}$$

for all i and k. This means that the total cost of k periods of delay is increasing in progress. Last, note that

$$\lim_{k\to\infty} \Lambda_{i,k} = V_i$$

Hence, for perpetual delay the entire value of the project is lost as there is never any progress beyond tasks $j \leq (i-1)$.

Next, define the marginal cost of a further period of delay on task i when there have already been k periods of delay

$$\begin{aligned}
\lambda_{i,k} &\equiv \Lambda_{i,k} - \Lambda_{i,k+1} \\
&= (1-\delta^k)V_i - (1-\delta^{k+1})V_i \\
&= \delta^k(\delta - 1)V_i \leq 0 \quad \forall i, k
\end{aligned}$$

It is easily seen that for all i

$$\lim_{k\to\infty} \lambda_{i,k} = 0$$

so that marginal cost of delay is decreasing in the number of periods of delay.

Last, it is noted that for all i and k

$$\lambda_{i,k} > \lambda_{i+1,k}$$

and thus marginal cost of delay is increasing in progress, i.e. becomes more negative as project completion approaches. When writing the optimal contract, the principal will seek to specify the function such that the agent's expected wage decreases in delay at rate $\lambda_{i,k}$ subject to suitable incentive compatibility and individual rationality constraints. This leads to the following conjectures:

CONJECTURE 1. *The optimal wage function for given i will be a convex, decreasing function of delay.*

CONJECTURE 2. *The optimal wage function will be steeper the closer the agent is to project completion, i.e. a given delay will be "punished" harder for later tasks.*

4.2. Long-Term Commitment. Analysing the case of full commitment is potentially a very difficult exercise. In the case of short-term commitment (contracting for the completion of a single task), there is only one state variable to keep track of, namely the number of periods of delay d_t. When analysing longer commitment, say for $m \leq n$ tasks, there are two state variables, i and $\{d_j\}_{j=1}^{i}$. The first indicates which task the agent is currently working on, while the second keeps track of the experienced delays of all previously completed tasks.

In general, models of repeated moral hazard yield the result that the longer the duration of the interaction (i.e. the longer the horizon), the less severe does the incentive problem become. In other words, as the horizon becomes very large, the optimal full commitment contract approximately achieves the first best. This model is no different in this respect, but there is a slight twist. As the scale of the project tends to infinity, the implemented effort will tend to zero. In other words, the project is not undertaken, which is indeed the first best under infinite scale.

Alternatively, one can ask how incentives are influenced by the length of the contract. I suspect that as the number of tasks included in the contract increases, the moral hazard problem is reduced, thereby increasing efficiency. Specifically, if the principal can choose a contract which includes the completion of $m \leq n$ tasks, he will find it optimal to set $m = n$.

A strategy of a proof follows: When considering the optimal contract with spot commitment, the principal could only provide incentives by setting the wage in case of immediate success, i.e. no delay. When commitment is assumed, the principal can sweeten the deal by not only offering the agent a wage for completion of the current task, but also by promising the agent wages for future task completions and making this promise a function of current delay. In a sense, successful completion of a task has an options value, in that completion allows the agent to proceed with (and be rewarded for) further tasks. The more tasks are included in the contract, the larger this options value becomes. As the value of the option increases, the agent is prepared to pay a higher price (disutility of effort). It will therefore be cheaper for the principal to induce higher effort.

Technically, the proof would consist of showing that the short-term contract improves upon the spot contracts. Next, it would be shown by induction that any contract can be improved upon by adding the completion of one task further to the contract.

Another possible avenue to obtain and characterise the optimal long-term contract is that developed by Spear and Srivastava (1987) who study an infinitely repeated version of Holmstrom's (1979) seminal paper on moral hazard. They show that under suitable simplifying assumptions, the dynamic problem can be reduced to a relatively simple calculus of variations problem. This is done by letting the agent's expected future utility be the state variable in the principal's dynamic optimisation problem. A possible complication in using their approach is that in their model, the outcome is a continuous variable, and they can therefore characterise the optimal contract in terms of thresholds of the outcome such that the agent is rewarded for outcomes above the threshold and punished for outcomes below it. In the current setup, the outcome is discrete (i.e. success or failure). However, it may be possible to treat the number of periods of delay on a task as the outcome, and determine a threshold number of periods of delay d^* playing a similar role.

Another extension it would be interesting to analyse is to derive the optimal contract subject to budgetary horizons. That is, a situation where (for legal reasons) the principal cannot commit to contracts that extend beyond $h < n$ periods (the no-commitment scenario analysed above is the special case with $h = 1$). This type of constraint is somewhat similar to that of annual budget caps on development projects extending over several years. I conjecture that budgetary horizons will limit the effectiveness of contracts in providing incentives, and thus lead to inefficient delays in project completion.

Last, there may be an interesting application to staged financing in which a start-up firm contracts with a venture capitalist for the financing of a project. It may be possible to study the notion of milestone payments in a rigorous manner. This would be interesting since this type of payment schemes are often observed in practice, notably within software, defence and pharmaceuticals development.

Part 3

Appendices

APPENDIX A

Appendices to Part 1

1. Proof of Lemmata 1 and 2

Consider the space of function on $Y_t \times X_t \times \mathbb{R}$ and define the operator $M : S[Y_t \times X_t \times \mathbb{R}] \to S[Y_t \times X_t \times \mathbb{R}]$ by

$$MV(z_t, x_t, \theta_t) = \max\{R(z_t, x_t, \theta_t), \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t]\}$$

Fix a sequence of strategies, and by implication a sequence z^t. It will now be shown that for each t, M is a contraction mapping on the space $S[Y_t \times X_t \times \mathbb{R}]$ with the sup-norm. Let $V(z_t, x_t, \theta_t) > \widehat{V}(z_t, x_t, \theta_t)$ for all (z_t, x_t, θ_t). Then

$$\begin{aligned} MV(z_t, x_t, \theta_t) &= \max\{R(z_t, x_t, \theta_t), \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t]\} \\ &\geq \max\{R(z_t, x_t, \theta_t), \delta E[\widehat{V}(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t]\} \\ &= M\widehat{V}(z_t, x_t, \theta_t) \end{aligned}$$

Thus the mapping M satisfies monotonicity. Next, let $a > 0$. Thus

$$\begin{aligned} M[V(z_t, x_t, \theta_t) + a] &= \max\{R(z_t, x_t, \theta_t), \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) + a \,|\theta_t]\} \\ &= \max\{R(z_t, x_t, \theta_t), \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t] + \delta a\} \\ &\leq MV(z_t, x_t, \theta_t) + \delta a \end{aligned}$$

and the mapping M satisfies discounting. Therefore, by Blackwell's sufficiency conditions, M is a contraction mapping (with modulus δ) on $S[Y_t \times X_t \times \mathbb{R}]$. Since by assumption $R(z_t, x_t, \theta_t)$ is bounded and continuous in all arguments, it follows by the contraction mapping theorem that there exists a unique fixed point $V(z_t, x_t, \theta_t)$ such that $MV(z_t, x_t, \theta_t) = V(z_t, x_t, \theta_t)$, and furthermore that this fixed point is bounded and continuous in (z_t, x_t, θ_t) (see e.g. Stokey and Lucas, 1989 for details). Assume throughout that $x_t > 0$ and fix a sequence z^t.

Recall that by assumption $R(z_t, x_t, \theta_t)$ is strictly increasing in θ_t, weakly decreasing in z_t and weakly increasing in x_t. For $z_t < x_t$ it follows by the assumption of first-order stochastic dominance that the fixed point $V(z_t, x_t, \theta_t)$ is strictly increasing in θ_t. Also, for $\theta_t > \underline{\theta}$, $V(z_t, x_t, \theta_t)$ is weakly decreasing in z_t and weakly increasing in x_t.

Recall that $V(z_t, x_t, \theta_t) \geq 0$ for all θ_t. That is, an acquirer can always secure himself a payoff of zero by waiting indefinitely. On the other hand, $R(z_t, x_t, \theta_t) < 0$ for $\theta_t < \underline{\theta}$ which implies that in this range of the economic fundamental it is optimal to wait, i.e. $V(z_t, x_t, \theta_t) = \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \, |\theta_t]$. Now let $\theta_t \geq \underline{\theta}$ and consider an increase in θ_t. Both the value of raiding and that of waiting will increase. A simple argument shows that for sufficiently high θ_t, the value of raiding overtakes that of waiting. Assume that for all θ_t

$$\delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \, |\theta_t] \geq R(z_t, x_t, \theta_t) > 0$$

Specifically, this implies

$$\sup_{\theta_t} V(z_t, x_t, \theta_t) = \sup_{\theta_t} \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t]$$

which contradicts $\delta \in]0, 1[$. The assumption that $R(z_t, x_t, \theta_t) > \delta E[R(z_{t+1}, x_{t+1}, \theta_{t+1})|\theta_t]$ ensures that there is a unique crossing since it implies that the value of raiding increases at a higher rate than the value of waiting. In conclusion, for each sequence z^t there exists a unique finite $\widetilde{\theta}(z^t) \in]\underline{\theta}, \infty[$ such that

$$R(z_t, x_t, \widetilde{\theta}(z^t)) = \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \left| \widetilde{\theta}(z^t)\right)]$$

Since $V(z_t, x_t, \theta_t)$ is weakly increasing in x_t, so is $\widetilde{\theta}(z^t)$. Similarly, since $V(z_t, x_t, \theta_t)$ is weakly decreasing in z_t, $\widetilde{\theta}(z^t)$ is weakly decreasing in z_t. This also holds for any future measure of raiders, and thus $\widetilde{\theta}(z^t)$ is also weakly decreasing in z^t. The first-best trigger $\overline{\theta}(z^t)$ is just the value of $\widetilde{\theta}(z^t)$ for the sequence z^t with $z_s = 0$ for all $s \geq t$. Last, note that for all $\theta_t > \underline{\theta}$,and $z_t \geq x_t$, $R(z_t, x_t, \theta_t) > 0$ while $V(z_{t+1}, x_{t+1}, \theta_{t+1}) = 0$.Thus there exists a unique $z_t^* \in [x_t, y_t]$ such that

$$R(z_t^*, x_t, \theta_t) = \delta E[V(z_{t+1}, x_{t+1}, \theta_{t+1}) \,|\theta_t]$$

It follows from the discussion above that z_t^* is weakly decreasing in θ_t and weakly increasing in x_t.

2. Proof of Lemma 3

To prove Lemma 3, two separate results need to be established. First, it is shown that there is a unique signal such that indifference obtains exactly when receiving signal $s_t = k_t^*$. Second, it is shown that for lower signals waiting is optimal, while for higher signals raiding is optimal.

Denote by $\Delta_\sigma^*(s,k)$ the expected payoff gain to waiting after having received signal s and believing that all other players use strategies with cutoffs k. It follows that

$$\Delta_\sigma^*(s,k) \equiv \int_{-\infty}^{\infty} \Delta\left(y\left[1-F(\frac{k-\theta}{\sigma})\right], x, s\right)\frac{1}{\sigma}f(\frac{s-\theta}{\sigma})d\theta$$
$$= -\int_{-\infty}^{\infty} \Delta\left(y\left[1-F(\frac{k-\theta}{\sigma})\right], x, s\right)\frac{1}{\sigma}dF(\frac{s-\theta}{\sigma})$$

Now, set $s = k$. From $z = y\left[1 - F(k-\theta)\right]$, it follows that $F(\frac{k-\theta}{\sigma}) = (y-z)/y$ and thus

$$dF(\frac{k-\theta}{\sigma}) = \frac{-dz}{y}$$

Performing this change of variables then yields

$$\Delta_\sigma^*(k,k) = \int_0^y \Delta\left(z,x,k\right)\frac{1}{y}dz$$

In other words, the function $\Delta_\sigma^*(s,k)$ has been rewritten such that it is an integral over a uniform distribution of z over $[0,y]$. But generically, there is a unique k^* that solves

$$\int_0^y \Delta\left(z, x, k^*\right) \frac{1}{y} dz = 0$$

Thus that there is exactly one cutoff signal k^* at which an agent is exactly indifferent between raiding and waiting. It now has to be verified that there exists an equilibrium where the agent raids whenever $s > k^*$ and waits whenever $s < k^*$. In order to do this, recall that the game displays *action single crossing* (follows from Lemma 2), *state monotonicity* (follows from Lemma 1) and the following monotone likelihood ratio property (which follows from the assumption of affiliated variables): For $a > b$, $f(a-\theta)/f(b-\theta)$ is increasing in θ.

The expected payoff gain to waiting, given signal s, when all other players use cutoffs k is given by

$$\begin{aligned}
\Delta_\sigma^*(s,k) &\equiv \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{s-\theta}{\sigma}\right) \Delta\left(y\left(1-F\left(\frac{k-\theta}{\sigma}\right)\right), x, s\right) d\theta \\
&= \int_{-\infty}^{\infty} f\left(\frac{s-k}{\sigma} - m\right) \Delta\left(y\left[1-F(-m)\right], x, s\right) dm
\end{aligned}$$

by changing variables so that $m = (\theta - k)/\sigma$. Now rewrite the above expression as

$$\begin{aligned}
\Delta_\sigma^*(s,k) &= h(s,k,s') \\
&\equiv \int_{-\infty}^{\infty} g(m,s')\widetilde{f}(s,m) dm
\end{aligned}$$

where

$$g(m, s') = \Delta(1 - F(-m), x, s')$$

$$\widetilde{f}(s, m) = f(\frac{s-k}{\sigma} - m)$$

Because of the monotone likelihood ratio property, $h(., k, s')$ preserves the single crossing property of $\Delta(z, x, \theta)$ by a result in Athey (2002). That is, there exists $s^*(k, s')$ such that

$$h(s, k, s') < 0 \quad if \quad s > s^*(k, s')$$

$$h(s, k, s') > 0 \quad if \quad s < s^*(k, s')$$

By state monotonicity, $h(s, k, s')$ is strictly decreasing in s'. Now let $s > s'$ and suppose that

$$h(s, k, s) = 0$$

It follows that

$$h(s', k, s') > h(s', k, s) > h(s, k, s) = 0$$

where the first inequality comes from state monotonicity and the second comes from the action single crossing property. A symmetric argument holds for $s < s'$.

Thus

$$h(s,k,s) < h(s',k,s') \quad if \quad s > s'$$

$$h(s,k,s) > h(s',k,s') \quad if \quad s < s'$$

$$h(s,k,s) = h(s',k,s') = 0 \quad if \quad s = s'$$

This implies that there exists a best response function $\beta : \mathbb{R} \to \mathbb{R}$ such that

$$\Delta^*_\sigma(s,k) < 0 \quad if \quad s > \beta(k)$$

$$\Delta^*_\sigma(s,k) = 0 \quad if \quad s = \beta(k)$$

$$\Delta^*_\sigma(s,k) > 0 \quad if \quad s < \beta(k)$$

But there exists a unique k^* that solves

$$\Delta^*_\sigma(k^*,k^*) = \int_0^y \frac{1}{y}\Delta(z,x,k^*)dz = 0$$

Therefore $\beta(k) = k$. It has thus been shown that with a uniform prior, there exists a unique equilibrium in cutoff strategies such that

$$I_s(k) = \begin{cases} 1 & if \quad s > k^* \\ 0 & if \quad s < k^* \end{cases}$$

This proves Lemma 3 ∎

3. Proof of Lemma 4

It was shown in Lemma 3 that in the associated game with uniform prior and private values, there is a unique equilibrium sequence of cutoffs. What remains to be shown is that the game with general prior and private values comes arbitrarily close to the associated private values uniform priors game, as noise vanishes.

Assume that the realisations of the state are governed by G_t, and that payoffs depend on the realised state, and not the signals. Recall that $\Delta_\sigma(s,k)$ is the expected payoff gain to waiting when signal s has been observed and all other players use cutoffs k. With general prior, this is given by

$$\Delta_\sigma(s,k) = \frac{\int_{-\infty}^{\infty} g(\theta) f(\frac{s-\theta}{\sigma}) \Delta\left(y\left(1 - F\left(\frac{k-\theta}{\sigma}\right), x, \theta\right)\right) d\theta}{\int_{-\infty}^{\infty} g(\theta) f(\frac{s-\theta}{\sigma}) d\theta}$$

Changing variables so $u = (s-\theta)/\sigma$, yields

$$\Delta_\sigma(s,k) = \frac{\int_{-\infty}^{\infty} g(s-\sigma u) f(u) \Delta\left(y\left(1 - F\left(\frac{k-s}{\sigma} + u\right), x, s - \sigma u\right)\right) du}{\int_{-\infty}^{\infty} g(s-\sigma u) f(u) du}$$

Letting $m = (k-s)/\sigma$ yields

$$\Delta_\sigma(s, s+\sigma m) = \frac{\int_{-\infty}^{\infty} g(s-\sigma u) f(u) \Delta\left(y\left(1 - F\left(m + u\right), x, s - \sigma u\right)\right) du}{\int_{-\infty}^{\infty} g(s-\sigma u) f(u) du}$$

Evaluating at $\sigma = 0$ it follows that

$$\Delta_0(s,s) = \int_{-\infty}^{\infty} f(u) \Delta\left(y\left(1 - F\left(m+u\right)\right), x, s\right) du$$

Finally, substituting for u and m gives

$$\Delta_0(s,s) = \int_{-\infty}^{\infty} \frac{1}{\sigma} f(\frac{s-\theta}{\sigma}) \Delta\left(y\left(1 - F\left(\frac{s-\theta}{\sigma}\right), x, s\right)\right) d\theta$$

In other words, when noise is exactly zero, the common values game where history is informative is equal to the private values game with a uniform prior. What remains to be shown is that $\Delta_\sigma(s, s - \sigma u) \rightarrow \Delta^*_\sigma(s, s - \sigma u)$ uniformly as $\sigma \rightarrow 0$, where $u = (s - k)/\sigma$. In other words, one must ensure that the equivalence of the two games is not a result of a discontinuity at $\sigma = 0$. Instead of showing uniform convergence directly, I will proceed by showing convergence with respect to the uniform convergence norm. Convergence in this norm implies uniform convergence. First, note that there exist extreme signals $\underline{s}$ and $\overline{s}$ such that for all k: $\Delta_\sigma(s,k) > 0$ for $s < \underline{s}$ and $\Delta_\sigma(s,k) < 0$ for $s > \overline{s}$. This follows from the existence of dominance regions $[-\infty, \underline{\theta}]$ and $[\overline{\theta}, \infty]$, where there is a unique optimal action. One can thus pick any pair $\underline{s}$ and $\overline{s}$ such that $\underline{s} < \underline{\theta}$ and $\overline{s} > \overline{\theta}$, and restrict attention to the compact interval $S \equiv [\underline{s}, \overline{s}]$. Since S is compact and the second argument of the Δ_σ function is continuous with respect to s (i.e. the function $s - \sigma u$), the set $K \equiv [\underline{s} - \sigma u, \overline{s} - \sigma u]$ is also compact. Hence $\Delta_\sigma(s,k)$ takes values in a compact set. Next, define the sup-norm (or *uniform convergence norm)*

$$||\Delta|| \equiv \sup_{s,k}\{|\Delta(s,k)|\}$$

It has to be shown that $\Delta_\sigma(s,k)$ is continuous in the uniform convergence topology. I start by showing continuity of $\Delta_\sigma(s,k)$ with respect to the Euclidean metric. Fix s', k'. Since the function is continuous in both arguments, it follows that

$$\forall \varepsilon_1 > 0, \exists \delta_1 : |s - s'| < \delta_1 \Rightarrow |\Delta_\sigma(s,k) - \Delta_\sigma(s',k)| < \varepsilon_1 \forall k$$

$$\forall \varepsilon_2 > 0, \exists \delta_2 : |k - k'| < \delta_2 \Rightarrow |\Delta_\sigma(s,k) - \Delta_\sigma(s,k')| < \varepsilon_2 \forall s$$

This in turn implies that

$$\sqrt{(s-s')^2 + (k-k')^2} < \delta \equiv \sqrt{\delta_1^2 + \delta_2^2}$$

But then by the triangle inequality it follows that

$$\begin{aligned} |\Delta_\sigma(s,k) - \Delta_\sigma(s',k')| &= |\Delta_\sigma(s,k) - \Delta_\sigma(s',k) + \Delta_\sigma(s',k) - \Delta_\sigma(s',k')| \\ &\le |\Delta_\sigma(s,k) - \Delta_\sigma(s',k)| + |\Delta_\sigma(s',k) - \Delta_\sigma(s',k')| \\ &\le \varepsilon_1 + \varepsilon_2 \equiv \varepsilon \end{aligned}$$

and continuity with respect to the Euclidean metric follows. Denoting by $\mathbf{C}(S \times K)$ the space of continuous functions over $S \times K$, it follows that $\Delta_\sigma(s,k) \in \mathbf{C}(S \times K)$. But showing uniform convergence is equivalent to showing that as

$\sigma \longrightarrow 0$,

$$||\Delta_{\sigma_\eta} - \Delta^*_{\sigma_\eta}|| = ||\Delta_{\sigma_\eta} - \Delta_0||$$

$$= \sup_{s,k} \{|\Delta_\sigma(s,k) - \Delta_0(s,k)|\} \longrightarrow 0$$

with respect to the sup-norm. By substituting for the relevant functions and taking limits, the result follows ∎

APPENDIX B

Appendices to Part 2

1. Proof of Proposition 3

From the agent's production function, it follows that

$$x(t) = \theta^{-1} z'(t) \tag{1.1}$$

Since the principal knows the agent's efficiency parameter θ, he will write a contract such that the agent's participation constraint is binding. He thus wants to choose T such as to maximise

$$\begin{aligned} & e^{-rT} S - \int_0^T e^{-rt} \frac{x(t)^2}{2} dt \\ & = e^{-rT} S - \frac{1}{2} \int_0^T e^{-rt} [\theta^{-1} z'(t)]^2 dt \end{aligned} \tag{1.2}$$

where (1.1) has been used. The boundary conditions are $z(0) = 0$ and $z(T) = L$.

Since the objective function (1.2) depends on t and $z'(t)$ only (i.e. not on $z(t)$), the Euler equation is

$$z'(t) = c_0 \theta e^{rt} \tag{1.3}$$

where c_0 is a constant to be determined. Integrating and rearranging yields

$$z(t) = \frac{c_0 \theta}{r}(e^{rt} - 1) \tag{1.4}$$

where the boundary condition $z(0) = 0$ has been used to determine the constant of integration $c_1 = -1$. Since T is chosen freely, the following transversality condition has to hold:[1]

$$e^{-rT} \frac{[\theta^{-1} z'(t)]^2}{2} = re^{-rT}$$

Rewriting this condition, it follows that

$$z'(T) = \theta \sqrt{2rS} \tag{1.5}$$

From (1.3) and (1.5), it follows that

$$c_0 \theta e^{rT} = \theta \sqrt{2rS} \tag{1.6}$$

Using the boundary condition $z(T) = L$ in (1.4) yields

$$c_0 \theta (e^{rT} - 1) = rL \tag{1.7}$$

The system of two equations (1.6)-(1.7) in the two unknowns c_0 and T can then be solved to get the optimal termination date $T^*(\theta)$.

[1]See Kamien & Schwartz (1991).

I will now determine the principal's cost of completion as a function of the completion date T. For given T, I have to find the path $x(t)$ that minimises

$$\int_0^T e^{-rt} \frac{[\theta^{-1} z'(t)]^2}{2} dt$$

The Euler equation is

$$z'(t) = c_3 \theta e^{rt}$$

Integrating and rearranging yields

$$z(t) = c_3 \left(\frac{\theta}{r}\right) [e^{rt} + c_4] \tag{1.8}$$

This, and the boundary conditions yield

$$c_3 = \frac{rL}{\theta(e^{rT} - 1)}$$

$$c_4 = -1$$

Substituting these in (1.8) gives

$$z(t) = \frac{(e^{r_A t} - 1)L}{e^{r_A T} - 1} \Rightarrow$$

$$z'(t) = \frac{e^{rt} rL}{e^{rT} - 1}$$

Therefore, from (1.1) the effort path is given by

$$x(t) = \frac{e^{rt} r L}{\theta(e^{rT} - 1)}$$

Note that effort is strictly increasing over time. This is an artefact of discounting. Also note that effort is a strictly decreasing function of efficiency θ. The total discounted disutility of effort is thus:

$$\begin{aligned} c(\theta, T(\theta)) &= \frac{1}{2} \int_0^T e^{-rt} x(t)^2 dt = \frac{1}{2} \int_0^T e^{-rt} \left[\frac{e^{rt} r L}{\theta(e^{rT} - 1)} \right]^2 dt \\ &= \frac{r L^2}{2\theta^2 (e^{rT} - 1)} \end{aligned}$$

From this, it is easy to verify that

$$c'(\theta, T) < 0, \quad c''(\theta, T) > 0, \quad c'''(\theta, T) < 0$$

This completes the proof ∎

2. Sufficiency of Monotonicity Condition

This proof follows that in Laffont and Tirole (1998) and is included for completeness. Given that

$$U_2(\theta,\theta) = 0$$

$$dT(\theta)/d\theta \leq 0$$

hold, truth telling, i.e. $\widehat{\theta}^*(\theta) = \theta$ is a global optimum for type θ (the local second order condition holds globally). To see this, assume that type θ strictly prefers to report type $\widehat{\theta}$ instead of θ:

$$U(\theta,\widehat{\theta}) > U(\theta,\theta) \Longleftrightarrow$$

$$\int_{\theta}^{\widehat{\theta}} U_2(\theta,x)dx > 0$$

Using the first order condition for truth telling (3.2), this is equivalent to writing

$$\int_{\theta}^{\widehat{\theta}} \left[U_2(\theta,x) - U_2(x,x)\right] dx > 0 \Longleftrightarrow$$

$$\int_{\theta}^{\widehat{\theta}} \int_{x}^{\theta} U_{12}(y,x)dydx > 0$$

From equations (3.5) and (3.6), it follows that

$$U_{21}(y,x) = -c''(y,T(x))(dT(x)/dx) > 0 \Rightarrow$$

$$-\int_{\theta}^{\widehat{\theta}} \int_{x}^{\theta} c''(y,T(x))(dT(x)/dx)dydx > 0$$

Now, if $\widehat{\theta} > \theta$, it follows that for all $x \in [\theta, \widehat{\theta}]$, $x > \theta$. This contradicts the inequality. Also, if $\widehat{\theta} \leq \theta$, it holds that for all $x \in [\widehat{\theta}, \theta]$, $x \leq \theta$. This also contradicts the inequality. The conclusion is that reporting the truth is a global optimum ∎

3. Euler and Transversality Conditions for the Two-Type Case

In this appendix, I develop the tools needed to solve the asymmetric information two type case. The problem is non-standard as it involves an objective function with two free end times as opposed to the standard free end time problems as presented for example in Kamien and Schwartz (1991).

I have to choose a pair of times (t_1, t_2) such as to maximise an objective of the general form

$$\int_{t_0}^{t_1} F(t, x, y, x', y')dt + G(t_1, t_2, x_1, y_2) + \int_{t_0}^{t_2} H(t, x, y, x', y')dt + J(t_1, t_2, x_1, y_2)$$

where $(x_0, y_0, x_1, y_2) = (x(t_0), y(t_0), x(t_1), y(t_2))$ are given boundary conditions. All functions are twice continuously differentiable. Let $x(t)$ be defined on $[t_0, t_1]$ and $y(t)$ on $[t_0, t_2]$. Letting $\widetilde{t} \equiv \max\{t_1, t_2\}$, extend $x(t)$ or $y(t)$ so that they share domain $[t_0, \widetilde{t}]$. Let $(x^*(t), y^*(t))$ defined on $[t_0, \widetilde{t}]$ be optimal paths. Now consider a pair of arbitrary, nearby admissible comparison paths $(x(t), y(t))$ defined on $[t_0, t_1]$ and $[t_0, t_2]$ respectively. Define the difference functions $h(t) = x(t) - x^*(t)$ and $m(t) = y(t) - y^*(t)$ on $[t_0, t_1 + \delta t_1]$ and $[t_0, t_2 + \delta t_2]$ respectively. Extend either so they share domain $[t_0, \widetilde{t} + \delta \widetilde{t}]$. By definition, $h(t_0) = m(t_0) = 0$. For ease of notation let $b = (x, y)$, $l = (h, m)$, $b' = (x', y')$, $l' = (h', m')$, $\overline{b} = (x_1, y_2)$ and $\overline{s} = (t_1, t_2)$. This can now be used to rewrite the objective as

$$\int_{t_0}^{t_1} F(t, b, b')dt + G(\overline{s}, \overline{b}) + \int_{t_0}^{t_2} H(t, b, b')dt + J(\overline{s}, \overline{b})$$

Now evaluate this along the families of trajectories $x^*(t)+ah(t)$ and $y^*(t)+am(t)$ on $[t_0, t_1 + a\delta t_1]$ and $[t_0, t_2 + a\delta t_2]$ respectively. This yields a function of the variable a alone:

$$g(a) = \int_{t_0}^{t_1+a\delta t_1} F(t, b^* + al, b^{*\prime} + al')dt + G(\overline{s} + a\delta\overline{s}, \overline{b} + a\delta\overline{b})$$
$$+ \int_{t_0}^{t_2+a\delta t_2} H(t, b^* + al, b^{*\prime} + al')dt + J(\overline{s} + a\delta\overline{s}, \overline{b} + a\delta\overline{b})$$

Since the paths w^* are optimal by assumption, it must hold that $g'(0) = 0$, which is the usual first order condition. Now use Leibnitz's rule whose general form is

$$g(a) = \int_{u(a)}^{v(a)} f(t, a)dt \Rightarrow$$
$$g'(a) = f(v(a), a)v'(a) - f(u(a), a)u'(a) + \int_{u(a)}^{v(a)} \frac{\partial f(t, a)}{\partial a}dt$$

First, apply Leibnitz's and the chain rule to the first term

$$g(a)_1 = \int_{t_0}^{t_1+a\delta t_1} F(t, b^* + al, b^{*\prime} + al')dt$$

to obtain

$$g_1'(a) = F(t_1 + a\delta t_1, b^* + al, b^{*'} + al')\,|_{t_1+a\delta t_1}\, \delta t_1$$
$$+ \int_{t_0}^{t_1+a\delta t_1} (F_x h + F_y m + F_{x'} h' + F_{y'} m')dt$$

and similarly,

$$g_2'(a) = H(t_2 + a\delta t_2, b^* + al, b^{*'} + al')\,|_{t_2+a\delta t_2}\,\delta t_2$$

$$+ \int_{t_0}^{t_2+a\delta t_2} (H_x h + H_y m + H_{x'} h' + H_{y'} m')dt$$

By the chain rule, the derivatives of the gain functions are

$$\frac{\partial G(t_1 + a\delta t_1, t_2 + a\delta t_2, x_1 + a\delta x_1, y_2 + a\delta y_2)}{\partial a}$$

$$= |G_t\delta t + G_s\delta s + G_x\delta x + G_y\delta y|_{(t_1+a\delta t_1, t_2+a\delta t_2, x_1+a\delta x_1, y_2+a\delta y_2)}$$

$$\frac{\partial J(t_1 + a\delta t_1, t_2 + a\delta t_2, x_1 + a\delta x_1, y_2 + a\delta y_2)}{\partial a}$$

$$= |J_t\delta t + J_s\delta s + J_x\delta x + J_y\delta y|_{(t_1+a\delta t_1, t_2+a\delta t_2, x_1+a\delta x_1, y_2+a\delta y_2)}$$

Substituting these expression back and evaluating at $a = 0$, yields

$$g'(0) = \int_{t_0}^{t_1} (F_x h + F_y m + F_{x'} h' + F_{y'} m')dt$$

$$+ F\delta t_1\,|_{t_1} + |G_t\delta t + G_s\delta s + G_x\delta x + G_y\delta y|_{(t_1,t_2)}$$

$$+ \int_{t_0}^{t_2} (H_x h + H_y m + H_{x'} h' + H_{y'} m')dt$$

$$+ H\delta t_2\,|_{t_2} + |J_t\delta t + J_s\delta s + J_x\delta x + J_y\delta y|_{(t_1,t_2)}$$

$$= 0$$

Breaking up the integrals and integrating by parts the terms containing derivatives with respect to x' and y', it follows after rearranging that

$$
\begin{aligned}
g'(0) = &\int_{t_0}^{t_1} \left((F_x - \frac{dF_{x'}}{dt})h + (F_y - \frac{dF_{y'}}{dt})m \right) dt \\
&+ \int_{t_0}^{t_2} \left((H_x - \frac{dH_{x'}}{dt})h + (H_y - \frac{dH_{y'}}{dt})m \right) dt \\
&+ |F_{x'}h + F_{y'}m + F\delta t_1|_{t_1} + |G_t\delta t_1 + G_s\delta t_2 + G_x\delta x_1 + G_y\delta y_2|_{(t_1,t_2)} \\
&+ |H_{x'}h + H_{y'}m + H\delta t_2|_{t_2} + |J_t\delta t_1 + J_s\delta t_2 + J_x\delta x_1 + J_y\delta y_2|_{(t_1,t_2)} \\
&= 0
\end{aligned}
$$

Now exploit that the admissible comparison functions were arbitrary. That is, pick functions such that $m(t_1) = m(t_2) = h(t_1) = h(t_2) = 0$. The first and last of these equalities follow directly from the definition of admissible comparison functions. This yields the following equation:

$$
\begin{aligned}
&\int_{t_0}^{t_1} \left[(F_x - \frac{dF_{x'}}{dt})h + (F_y - \frac{dF_{y'}}{dt})m \right] dt \\
&+ \int_{t_0}^{t_2} \left[(H_x - \frac{dH_{x'}}{dt})h + (H_y - \frac{dH_{y'}}{dt})m \right] dt \\
&= 0
\end{aligned}
$$

Since the functions $h(t)$ and $m(t)$ are arbitrary, one cannot generally expect them to be zero. One must therefore look at the integrands. A necessary condition for the equation to hold is that each term in the integrands are equal to zero for all t. As expected, this is just yields the usual Euler equations. One can thus conclude that the Euler equations are still necessary conditions for optimality. Substituting these conditions back into $g'(0) = 0$ yields

$$\left|F_{x'}h + F_{y'}m + F\delta t_1\right|_{t_1} + \left|G_t\delta t_1 + G_s\delta t_2 + G_x\delta x_1 + G_y\delta y_2\right|_{(t_1,t_2)}$$

$$+ \left|H_{x'}h + H_{y'}m + H\delta t_2\right|_{t_2} + \left|J_t\delta t_1 + J_s\delta t_2 + J_x\delta x_1 + J_y\delta y_2\right|_{(t_1,t_2)}$$

$$= 0$$

One cannot generally expect these terms to cancel out. Thus, necessary conditions for this equation to hold is that

$$\left|F_{x'}h + F_{y'}m + F\delta t_1\right|_{t_1} + \left|G_t\delta t_1 + G_s\delta t_2 + G_x\delta x_1 + G_y\delta y_2\right|_{(t_1,t_2)} = 0$$

$$\left|H_{x'}h + H_{y'}m + H\delta t_2\right|_{t_2} + \left|J_t\delta t_1 + J_s\delta t_2 + J_x\delta x_1 + J_y\delta y_2\right|_{(t_1,t_2)} = 0$$

The arbitrary functions $h(t)$ and $m(t)$ are still appearing in these equations. To eliminate these, make the following approximations:

$$h(t_1) \approx \delta x_1 - x^{*'}(t_1)\delta t_1$$

$$m(t_2) \approx \delta y_2 - y^{*'}(t_2)\delta t_2$$

Using these approximations (and the fact that δt_1 and δt_2 are free), the sought transversality conditions are obtained:

$$F - F_{x'}x^{*\prime} + G_t|_{(t_1,t_2)} = 0 \quad \text{at} \quad t = t_1$$

$$H - H_{y'}y^{*\prime} + J_t|_{(t_1,t_2)} = 0 \quad \text{at} \quad t = t_2$$

The four boundary conditions, the four Euler equations and the two transversality conditions can now be used to find the optimal pair of terminal times (t_1, t_2). In the problem at hand, it is apparent that the objective does not include the state variables themselves, but only their derivatives $\overline{z}'(t)$ and $\underline{z}'(t)$. This has implications for the Euler equations. Namely, they reduce to

$$\frac{dF_{x'}}{dt} = \frac{dF_{y'}}{dt} = \frac{dH_{x'}}{dt} = \frac{dH_{y'}}{dt} = 0$$

This in turn implies that the following derivatives are constants:

$$F_{x'} = k_1, \quad F_{y'} = k_2, \quad H_{x'} = k_3, \quad H_{y'} = k_4 \tag{3.1}$$

4. The Optimal Contract in the Quadratic Two-Type Case

I now analyse the optimal contract when the agent can have one of two efficiency types $\theta \in \{\underline{\theta}, \overline{\theta}\}$, and $q = \Pr(\theta = \overline{\theta})$. By the assumption that the principal can observe only whether the task is completed or not, the assumption that the completion depends on cumulative effort only, and by the assumption that the principal can perfectly commit to any contract written at the beginning of the game, this dynamic problem is essentially transformed into a static one. It is therefore possible to use the methodology of static adverse selection. Specifically, I can invoke the revelation principle. The principal can thus, without loss of generality, offer a menu of contracts $\{(\overline{w}, \overline{T}); (\underline{w}, \underline{T})\}$, one designed specifically for each type.

The principal wishes to maximise his utility subject to the following constraints:

$$\overline{w} \geq \frac{1}{2} \int_0^{\overline{T}} e^{-rt} \overline{x}(t)^2 dt \tag{4.1}$$

$$\underline{w} \geq \frac{1}{2} \int_0^{\underline{T}} e^{-rt} \underline{x}(t)^2 dt \tag{4.2}$$

$$\overline{w} - \frac{1}{2} \int_0^{\overline{T}} e^{-rt} \overline{x}(t)^2 dt \geq \underline{w} - \frac{1}{2} \int_0^{\underline{T}} e^{-rt} \overline{x}(t)^2 dt \tag{4.3}$$

$$\underline{w} - \frac{1}{2} \int_0^{\underline{T}} e^{-rt} \underline{x}(t)^2 dt \geq \overline{w} - \frac{1}{2} \int_0^{\overline{T}} e^{-rt} \underline{x}(t)^2 dt \tag{4.4}$$

where $\overline{w} = w(\overline{\theta})$, $\underline{w} = w(\underline{\theta})$, $\overline{x}(t) = x(\overline{\theta}, t)$ and $\underline{x}(t) = x(\underline{\theta}, t)$. The first two constraints, (4.1) and (4.2) are the usual participation (or individual rationality) constraints, securing that the agent enjoys non-negative utility when choosing the contract designed for him. The constraints (4.3) and (4.4) are incentive compatibility constraints, ensuring that no type has an incentive to mimic being of the other type.

As usual, the individual rationality for the inefficient type will be binding in the optimal contract. Thus

$$\underline{w} = \frac{1}{2}\int_0^{\underline{T}} e^{-rt}\underline{x}(t)^2 dt \tag{4.5}$$

Also, the incentive compatibility constraint of the efficient type will be binding. This yields

$$\overline{w} - \frac{1}{2}\int_0^{\overline{T}} e^{-rt}\overline{x}(t)^2 dt = \underline{w} - \frac{1}{2}\int_0^{\underline{T}} e^{-rt}\overline{x}(t)^2 dt \Leftrightarrow$$

$$\begin{aligned}\overline{w} &= \underline{w} - \frac{1}{2}\int_0^{\underline{T}} e^{-rt}\overline{x}(t)^2 dt + \frac{1}{2}\int_0^{\overline{T}} e^{-rt}\overline{x}(t)^2 dt \\ &= \frac{1}{2}\left[\int_0^{\overline{T}} e^{-rt}\underline{x}(t)^2 dt + \int_{\overline{T}}^{\underline{T}} e^{-rt}\left[\underline{x}(t)^2 - \overline{x}(t)^2\right] dt\right]\end{aligned} \tag{4.6}$$

Note that the individual rationality constraint of the inefficient type (4.2) and incentive compatibility of the efficient type (4.3) imply individual rationality of the efficient type (4.1). To see this, substitute the inefficient type's individual rationality constraint for $\underline{w}$ in the efficient type's incentive compatibility constraint and rearrange:

$$\overline{w} \geq \frac{1}{2}\int_0^{\overline{T}} e^{-rt}\overline{x}(t)^2 dt + \frac{1}{2}\int_0^{\underline{T}} e^{-rt}\left[\underline{x}(t)^2 - \overline{x}(t)^2\right] dt$$

The first term on the right hand side, is just the efficient type's individual rationality constraint. Now I just need to show that the second term is positive. From appendix B.1, I have explicit expressions for the effort minimising paths

$\underline{x}(t)$ and $\overline{x}(t)$. Using these, positivity of the second term reduces to the following constraint:

$$(\overline{\theta}e^{rT}-1)^2 \geq (\underline{\theta}e^{rT}-1)^2$$

which is trivially satisfied since $\overline{\theta} > \underline{\theta}$. I can thus proceed solving the principal's programme ignoring the efficient types participation constraint.

The principal's maximisation problem is:

$$\max_{(\overline{w},\underline{w},\overline{T},\underline{T})} \{q[e^{-r\overline{T}}S-\overline{w}]+(1-q)[e^{-r\underline{T}}S-\underline{w}]\}$$

with $(\overline{w},\underline{w})$ given by (4.5) and (4.6).

After substituting for the binding constraints, the objective can be rewritten as

$$qe^{-r\overline{T}}S+(1-q)e^{-r\underline{T}}S-\int_0^{\underline{T}}\frac{e^{-rt}}{2}\left[\underline{x}(t)^2-q\overline{x}(t)^2\right]dt-\int_0^{\overline{T}}\frac{qe^{-rt}}{2}\overline{x}(t)^2dt$$

Recall that $z'(t)=\theta^{-1}x(t)$, and thus $\underline{z}'(t)=\underline{\theta}^{-1}\underline{x}(t)$ and $\overline{z}'(t)=\overline{\theta}^{-1}\overline{x}(t)$. I now translate the problem at hand to the format of the proof in appendix B.3. Let

$t_1 = \overline{T}$, $t_2 = \underline{T}$ and

$$x(t) = \overline{\theta}^{-1}\overline{z}'(t)$$

$$y(t) = \underline{\theta}^{-1}\underline{z}'(t)$$

$$F = \frac{qe^{-rt}}{2}\left[\overline{\theta}^{-1}\overline{z}'(t)\right]^2$$

$$H = \frac{e^{-rt}}{2}\left[\left[\underline{\theta}^{-1}\underline{z}'(t)\right]^2 - q\left[\overline{\theta}^{-1}\overline{z}'(t)\right]^2\right]$$

$$G = qe^{-r\overline{T}}S$$

$$J = (1-q)e^{-r\underline{T}}S$$

The derivatives (3.1) are as follows:

$$F_{x'} = qe^{-rt}\overline{\theta}^{-1}\overline{z}'(t) = k_1 \Leftrightarrow \overline{z}'(t) = q^{-1}e^{rt}\overline{\theta}k_1$$

$$F_{y'} = 0$$

$$H_{x'} = -qe^{-rt}\overline{\theta}^{-1}\overline{z}'(t) = k_3 \Leftrightarrow \overline{z}'(t) = -q^{-1}e^{rt}\overline{\theta}k_3$$

$$H_{y'} = -e^{-rt}\underline{\theta}^{-1}\underline{z}'(t) = k_4 \Leftrightarrow \underline{z}'(t) = -e^{rt}\underline{\theta}k_4$$

It turns out that the second equation does not contribute, and that the first and third are identical. Integrating these expressions and using the boundary conditions $\overline{z}'(0) = \underline{z}'(0) = 0$, yields

$$\overline{z}(t) = \frac{\overline{\theta} k_1}{qr}\left(e^{rt} - 1\right) \Rightarrow \tag{4.7}$$

$$\overline{z}'(\overline{T}) = \frac{\overline{\theta} k_1}{q} e^{r\overline{T}} \tag{4.8}$$

$$\underline{z}(t) = \frac{\underline{\theta} k_4}{r}(1 - e^{rt}) \Rightarrow$$

$$\underline{z}'(\underline{T}) = -\underline{\theta} k_4 e^{r\underline{T}} \tag{4.9}$$

Now turn to the transversality conditions. First I find the partials of the gains functions:

$$G_t'(\overline{T}) = -rqe^{-r\overline{T}} S$$

$$J_t'(\underline{T}) = -r(1-q)e^{-r\underline{T}} S$$

The two transversality conditions thus become

$$qe^{-r\overline{T}}\left[\overline{\theta}^{-1}\overline{z}'(\overline{T})\right]^2 = 2rqe^{-r\overline{T}} S$$

$$e^{-r\underline{T}}\left[\left(\underline{\theta}^{-1}\underline{z}'(\underline{T})\right)^2 - q\left(\overline{\theta}^{-1}\overline{z}'(\underline{T})\right)^2\right] = 2r(1-q)e^{-r\underline{T}} S$$

Rearranging, this yields

$$\overline{z}'(\overline{T}) = \overline{\theta}\sqrt{2rS} \tag{4.10}$$

$$\underline{z}'(\underline{T}) = \underline{\theta}\sqrt{(1-q)2rS + \frac{q}{\overline{\theta}^2}\left[\overline{z}'(\underline{T})\right]^2}$$

From (4.7), it follows that

$$\overline{z}'(\underline{T}) = \frac{\overline{\theta}k_1}{q}e^{r\underline{T}}$$

This in turn implies that

$$\underline{z}'(\underline{T}) = \underline{\theta}\sqrt{(1-q)2rS + \frac{k_1^2}{q}e^{2r\underline{T}}} \tag{4.11}$$

Equalising (4.8) and (4.10) yields:

$$\overline{\theta}k_1 e^{r\overline{T}} = q\overline{\theta}\sqrt{2rS} \Rightarrow$$

$$\overline{T} = r^{-1}\log\left[\frac{\overline{\theta}\sqrt{2rS}}{\overline{\theta}\sqrt{2rS} - rL}\right]$$

where it has been used that

$$k_1 = \frac{qrL}{\overline{\theta}(e^{r\overline{T}} - 1)}$$

Naturally, one needs to impose $\overline{\theta}\sqrt{2rS} > rL$. I need to find the exact value of k_1, and I do so by evaluating it at $\overline{T}$. I get

$$k_1 = q\sqrt{2rS} - \frac{qrL}{\overline{\theta}} \Rightarrow$$

$$k_1^2 = q^2\left[\sqrt{2rS} - \frac{rL}{\overline{\theta}}\right]^2$$

Also I have that

$$k_4 = \frac{rL}{\underline{\theta}(1 - e^{r\underline{T}})}$$

Substituting this in expression (4.9), it follows that

$$\underline{z}'(\underline{T}) = -\frac{rLe^{r\underline{T}}}{1 - e^{r\underline{T}}}$$

Equalising this with (4.11) after having eliminated k_1, rearranging and defining

$$\rho = e^{r\underline{T}}$$

$$\alpha = 2r(1-q)S$$

$$\beta = q\left(\sqrt{2rS} - \frac{rL}{\overline{\theta}}\right)^2$$

$$\gamma = -\left(\frac{rL}{\underline{\theta}}\right)^2$$

yields

$$\underline{T} = r^{-1}\log(\rho)$$

where ρ is the relevant solution to the polynomial

$$\alpha(\rho - 1)^2 + \beta\rho^2(\rho - 1)^2 + \gamma\rho^2 = 0$$

This is a fourth degree equation in ρ, which in principle can be solved. However, the complicated form of the solutions does not facility comparison with the first-best. For specific parameterisations of the model, it can be confirmed that the optimal contract indeed distorts the deadline for the inefficient type.

5. On the Timing of Transfers

Since I have assumed the existence of perfect capital markets guaranteeing a riskless return r on investments, both the agent and the principal are indifferent as to the timing of payments. To see this, consider an amount of disutility a incurred at time $t = t'$. The principal has to pay the agent an amount equal to $ae^{-rt'}$. With our assumption that the agent is paid on delivery, i.e. at $t = T > t'$, the principal has to pay an amount $b > a$ such that $ae^{-rt'} = be^{-rT}$. Thus $b = ae^{r(T-t')}$. The agent is indifferent between these two transfers. To see that the principal is also indifferent, note that if, instead of paying $ae^{-rt'}$ at $t = t'$, he invests the amount in the capital markets, he will, at $t = T$, receive the amount $d = (ae^{-rt'})e^{r(T-t')} = ae^{r(T-2t')}$. In date $t = t'$ units, this is worth $de^{-r(T-t')} = e^{-r(T-t')}ae^{r(T-2t')} = ae^{-rt'}$. Therefore no generality is lost by adopting the convention that the transfer to the agent is executed on the date of project completion.

Bibliography

[1] ANDERSON, S. P., A. DE PALMA AND J.-F. THISSE, 1992. *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge Massachusetts.

[2] ANDRADE, A. AND E. STAFFORD, 1999. Investigating the Economic Role of Mergers, *mimeo.*

[3] ANDRADE, A., M. MITCHELL AND E. STAFFORD, 2001. New Evidence and Perspectives on Mergers, *Journal of Economic Perspectives*, 15(2), 103-120.

[4] ARVAN, L. AND A. P. N. LEITE, 1990. Cost Overruns in Long Term Projects, *International Journal of Industrial Organization,* 8, 443-467.

[5] ATHEY, S., 2002. Monotone Comparative Statics Under Uncertainty, *Quarterly Journal of Economics*, CXVII (1), 187-223.

[6] BANERJEE, A., (1992). A Simple Model of Herd Behavior, *Quarterly Journal of Economics*, 107, 797-818.

[7] BARKOULAS, J. T., C. F. BAUM AND A. CHAKRABORTY, 2001. Waves and Persistence in Merger and Acquisition Activity, *Economics Letters*, 70, 237-243.

[8] BIKHCHANDANI, S., D. HIRSHLEIFER AND I. WELCH, 1992. A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades, *Journal of Political Economy*, 100, 992-1026.

[9] BITTLINGMAYER, G., 1996. Mergers as a Form of Investment, *Kyklos*, 49(2), 127-153.

[10] BLAIR, M. M. AND M. A. SCHARY, 1993. Industry-Level Pressures to Restructure, in M. M. Blair (ed.), The Deal Decade: What Takeovers and Leveraged Buyouts Mean for Corporate Governance, *The Brookings Institution*, Washington D. C.

[11] Boeckem, S., 2001. Mergermania or "After You", A Theory of Merger Waves with Evidence from the Oil Industry, *mimeo.*

[12] Bulow, J. I., J. D. Geanakoplos and P. D. Klemperer, 1985. Multimarket Oligopoly: Strategic Substitutes and Complements, *Journal of Political Economy*, 93(3), 488-511.

[13] Bulow, J., M. Huang and P. Klemperer, 1999. Toeholds and Takeovers, *Journal of Political Economy*, 107(3), 427-454.

[14] Bulow, J. and P. Klemperer, 1994. Rational Frenzies and Crashes, *Journal of Political Economy*, 102(1), 1-23.

[15] Burdzy, K., D. M. Frankel and A. Pauzner, 2001. Fast Equilibrium Selection by Rational Players Living in a Changing World, *Econometrica*, 69(1), 163-89.

[16] Burkart, M., 1995. Initial Shareholding and Overbidding in Takeover Contests, *Journal of Finance*, 5, 1491-1515.

[17] Cabral, I., 2000. Merger Waves and Informational Cascades, *mimeo.*

[18] Carlsson, H. and E. van Damme, 1993. Global Games and Equilibrium Selection, *Econometrica*, 61(5), 989-1018.

[19] Chamley, C., 1999. Coordinating Regime Switches, *Quarterly Journal of Economics*, 114(3), 869-905.

[20] Chamley, C.,2002. Dynamic Speculative Attacks, *mimeo.*

[21] Chappell, Jr., H. W. and D. C. Cheng, 1984. Firm's Acquisition Decisions and Tobin's q Ratio, *Journal of Economics and Business*, 36, 29-42.

[22] Chow, Y. S., H. Robbins and D. Siegmund, 1991. The Theory of Optimal Stopping, *Dover Publications, Inc.*, New York.

[23] Cramton, P., 1998. Auctions and Takeovers, in the New Palgrave Dictionary of Law and Economics, Peter Neuman (ed.), London: *MacMillan Press*, 1, 122-125.

[24] Cramton, P. and A. Schwartz, 1991. Using Auction Theory to Inform Takeover Regulation, *Journal of Law, Economics, and Organization*, 7(1), 27-53.

[25] CUKIERMAN, A. AND Z. F. SHIFFER, 1976. Contracting for Optimal Delivery Time in Long-Term Projects, *Bell Journal of Economics,* 7(1),132-149.

[26] DIXIT, A. K. AND R. S. PINDYCK, 1994. Investment Under Uncertainty, *Princeton University Press.*

[27] ENGELBRECHT-WIGGANS, R. AND R. J. WEBER, 1979. An Example of a Multi-Object Auction Game, *Management Science,* 25(12), 1272-1277.

[28] FARIA, A. L., 2002. Mergers and the Market for Organizational Capital, *mimeo.*

[29] FARRELL, J. AND C. SHAPIRO, 1990. Horizontal Mergers: An Equilibrium Analysis, *American Economic Review, 81(1),* 107-26.

[30] FAULI-OLLER, R., 2000. Takeover Waves, *Journal of Economics and Management Strategy,* 9(2), 189-210.

[31] FAULI-OLLER, R. AND M. MOTTA, 1996. Managerial Incentives for Takeovers, *Journal of Economics and Management Strategy, 5(4),* 497-514.

[32] FISHMAN, M. J., 1988. A Theory of Preemptive Takeover Bidding, *Rand Journal of Economics,* 19(1), 88-101.

[33] FRANKEL, D. AND A. PAUZNER, 2000. Resolving Indeterminacy in Dynamic Settings: The Role of Shocks, *Quarterly Journal of Economics,* 115, 285-304.

[34] FRANKEL, D., S. MORRIS AND A. PAUZNER, 2000. Equilibrium Selection in Global Games with Strategic Complementarities, *forthcoming in Journal of Economic Theory.*

[35] FRIDOLFSSON S.-O.AND J. STENNEK, 2000. Why Mergers Reduce Profits, and Raise Share Prices - A Theory of Preemptive Mergers, *mimeo.*

[36] FUDENBERG, D. AND J. TIROLE, 1985. Preemption and Rent Equalization in the Adoption of New Technology, *Review of Economic Studies,* 52(3), 383-401.

[37] GALE, D., 1995. Dynamic Coordination Games, *Economic Theory,* 5, 1-18.

[38] GOLBE, D. L. AND L. J. WHITE, 1988. A Time-Series Analysis of Mergers and Acquisitions in the U.S. Economy, in *Corporate Takeovers: Causes and Consequences,* A. J. Auerbach (ed.), *University of Chicago Press/NBER.*

[39] GOLDSTEIN, I. AND A. PAUZNER, 2000. Demand Deposit Contracts and the Probability of Bank Runs, *mimeo.*

[40] GORDY, M. B., 1998. Computationally Convenient Distributional Assumptions for Common-Value Auctions, *Computational Economics*, 12(1), 61-78.

[41] GORT, M., 1969. An Economic Disturbance Theory of Mergers, *Quarterly Journal of Economics*, 83(4), 624-642.

[42] GORTON, G., M. KAHL AND R. ROSEN, 1999. Eat or Be Eaten: A Theory of Mergers and Merger Waves, *mimeo.*

[43] GOWRISANKARAN, G., 1999. A Dynamic Model of Endogenous Horizontal Mergers, *Rand Journal of Economics, 30(1)*, 56-83.

[44] GOWRISANKARAN, G. AND T. J. HOLMES, 2000. Do Mergers Lead to Monopoly in the Lomg Run? Results from the Dominant Firm Model, *mimeo.*

[45] GRENADIER, S., 1996. The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets, *Journal of Finance*, LI(5), 1653-1679.

[46] GRENADIER, S. (ed.), 2000. Game Choices: The Intersection of Real Options and Game Theory, *Risk Books/Risk Publications.*

[47] GROSSMAN, G. M. AND C. SHAPIRO, 1986. Optimal Dynamic R&D Programs, *Rand Journal of Economics,* 17(4), 581-593.

[48] GUESNERIE, R. AND J.-J. LAFFONT, 1984. A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-managed Firm, *Journal of Public Economics*, 25, 329-369.

[49] GUL, F. AND R. LUNDHOLM (1995): Endogenous Timing and the Clustering of Agent's Decisions, *Journal of Political Economy*, 103(5), 1039-1066.

[50] HARRIS, E. G., 1994. Why One Firm is the Target and the Other the Bidder in Single-Bidder, Synergistic Takeovers, *Journal of Business*, 67(2), 263-280.

[51] HARSTAD, R. M., 1990. Alternative Common-Value Auction Procedures: Revenue Comparisons with Free Entry, *Journal of Political Economy*, 98(2), 421-429.

[52] HAUSCH, D. B. AND L. LI, 1993. A Common Value Auction Model with Endogenous Entry and Information Acquisition, *Economic Theory*, 3, 315-334.

[53] HOLMSTROM, B., 1979. Moral Hazard and Observability, *Bell Journal of Economics*, 10(1), 74-91.

[54] HVIID, M. AND C. PRENDERGAST, 1993. Merger Failure and Merger Profitability, *Journal of Industrial Economics*, XLI(4), 371-386.

[55] JENKINS, A. M., J. D. NAUMANN AND J. C. WETHERBE, 1984. Empirical Investigation of Systems Development Practices and Results, *Information and Management*, 7.

[56] JOVANOVIC, B. AND P. L. ROUSSEAU, 2002. The Q-Theory of Mergers, *mimeo.*

[57] KAMIEN, M. I. AND N. L. SCHWARTZ, 1971. Expenditure Patterns for Risky R and D Projects, *Journal of Applied Probability*, 8, 60-73.

[58] KAMIEN, M. I. AND N. L. SCHWARTZ, 1991. Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, 2nd Edition, *North-Holland.*

[59] KAMIEN, M. I. AND I. ZANG, 1990. The Limits of Monopolization Through Acquisition, *Quarterly Journal of Economics*, 105(2), 465-499.

[60] KAMIEN, M. I. AND I. ZANG, 1993. Monopolization by Sequential Acquisition, *Journal of Law, Economics & Organization*, 9(2), 205-229.

[61] KLEMPERER, P., 1997. Auctions with Almost Common Values: The "Wallet Game" and its Applications, *European Economic Review*, 42(3-5), 757-69.

[62] KREMER, M., 1993. The O-Ring Theory of Economic Development, *Quarterly Journal of Economics*, 108(3), 551-575.

[63] LAFFONT, J.-J. AND J. TIROLE, 1998. A Theory of Incentives in Procurement and Regulation, *MIT Press*, Cambridge, Massachusetts.

[64] LAMBRECHT, B. M., 2001. The Timing and Terms of Takeovers under Uncertainty: A Real Options Approach, *mimeo.*

[65] LANG, L. H. P., R. M. STULZ AND R. A. WALKLING, 1989. Managerial Performance, Tobin's Q, and the Gains from Successful Tender Offers, *Journal of Financial Economics*, 24(1), 137-154.

[66] LEVIN, J., 2001. A Note on Global Equilibrium Selection in Overlapping Generations Games, *mimeo.*

[67] LEWIS, T. R., 1986. Reputation and Contractual Performance in Long-Term Projects, *Rand Journal of Economics*, 17(2), 141-157.

[68] LUCAS, JR., R. E., 1971. Optimal Management of a Research and Development Project, *Management Science*, 17(11), Theory Series, 679-697.

[69] MANSFIELD, E., J. SCHNEE AND S. WAGNER, 1971. Overruns and Errors in Estimating Development Cost, Time, and Outcome, reprinted in Innovation, Technology and the Economy, The Selected Essays of Edward Mansfield, Vol. I, 1995, *Edward Elgar Publishing Limited.*

[70] MARSHALL, A. W. AND W. H. MECKLING, 1962. Predictability of the Costs, Time, and Success of Development, in The Rate and Direction of Inventive Activity: Economic and Social Factors, *Princeton University Press.*

[71] MATHEWSON, F. AND N. QUIGLEY, 1998. To Merge or Not To Merge: Is That Really the Question?, *paper presented at the Policy Conference on Financial Institution Reform*, C. D. Howe Institute.

[72] MCDONALD, R. AND D. SIEGEL, 1986. The Value of Waiting to Invest, *Quarterly Journal of Economics*, 101, 707-728.

[73] MILGROM, P. R. AND R. J. WEBER, 1982. A Theory of Auctions and Competitive Bidding, *Econometrica*, 50(5), 1089-1122.

[74] MITCHELL, M. L. AND J. H. MULHERIN, 1996. The Impact of Industry Shocks on Takeover and Restructuring Activity, *Journal of Financial Economics*, 41, 193-229.

[75] MOEL, A. AND P. TUFANO, 2002. When Are Real Options Exercised? An Empirical Study of Mine Closings, *Review of Financial Studies*, 15(1), 35-64.

[76] MOLNAR, J., 2000. Preemptive Horizontal Mergers: Theory and Evidence, *mimeo.*

[77] MOERCK, R., A. SHLEIFER AND R. W. VISHNY, 1988. Characteristics of Targets of Hostile and Friendly Takeovers, *in Corporate Takeovers: Causes and Consequences*, A. J. Auerbach (ed.), University of Chicago Press/NBER.

[78] MORRIS, S. AND H. S. SHIN, 1998. Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks, *American Economic Review*, 88(3), 587-597.

[79] MORRIS, S. AND H. S. SHIN, 2002. Global Games: Theory and Applications, *forthcoming in Advances in Economics and Econometrics*, the Eighth World Congress (M. Dewatripont, L. Hansen and S. Turnovsky, eds.), Cambridge University Press.

[80] MUELLER, D. C., 1969. A Theory of Conglomerate Mergers, *Quarterly Journal of Economics*, 83, 634-659.

[81] MUSGROVE, P., 1985. Why Everything Takes 2.71828 . . . Times as Long as Expected, *American Economic Review*, 75(1), 250-252.

[82] NELSON, R., 1959. Merger Movements in American Industry, 1895-1956, *Princeton University Press/NBER.*

[83] NILSSEN, T. AND L. SORGARD, 1998. Sequential Horizontal Mergers, *European Economic Review*, 42, 1667-1682.

[84] NITSCHE, R., 2001. Takeover Bids in Merger Races, *mimeo.*

[85] NORBAK, P.-J. AND L. PERSSON, 2001. Investment Liberalization - Who Benefits from Cross-Border Mergers, *mimeo.*

[86] NORBAK, P.-J. AND L. PERSSON, 2002. Cross-Border Acquisitions and Greenfield Entry, *mimeo.*

[87] ODONOGHUE, T. AND M. RABIN, 1999. Incentives for Procrastinators, *Quarterly Journal of Economics*, 114(3), 769-816.

[88] OYAMA, D., 2001. Booms and Slumps in a Game of Sequential Investment with Changing Fundamentals, *mimeo.*

[89] PECK, M. J. AND F. M. SCHERER, 1962. The Weapons Acquisition Process, *Harvard University Press.*

[90] PERRY, M. K. AND R.H. PORTER, 1985. Oligopoly and the Incentive for Horizontal Merger, *American Economic Review*, 75, 219-227.

[91] RHODES-KROPF, M. AND S. VISWANATHAN, 2002. Market Valuation and Merger Waves, *mimeo.*

[92] ROCHET, J.-C. AND X. VIVES, 2001. Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?, *mimeo.*

[93] RODRIGUES, V., 2002. Merger Waves, *mimeo.*

[94] ROLL, R., 1986. The Hubris Hypothesis of Corporate Takeovers, *Journal of Business*, 59(2), Part 1, 197-216.

[95] SALANT, S. W., S. SWITZER AND R.J. REYNOLDS, 1983. Losses From Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium, *Quarterly Journal of Economics,* XCVIII(2), 185-199.

[96] SALONER, G., 1987. Predation, Mergers, and Incomplete Information, *Rand Journal of Economics,* 18(2), 165-186.

[97] SERVAES, H., 1991. Tobin's Q and the Gains from Takeovers, *Journal of Finance*, 46(1), 409-419.

[98] SHELTON, L. M., 2000. Merger Market Dynamics: Insights into the Behavior of Target and Bidder Firms, *Journal of Economic Behavior & Organization*, 41, 363-383.

[99] SHLEIFER, A. AND R. W. VISHNY, 2001. Stock Market Driven Acquisitions, *NBER Working Paper 8439.*

[100] SHUGHART, W. F. II AND R. D. TOLLISON, 1984. The Random Character of Merger Activity, *Rand Journal of Economics*, 15(4), 500-509.

[101] SMITH, K. W. AND A. J. TRIANTIS, 1995. The Value of Options in Strategic Acquisitions, in L. Trigeorgis (ed.), Real Options in Capital Investment; Models, Strategies, and Applications, *Praeger*, Westport Connecticut.

[102] SOBEL, J., 1992. How to Count to One Thousand, *The Economic Journal,* 102(410), 1-8.

[103] SPEAR, S. E. AND S. SRIVASTAVA, 1987. On Repeated Moral Hazard with Discounting, *Review of Economic Studies*, 54(4), 599-617.

[104] STOKEY, N. L. AND R. E. LUCAS JR., 1989. Recursive Methods in Economic Dynamics, *Harvard University Press*, Cambridge, Massachusetts.

[105] TOWN, R. J., 1993. Merger Waves and the Structure of Mergers and Acquisitions Time Series, *Journal of Applied Econometrics*, 7, 83-100.

[106] TEALL, J. L., 1992. Merger Activity and Managerial Compensation, *Review of Financial Economics*, 2(1), 55-67.

[107] WESTON, J. F., K. S. CHUNG AND S. E. HOAG, 1990. Mergers, Restructuring, and Corporate Control, *Prentice Hall, Inc.,* Englewood Cliffs, New Jersey.

[108] WILSON, R. B., 1993. Nonlinear Pricing, *Oxford University Press.*