STRATEGIC MERGER WAVES: A THEORY OF MUSICAL CHAIRS

BY FLAVIO TOXVAERD

April 2003

Abstract. This paper proposes an explanation of merger waves based on a dynamic preemption game. A set of acquirers compete over time for scarce targets. At each point in time, an acquirer can either postpone a takeover attempt, or raid immediately. By postponing the takeover attempt, an acquirer may gain from more favorable future market conditions, but runs the risk of being preempted by rivals. I first consider a complete information model and show that the above tradeoff leads to a continuum of subgame perfect equilibria in monotone strategies that are strictly Pareto ranked. All these equilibria share the feature that all acquirers rush simultaneously in merger waves. The model is then extended to a dynamic global game by introducing slightly noisy private information about merger profitability. This game is shown to have a unique Markov perfect Bayesian equilibrium in monotone strategies. The comparative dynamics predictions of the model are related to stylized facts.

Keywords: Merger waves, preemption, dynamic global games, real options games.

JEL Classification: C73, D92, G34, L13.

∗Address for correspondence: Universidade Nova de Lisboa, Faculdade de Economia, Campus de Campolide, P-1099-032 Lisboa, Portugal.

†I am pleased to acknowledge support and guidance from Luis Cabral and insightful suggestions from Chryssi Giannitsarou and Hyun Song Shin. I also benefited from comments and conversations with Pascal Courty, David Frankel, Stephen Morris, Marco Ottaviani, Ady Pauzner and Xavier Vives. I thank seminar participants at Universitat Autonoma de Barcelona, Universitat Pompeu Fabra, Universidade Nova de Lisboa, Cambridge University, London School of Economics, London Business School, Hebrew University of Jerusalem, Tel Aviv University, University of Copenhagen, IUI Stockholm, Chicago GSB, Yale SOM, NYU Stern, HEC Montreal and participants at the 2002 meetings of the European Economic Association in Venice, the 2002 Econometric Society European Meetings in Venice and the 2002 EARIE meetings in Madrid. This research has been supported by a Marie Curie Fellowship of the European Union programme Marie Curie Training Site under contract number HPMT-CT-2000-00172.
1. **Introduction**

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

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 aim of this paper is to build a theory that can explain why mergers happen in waves, incorporating both dependence on exogenous factors, such as aggregate activity, and strategic interdependence between firms’ decisions.

Merger wave theories can be categorized 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 organization literature has focused almost exclusively on strategic interaction through the product market. Unfortunately, standard oligopoly theory is not entirely satisfactory in explaining merger waves. In fact, the simplest model may predict 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 \(n\) for some \(n\). In other words, there are industry sizes under which the incentive for any two firms to merge decreases as the number of mergers in the industry increases.\(^3\) 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 into merger waves that consist of vertical or conglomerate mergers.

At the other end, non-strategic theories of merger waves emphasize the effects of exogenous factors such as deregulation, globalization or the introduction of new technologies. In this context, merger waves are characterized 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

---

\(^1\)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).

\(^3\)Fauli-Oller (2000) and Rodrigues (2002) build simple four-firm oligopoly models in which some equilibria can be characterized as merger waves. However, they predict that merger waves should coincide with declining markets, which is in contrast to the evidence.
in merger activity may follow from technological shocks as physical 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. Last, Faria (2002) presents a model in which mergers serve as a vehicle for the transfer of managerial skills (intangible assets). In this setting, interactions between market participants or strategic considerations are absent.

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. In the present work, 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 stochastic process, but merger waves occur as a result of strategic interaction. A strategic merger wave in the current setting will be interpreted as a situation in which the exogenous economic conditions prompting a firm to seek a merger vary discontinuously with the merger activity of other firms.

The model builds on three simple features. 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 capitalization. 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 assets. Such assets could be access to restricted (geographical) markets, existing customer 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. Viewing an acquisition as an irreversible investment (or at least partially irreversible in the short run) is plausible, and the merger decision can then reasonably be viewed as the problem of optimally exercising a real option. Delaying a merger may allow firms to look for the best fit; or it may be that the returns from the merger are realized 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.

The third feature of the present model is that competition for targets is imperfect. Specifically, what is ruled out is competition à la Bertrand with homogeneous products, where all

---

4There is an extensive finance literature that distinguishes acquirers from targets along values of Tobin’s q. See e.g. Servaes (1991) and references therein.

5The view of mergers as investments can be traced back to Mueller (1969). See also Bittlingmayer (1996).

6For a thorough exposition of real options, see e.g. Dixit and Pindyck (1994).
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 existence of white knights and the fact that target management sometimes accept offers that are not the most attractive, suggest that this is a plausible modeling assumption.

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 realizations 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. The timing of the merger wave can thus be predicted.

The tradeoff 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.

The musical chairs metaphor is routinely used in the business press to describe the environment and conditions leading to merger waves. Commenting on an impending merger wave in the international brewing industry, one analyst observed that

International brewers are playing a game of musical chairs. Investors want beer companies to have exposure to both developed and developing markets to balance steady cash flow and new growth opportunities. In a relatively fragmented industry, one of the best ways to do this is through acquisitions - and there are only so many targets. [Furthermore], the easy pickings are dwindling in number.7

Last, a commentator described an expected merger wave in the international industry for legal services by stating that

One of the images accountants like to use when describing the strategic thinking of law firms is that of an enormous and slightly lascivious game of musical chairs.

The music is almost over and all the big Australian law firms are circling the room, trailing their coats in the direction of a handful of global law firms and the Big Five professional services firms. If the Australians are lucky, the music might last just long enough for them to attract a merger partner [...]. But if they delay, all the international merger candidates will be snapped up by the lucky few [...].

As these quotes make clear, analysts and industry participants themselves view the tradeoff between preemption (strategic considerations) and exogenous economic factors (non-strategic considerations) as crucial for the decision on if and when to seek a merger. This fact lends strong support to the present modeling approach.

Methodologically, the present model has features in common with several strands of literature. 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 (2002) and Frankel, Morris and Pauzner (2000). 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, Morris and Shin, 1999, 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. Also, I consider a setting with simultaneous moves.

There is a large literature on rational herding and informational cascades. In such models, information is inferred from observing others’ actions. In contrast, players in the present model use available information to infer others’ actions. My model is close in spirit to Bulow and Klemperer’s (1994) model of rational frenzies, but differs 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 decentralized 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 characterized as a stampede (“sudden mass movement”) rather than as a frenzy (“brief delirium that is almost insanity”).

Although constructed with merger waves in mind, the presented methodology is kept at a fairly general level and may thus have applications to other models of mass movements, such as models of speculative attacks, crashes etc. A discussion of possible applications is offered in the concluding section.

---

The basic setup is described in Section 2, which also exposes some key properties of the model. In Section 3, the complete information version of the model is analyzed and merger wave equilibria are characterized. Section 4 extends the model to a dynamic global game by introducing incomplete information, and shows that when information is very precise, the timing of the merger wave can be uniquely determined. Comparative analysis is performed in Section 5, and a simple example is contained in Section 6. Section 7 offers concluding remarks and a discussion of some methodological issues. Finally, proofs of most lemmata are relegated to the Appendix.

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.\(^{10}\) 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. Denote by \( a^i_t \in A^i_t = \{0, 1\} \) player \( i \)'s action at time \( t \), with \( a^i_t = 0 \) denoting waiting and \( a^i_t = 1 \) denoting raiding. 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 \( B_t : Z_t \times X_t \times \mathbb{R} \to \mathbb{R} \) with von Neumann-Morgenstern expected payoff \( R(z_t, x_t, \theta_t) \) and remains inactive in all future periods.\(^{11}\) The single dimensional variable \( \theta_t \in \mathbb{R} \) represents some economic fundamental that influences merger profitability.

The expected payoff \( R(z_t, x_t, \theta_t) \) from participating in the allocation game is called the raiding value, and should 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

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

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

\[
\Delta(z_t, x_t, \theta_t) = W(z_t, x_t, \theta_t) - R(z_t, x_t, \theta_t)
\]

In the event that \( \Delta(z_t, x_t, \theta_t) < 0 \), raiding is the dominant strategy, while waiting is dominant for \( \Delta(z_t, x_t, \theta_t) > 0 \).

Next, make the following assumptions:

**A1** \( y_0 > x_0 \).

**A2** The raiding value \( R(z_t, x_t, \theta_t) \) is bounded, continuous in all arguments, strictly increasing in \( \theta_t \), weakly decreasing in \( z_t \) and weakly increasing in \( x_t \) with \( R(z_t, 0, \theta_t) = 0 \).

\(^{10}\)Note that the analysis of neither the complete nor the incomplete information games depends on the continuum player assumption. Furthermore, the results of the complete information game do not depend on symmetry between the acquirers. It is still an open question whether the results derived in the incomplete information setting generalize to acquirers with non-informational asymmetries.

\(^{11}\)It is implicitly assumed that there exists a unique equilibrium in the allocation game.
A3 The process \( \{ \theta_t \}_{t=0}^{\infty} \) is first-order Markov such that \( \theta_t | \theta_{t-1} \sim G \), with density function \( g \), and for \( \theta_{t-1} > \theta'_{t-1} \), \( G(\theta_t | \theta'_{t-1}) > G(\theta_t | \theta_{t-1}) \).

A4 The raiding value and the stochastic process \( \{ \theta_t \}_{t=0}^{\infty} \) are such that for all \( z_{t+1}, x_{t+1} \)

\[
R(z_t, x_t, \theta_t) - \delta E[R(z_{t+1}, x_{t+1}, \theta_{t+1}) | \theta_t]
\]

is strictly increasing in \( \theta_t \).

A5 If \( a^i_t = 1 \) then \( A^i_s = \emptyset \) for all \( s > t \).

Assumption A1 captures the notion that targets are scarce. Assumption A2 ensures that the raiding value is bounded and well behaved, and further that expected payoffs from raiding are decreasing in the intensity of competition (measured either as an increase in the measure of competitors \( z_t \) or as the absolute scarcity of remaining targets \( x_t \)). Last, the value of raiding is assumed to be increasing in the economic fundamental, such that higher realizations increase the benefits of merging, controlling for the level of competition. Assumption A3 states that there is persistence in the evolution of the economic fundamental, such that a higher realization of \( \theta_t \) today shifts the distribution of future realizations according to first-order stochastic dominance. Assumption A4 is a joint condition on the raiding value and the stochastic process which ensures that the raiding value increases at a higher rate than the waiting value. Last, assumption A5 simply states that a merger is irreversible.

Note that assumptions A2-A5, i.e. irreversibility and persistent shocks to merger profitability yield a value of delay, a standard insight of the real options literature. The requirement that the raiding value be strictly increasing in the economic fundamental imposes restrictions on the allocation mechanism \( B_t \). Specifically, it amounts to the assumption of market imperfections such that, controlling for the level of competition, the expected benefit from merging is non-negative for the raider. It is implicitly assumed that no raider gets rationed in the allocation game as long as \( z_t < x_t \). Thus, if an acquirer decides to raid, he will either merge with a target and leave the game, or be rationed, in which case the game is over. Thus, it follows that the laws of motion for the endogenous state variables are given by

\[
\begin{align*}
x_t & \equiv \max \left\{ x_0 - \sum_{r=0}^{t-1} z_r, 0 \right\} \\
y_t & \equiv \max \left\{ y_0 - \sum_{r=0}^{t-1} z_r, y_0 - x_0 \right\}
\end{align*}
\]

This couple 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.

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

---

12 The assumption that targets are scarce is also made by Klemperer (1997), Bulow, Huang and Klemperer (1999), Rhodes-Kropf and Viswanathan (2002) and is virtually standard in the corporate finance literature on takeover auctions. Conceivably, this assumption could be dispensed with by also dispensing with the assumption of unit demand.

13 This assumption does not rule out the possibility that the target obtains most of the gains from trade. What it does rule out is perfect rent equalization.
Definition 1. (Merger Triggers)

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

**Strategic Trigger**: $\tilde{\theta}(z_t^*) \equiv \inf \{\theta_t : \Delta(z_t, x_t, \theta_t) \leq 0\}$

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

That is, $\theta$ is the lowest value of $\theta_t$ at which the raiding value is non-negative. The term Marshallian trigger is borrowed from the real options literature. Next, the strategic trigger $\tilde{\theta}(z_t^*)$ is the lowest value of $\theta_t$ such that raiding in the current period dominates waiting, given a sequence of current and future measures of raiders. Last, the first-best trigger $\theta(z_t^*)$ is the lowest value of $\theta_t$ such that, even in the absence of competitive pressure, delaying a takeover one period further is not optimal.\(^{14}\) The first-best trigger $\theta(z_t^*)$ is simply the strategic trigger $\tilde{\theta}(z_t^*)$ evaluated at $z_t^*$ with $z_s = 0, s \geq t$. Clearly, $\theta(z_t^*) \in [\bar{\theta}, \tilde{\theta}(z_t^*)]$ and $\tilde{\theta}(z_t^*) \rightarrow \theta(z_t^*)$ as $x_t \rightarrow y_t$.

The following lemmata simply state that the above merger triggers exist and are unique:

**Lemma 2.** (Single Crossing in $\theta_t$) For any sequence $z_t$ there exists a unique $\tilde{\theta}(z_t^*) \in [\bar{\theta}, \infty[$ such that $\Delta(z_t, x_t, \theta_t) > 0$ for $\theta_t < \tilde{\theta}(z_t^*)$, $\Delta(z_t, x_t, \theta_t) = 0$ for $\theta_t = \tilde{\theta}(z_t^*)$ and $\Delta(z_t, x_t, \theta_t) < 0$ for $\theta_t > \tilde{\theta}(z_t^*)$. Furthermore, $\tilde{\theta}(z_t^*)$ is weakly increasing in $x_t$ and weakly decreasing in $z_t$ and $z_t^*$.

**Proof.** See Appendix A \(\blacksquare\)

**Lemma 3.** (Single Crossing in $z_t$) For any sequence $z_t$ 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 $\theta_t$.\(^{15}\)

**Proof.** See Appendix A \(\blacksquare\)

These results have a straightforward interpretation. First, given any level of future and present competition, the expected value for an acquirer contemplating whether to raid or wait increases in the economic fundamental. In fact, both the value of raiding and waiting increase. But as the economic fundamental 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 $\theta_t$ 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, $\theta_t$ is taken to be any variable determined at aggregate level, such as technological progress, interest rates, a stock market index or other macroeconomic variable.

\(^{14}\) Although not explicit in the adopted notation, $\tilde{\theta}$ at time $t$ may depend on the state variable $x_t$, in which case a higher stock of targets would increase the first-best trigger.

\(^{15}\) If $[\bar{\theta}, \tilde{\theta}(z_t^{t+1})] = \emptyset$ for all $z_t^{t+1}$ then for $\theta_t \geq \tilde{\theta}$, $\Delta(z_t, x_t, \theta_t) \leq 0$ for all $z_t \in [0, y_t]$. It is thus assumed throughout that $[0, \tilde{\theta}(z_t^{t+1})] \neq \emptyset$ for some $z_t^{t+1}$. 
Remark 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 realization, 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, $\theta_t$ can parameterize a flow which the winner of the allocation process obtains at a further sunk non-stochastic 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 $\theta_t$ is a sufficient statistic for the value of raiding, controlling for the 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 $\zeta$ describes 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 such a setup, assumptions A3-A4 can be dispensed with. For example, shocks that are identically and independently distributed over time can be handled and is actually easier to analyze. For an excellent treatment of optimal stopping problems of this type see Chow, Robbins and Siegmund (1991).

3. The Complete Information Game

For the complete information game it is assumed that both past and current realizations of the economic fundamental $\theta_t$ are common knowledge. Let $h_t = (\theta_0, ..., \theta_{t-1}; z_0, ..., z_{t-1})$ 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 4. A monotone Markovian strategy for the complete information game is a mapping $a : \mathbb{R}^{2t} \times \mathbb{R} \to \mathbb{R}$ such that $(h_t, \theta_t) \mapsto a(h_t, \theta_t) \equiv k_t$.

Thus, for a history $h_t$ and current fundamental $\theta_t$, 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 $a_t$ will thus be

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

where 1 stands for raid and 0 stands for wait and $a_{k_t}(\theta_t)$ denotes an indicator function. With this definition in place, the following result can be stated:
Proposition 5. (Merger Waves) For any sequence \( z^{t+1} \) and history \( h_t \in H_t \), any cutoff \( k_t \in [\bar{\theta}(z^{t+1}), \hat{\theta}(z^{t+1})] \) constitutes an equilibrium strategy. Furthermore, there exist no equilibria in asymmetric strategies.

Proof. The first part of the proposition follows immediately from the definitions and Lemma 3. To see the second part, fix a sequence of future cutoff strategies \( \{k_i\}_{i=t+1}^{\infty} \) (and thus \( z^{t+1} \)) and consider period \( t \). Let a measure \( \mu^i \), \( i = a,b \) use strategies with cutoff \( k_i^a \) with \( k_i^a < k_i^b \) and \( \mu^a + \mu^b = 1 \). Recall that a cutoff \( k_t \) is only an equilibrium strategy if it is optimal for any realization of the economic fundamental \( \theta_t \). For \( \theta_t \in [\bar{\theta}, k_t^a] \) all acquirers wait and the asymmetric strategies can coexist in equilibrium. Similarly, for \( \theta_t \in [k_t^b, \hat{\theta}(z^{t+1})] \) 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 realization in this range prompts a measure \( \mu^a \) to raid and a measure \( \mu^b \) to wait. If \( \mu^a \geq z_t^a \), the cutoff \( k_t^b \) cannot be an equilibrium strategy. Similarly, if \( \mu^b < z_t^a \), the cutoff \( k_t^a \) cannot be an equilibrium strategy. Thus equilibria in asymmetric strategies do not exist.

Thus there is a continuum of equilibrium cutoffs in each period. These are strictly Pareto ranked, with higher cutoffs dominating lower ones. It follows that even the Pareto efficient subgame perfect equilibrium is strictly inefficient, as it induces firms to merge at a suboptimally low level of the economic fundamental.\(^{16}\)

Corollary 6. (Indeterminacy) A merger wave may be triggered in any period \( t \) where \( \theta_t \in [\bar{\theta}(x_t, 0), \theta(z^{t+1})] \). Furthermore, a merger wave may be triggered no earlier than \( t = \inf \{t : \theta_t \geq \bar{\theta}\} \) and no later than \( \bar{t} = \inf \{t : \theta_t \geq \hat{\theta}(z^t)\} \).

Proof. Follows immediately from Proposition 5.

As usual in this type of models, there is no clear way in which to determine the equilibrium outcome. At this stage, the only thing that can be said is that \( \bar{t} \) is the earliest time and \( \bar{t} \) the very latest time at which a rush can occur. Actually, this claim is weaker than necessary, for there is indeed no equilibrium with acquirers rushing at time \( \bar{t} \). Postponing a takeover this long is inconsistent with the pressure caused by any reasonable threat of preemption. An immediate result is thus that in all equilibria, a merger wave will happen in finite time with probability one. In other words, there is no equilibrium in which all acquirers postpone their takeovers indefinitely.

It has been argued (see 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

\(^{16}\)The game also has an equilibrium in mixed strategies. To characterize this equilibrium, recall that by definition, a raid of measure \( z_t^a \) leaves an acquirer exactly indifferent between raiding and waiting. This yields a straightforward expression for the mixed strategy equilibrium. Assume that all acquirers at time \( t \) raid with probability \( p_t \). The measure of acquirers is thus simply given by \( p_t y_t \). This in turn implies that, in the mixed strategy equilibrium, acquirers raid with probability \( p_t^* = z_t^a / y_t \). It follows trivially that when \( z_t^a \) is undefined, the mixed strategy equilibrium is degenerate, i.e. the unique pure strategy equilibrium is to raid.
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 are the assumptions of complete information and common knowledge. These assumptions imply that acquirers are perfectly able to predict rivals’ behavior 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 realization of the fundamental variable $\theta_t$ is no longer common knowledge. Recall that $G(\theta_t | \theta_{t-1})$ denotes the distribution of $\theta_t$, conditional on past information, and denote by $g$ the corresponding probability density function. The next step is to specify the information technology, which is characterized by the following assumptions.

A6 At time $t$, acquirer $i$ receives signal $s_{it} = \theta_t + \sigma \varepsilon_{it}$ with $\varepsilon_{it} \sim F$ (density $f$) identically and independently distributed over $t$ and across $i$.

A7 For $a > b$, $f(a - \theta)/f(b - \theta)$ is increasing in $\theta$.

In A6, the scalar $\sigma > 0$ measures the precision of the signal. Assumption A7 simply states that the distribution of noise $F$ satisfies the monotone likelihood ratio property. That is, an increase in any signal shifts the distribution of other signals in the sense of first-order stochastic dominance. This fact is important, as signals are not only used to estimate $\theta_t$, but also to make inferences about other acquirers’ signals. The monotone likelihood ratio property is implied by the assumption that the variables $\theta_t$ and $\{s_{it}\}_{i \in Y_t}$ are affiliated.17

In this setting, define the following:

Definition 7. A monotone Markovian strategy for the incomplete information game is a mapping $a : \mathbb{R}^{2t} \times \mathbb{R} \to \mathbb{R}$ such that $(h_t, s_t) \mapsto a(h_t, s_t) \equiv k_t$.

Given a strategy $k_t$ the chosen action $a_t$ will thus be given by

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

where again 1 denotes raid and 0 denotes wait and $a_{k_t}(s_t)$ denotes an indicator function. Thus, the choice variable is the cutoff level $k_t$. Given these strategies, the measure of raiders for given cutoff $k_t$ is determined by the distribution of signals. But given $\theta_t$, the distribution of a signal $s$ is given by $F\left(\frac{s - \theta_t}{\sigma}\right)$. One can use this to express the measure of raiders as

$$z_t = y_t \left[1 - F\left(\frac{k_t - \theta_t}{\sigma}\right)\right]$$

---

17 For a formal definition and further results on affiliated variables, see Milgrom and Weber (1982).
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 realize that history only serves to the extent that it yields a prior belief \( G(\theta_t | \theta_{t-1} ) \) on \( \theta_t \). 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 \( \Gamma_t \) with a simplified (associated) static game \( \Gamma^*_t \), 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 \( \Gamma_t \) 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 realization of the state and where history is informative about the distribution of this period’s realization) 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 truncated game converge as the horizon recedes so that equilibria of \( \Gamma(T) \), as \( T \to \infty \), are equilibria of the infinite horizon game \( \Gamma(\infty) \).

Using the terminology of Morris and Shin (2002), the problem to be solved in each period satisfies action single crossing (established in Lemma 3), state monotonicity (established in Lemma 2), limit dominance (follows from the existence of \( \bar{\theta} \) and \( \bar{\theta}(z') \)) and a monotone likelihood ratio property of the signal distribution (Assumption A7). Under these and some continuity conditions, Morris and Shin (2002) show that there exists a unique Bayesian equilibrium in monotone strategies. The key to the proof is to realize 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 a unique 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 \( k_t \).

Before continuing with the analysis, I will state a straightforward lemma that determines the posterior distribution and density of \( \theta \) given a signal \( s \), that will be useful in the sequel.

**Lemma 8.** (Postiors) The posterior density \( f_{\theta|s} \) and distribution \( F_{\theta|s} \) of \( \theta \) given some
signal $s$ are given by

$$f_{\theta|s}(\theta|s) = \frac{g(\theta) f \left( \frac{s-\theta}{\sigma} \right)}{\int_{-\infty}^{\infty} g(\theta) f \left( \frac{s-\theta}{\sigma} \right) \, d\theta}$$  \quad (1)$$

$$F_{\theta|s}(\theta|s) = \frac{\int_{0}^{\theta} g(\theta) f \left( \frac{s-\theta}{\sigma} \right) \, d\theta}{\int_{-\infty}^{\infty} g(\theta) f \left( \frac{s-\theta}{\sigma} \right) \, d\theta} = \frac{\int_{s}^{\infty} g(s-\sigma u) f(u) \, du}{\int_{-\infty}^{\infty} g(s-\sigma u) f(u) \, du}$$  \quad (2)

**Proof.** See Appendix B

To formally state the proposition of the existence of a unique perfect Bayesian equilibrium in monotone strategies, assume for now 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 a simplified associated game $\Gamma^*(T)$, where it is assumed that the received signal is a sufficient statistic of the state and $\theta$ is drawn from a uniform distribution on the real line. Throughout, an asterisk will denote quantities pertaining to the associated game. Although the prior of $\theta$ is an improper distribution (has infinite probability mass), it is possible to apply Lemma 8 by normalizing the prior density to one, i.e. $g(\theta) = 1$. The density of the posterior is then given by $f_{\theta|s}(\theta|s) = \sigma^{-1} f \left( \frac{s-\theta}{\sigma} \right)$ and the distribution by $F_{\theta|s}(\theta|s)$.

Let $\Delta^*_k(s, k)$ denote the expected payoff gain to “waiting” with respect to the posterior after having received signal $s$, and believing that all other players use strategies with cutoffs $k$. This is given by

$$\Delta^*_k(s, k) \equiv E_{\theta|s} \left[ \Delta \left( y \left[ 1 - F \left( \frac{k - \theta}{\sigma} \right) \right], x, s \right) \right]$$

$$= \int_{-\infty}^{\infty} \Delta \left( y \left[ 1 - F \left( \frac{k - \theta}{\sigma} \right) \right], x, s \right) \sigma^{-1} f \left( \frac{s - \theta}{\sigma} \right) \, d\theta$$  \quad (3)

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 $k_t$. By Lemma 8, this is given by

$$\Delta_\sigma(s_t, k_t) \equiv E_{\theta|s} \left[ \Delta \left( y_t \left[ 1 - F \left( \frac{k_t - \theta}{\sigma} \right) \right], x_t, \theta_t \right) \right]$$

$$= \int_{-\infty}^{\infty} \Delta \left( y_t \left[ 1 - F \left( \frac{k_t - \theta}{\sigma} \right) \right], x_t, \theta_t \right) g(\theta) f \left( \frac{s_t - \theta}{\sigma} \right) \, d\theta$$  \quad (4)

The differences between (3) and (4) are twofold. First, in (3), the signal $s$ replaces the economic fundamental $\theta$. 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 $\Delta^*_k(s, k)$ 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 the uniqueness result can be stated:

\footnote{Since the associated game is essentially static, time subscripts are omitted for ease of notation. This should cause no confusion.}
Lemma 9. (Uniqueness in Associated Game) For any history $h_t \in H_t$, there exists a unique cutoff signal $s_t^*$ in the associated static game $\Gamma_t$ such that: $\Delta^*_\sigma(s_t^*, s_t^*) = 0$, $\Delta^*_\sigma(s_t, s_t^*) > 0$ for $s_t < s_t^*$ and $\Delta^*_\sigma(s_t, s_t^*) < 0$ for $s_t > s_t^*$.

Proof. See Appendix C ■

Lemma 10. (Limit Uniqueness in Underlying Game) For any history $h_t \in H_t$, as $\sigma \to 0$, $\Delta_\sigma(s_t, s_t - \sigma \xi) \to \Delta^*_\sigma(s_t, s_t - \sigma \xi)$ uniformly.

Proof. See Appendix D ■

Lemma 9 states that any static associated game has a unique Bayesian equilibrium. Lemma 10 shows that when private information is very precise, any finite horizon version of the underlying game becomes arbitrarily close to a sequence of simplified static associated games. In other words, as $\sigma \to 0$, the period $t$ expected relative payoff function in the underlying game, converges uniformly to the relative payoff function of some simplified static game which has a unique Bayesian equilibrium in monotone strategies. Having established these results, the following proposition can be stated:

Proposition 11. (Uniqueness in Infinite Horizon Game) As $\sigma \to 0$, for each sequence of realizations $\{\theta_t\}_{t=0}^\infty$, there exists a unique sequence of cutoff signals $\{s_t^*\}_{t=0}^\infty$ such that for any history $h_t \in H_t$: $\Delta_\sigma(s_t^*, s_t^*) = 0$, $\Delta_\sigma(s_t, s_t^*) > 0$ for $s_t < s_t^*$ and $\Delta_\sigma(s_t, s_t^*) < 0$ 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 a 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 $s_t^*(T) \to s_t^*(\infty)$ as $T \to \infty$, where $s_t^*(T)$ is the equilibrium cutoff in period $t$ in the game truncated after period $T$ and $s_t^*(\infty)$ is the equilibrium cutoff at time $t$ in the infinite horizon game.$^{19}$ ■

$^{19}$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 $T$. 
Corollary 12. (Unique Timing of Merger Wave) The merger wave will be triggered at time $t^* \equiv \inf \{ t : \theta_t \geq \theta^*_t \}$.

Proof. Follows immediately from Proposition 11.

Recall that under complete information, there was a continuum of realizations of the economic fundamental which could constitute equilibrium cutoffs. The striking result of Proposition 11 is that under incomplete but very precise information, there exists only one equilibrium cutoff $\theta^*$ (which is of course a function of the state variables). In $2 \times 2$ 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, $\theta^*$ 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.

Figure 1: Triggers and sample path of $\theta_t$. 

![Figure 1: Triggers and sample path of $\theta_t$.](Image)
5. Comparative Analysis

Although the model presented here is quite general, it still allows for an identification of factors influencing the timing of mergers.\footnote{Of course $\theta^* = \theta$ cannot be excluded, in which case the equilibrium is degenerate.} 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 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 realizations closer to the strategic trigger shifts the risk-dominant trigger downwards. This is because the expected time until the equilibrium trigger is hit decreases. The model thus predicts that the higher the growth rate of the economic fundamental, the earlier the wave occurs. This is consistent with the stylized fact that M&A activity is particularly intense in industries with fast technological progress, such as those of information technology, telecoms and pharmaceuticals.

Third, a comparative dynamics result that follows from the real options literature is that the first-best trigger $\theta(z^t)$ is increasing in the volatility of the process $\{\theta_t\}_{t=0}^\infty$ \footnote{Getting sharp comparative statics results within the current setup is significantly complicated by the assumption of persistence. Under iid shocks however, results can be obtained with relative ease, see e.g. Sargent (1987). This is done by first considering how the first-best trigger $\theta(z^t)$ varies with the parameters of interest, and then arguing that the level of the unique equilibrium trigger is a monotone function of the first best trigger.} (see e.g. Dixit and Pindyck, 1994). To see this, consider a mean preserving spread of the distribution $G(\theta_t|\theta_{t-1})$ and recall that the waiting value is non-negative (a zero payoff can be achieved by always waiting). A mean preserving spread will increase the probability mass in the tails of the distribution. But while making high values of the economic fundamental more likely, and thereby high raiding values more likely, the downside is truncated by the fact that raiding will never occur for realizations of the economic fundamental below $\theta$. Thus, an increase in volatility increases the opportunity cost of an immediate raid, thereby making it optimal to postpone it further. Similarly, with very low volatility there is no incentive to postpone a merger (supposing of course that $\theta_t \geq \theta$ so that the raiders break even).

Fourth, the exact way in which the expected payoffs $R(z_t, x_t, \theta_t)$ in the allocation game $B_t$ 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 behavior, and there seems to be a consensus that most takeover contests, especially those that seek to replace inefficient management, resemble common-value auctions.\footnote{See discussion in Klemperer (1997), Cramton and Schwartz (1991) and Cramton (1998).} 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.
Strategic Merger Waves

Figure 2: Simulated merger activity.

...and acquirers. Leaving the exact nature of the allocation game unspecified and further assuming that the raiding value $R(z_t, x_t, \theta_t)$ is differentiable, consider the effects of increases in $R_\theta(z_t, x_t, \theta_t)$, $R_x(z_t, x_t, \theta_t)$ and $R_z(z_t, x_t, \theta_t)$, where subscripts denote partial derivatives. First, recall that the options value of postponing a merger derives entirely from the variability in the raiding value brought about by the stochastic evolution of the economic fundamental. Thus, the higher $R_\theta(z_t, x_t, \theta_t)$, the more an acquirer can benefit from high realizations of.

---

23 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. Although some work on auctions explicitly considers cases where there is a continuum of bidders, it may not be the most palatable framework to work in. One may assume that the allocation mechanism $B_t$ is initiated by a matching round in which each raider is assigned to a specific target. A finite player common values auction may be conducted at a second stage. Alternatively, recall that the continuum player assumption is not substantial and simply made for convenience.
the economic fundamental. In a sense, this result is equivalent to increased volatility of the economic fundamental, and so tends to delay the merger wave. Conversely, as $R_\theta(z_t, x_t, \theta_t)$ approaches zero, the merger wave will happen at the first time where $\theta_t \geq \theta$. An increase in $R_x(z_t, x_t, \theta_t)$ corresponds to an increase in the competitive pressure brought about by the scarcity of targets. As the scarcity is non-decreasing over time, a higher absolute value of $R_x(z_t, x_t, \theta_t)$ erodes the options value of delay, thus bringing forward the merger wave. Last, consider and increase in $R_z(z_t, x_t, \theta_t)$. Intuitively, this also corresponds to a worsening of the raiders’ terms of trade vis-à-vis the targets. Again, such an increase will tend to hasten the merger wave.

While the effects of increased volatility were discussed above, there is a second and somewhat subtler effect of increasing the level of uncertainty. Recall that knowledge of the process $\{\theta_t\}_{t=0}^\infty$ has two uses, namely for forecasting the future evolution and for generating a prior distribution $G(\theta_t | \theta_{t-1})$. The latter is influenced by the volatility of the stochastic process since it 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 $\theta_t | \theta_{t-1} \sim N(\theta_{t-1} + \mu_\theta, \sigma^2_\theta)$ and $s_{it} = \theta_t + \epsilon_{it}$ with $\epsilon_{it} \sim N(\mu_{\epsilon}, \sigma^2_{\epsilon})$. Thus, in a very volatile environment, uniqueness of a perfect Bayesian equilibrium in monotone strategies 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 a small measure of acquirers raid in a given period, the unique merger trigger in the proceeding period decreases, in turn making it more likely that some acquirers will receive signals above the trigger. In this way, the wave may gain momentum and ultimately result in a rush to merge. Figure 2 illustrates simulated merger activity and a sample path of the economic 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 realizations of the fundamental, there is no merger activity. As the fundamental increases, merger activity picks up. Note that towards the end, merger activity becomes insensitive to increases in the economic fundamental. This is simply because by then, there are not many remaining targets.

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. Mitchell and Mulherin (1996) cite their findings as support for the hypothesis that the shocks causing merger waves are industry-specific. My model’s predictions are consistent with this view. However, the previous discussion shows that industry-specific shocks need not be the sole

---

Note however, that although Gort (1969) emphasizes 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.
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

...[the evidence] suggests a formal model of [merger] activity as a function of a set of macroeconomic and industry-specific conditions [...]. [Mergers are] 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.

As mentioned above, it is a stylized 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 consistent with the predictions of 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 ends when all targets have been acquired.

Next, because the measure of raiders in any given period is endogenous and target scarcity is non-decreasing over time, the model is consistent with an interesting 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 of single-bidder and multiple-bidder takeovers. 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. Also, such an analysis could serve as a way of discriminating between the musical chairs theory and other models of merger waves.
6. An Example

In order to make the results less abstract, a very stylized but 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.$^{25}$

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 $\alpha$ and engages in bilateral bargaining. As one commentator on mergers in the US IT industry put it, 

[...]

in the context of musical chairs, those who do not carry enough weight to have their own chair, need to choose whose lap they are going to sit upon.$^{26}$

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.$^{27}$ Underlying these preferences lie non-modeled factors such as the tastes of target management, differences in corporate culture etc. Bidder $j$ seeks to maximize

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

where

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

is the probability that bidder $j$ wins the target and $\pi(\theta) = \delta^T V - c$. 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

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

$^{25}$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.


$^{27}$See e.g. Anderson, de Palma and Thisse (1992) for a thorough exposition.
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

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

Summing up, the raiding value is given by

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

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. Discussion

This paper set forth a theory to explain the occurrence of merger waves. Merger waves were derived as an equilibrium phenomenon, in a simple timing game. The model presented here has several satisfactory features. First, it derives from simple and intuitive assumptions, namely scarcity of targets, imperfect competition, uncertainty and irreversibility. Second, the analysis builds on quite general assumptions on payoff functions and the stochastic process determining the evolution of the economic fundamental. Third, the present model reconciles stylized facts with casual observation. Specifically, it encompasses both dependence of the merger decision on macroeconomic variables and strategic considerations. Fourth, it predicts patterns of M&A activity broadly consistent with what is in fact observed. Fifth, the model is fully dynamic. While interesting comparative statics results may be obtained from static models (see e.g. Rhodes-Kropf and Viswanathan, 2002), a fuller understanding of merger dynamics is obtained only through the study of full-blown dynamic models.

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. 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 does not exclude the possibility that target firms play a more 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.

It may be argued that a pure real options framework would predict similar patterns of M&A activity as the musical chairs theory (for a two-firm merger model in a real options framework see Lambrecht, 2001). At least two lines of defence against this argument are possible. First, as illustrated by the quoted statements in the Introduction, it is abundantly clear that strategic considerations often play an integral role in the merger decision. Thus a theory of merger waves ignoring this feature may be telling only part of the story. Second, on a more methodological level, merger wave equilibria in a pure real options framework
would be very sensitive to assumptions of target homogeneity/heterogeneity. 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 could be “cherry-picking” rather than merger waves. That is, for a given state of the economic fundamental, only the marginal firms would be desirable takeover targets and, without any competitive pressure from rival acquirers, there would be no pressure to rush in order to prevent preemption. Dixit and Pindyck (1994) discuss similar issues and warn against the serious problems involved in aggregation in real options settings.

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.

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) analyze 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 (formally, 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 decentralized setting. To my knowledge, only one other paper has dealt with sequential bidding games of this nature, namely Bulow and Klemperer’s (1994) 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.

Although the main contribution of this paper is the modeling of merger waves, it does offer some distinct methodological contributions to the global games literature which are worth pointing out. First, while the informational structure of the model is shared by Morris and Shin (1999), that model does not contain any intertemporal links through either payoff functions or through action spaces. Thus, their model is by construction a sequence of relatively independent static games which can be solved independently, while in the current
model this feature has to be derived. Second, the papers using the approach of Burdzy et al., 2000 rely explicitly on the fact that players do not move simultaneously, and it is thus not clear if that approach carries over to discrete time settings. This is because the heterogeneity needed for uniqueness in their setting is provided by the feature that players moving at different times are doing so under different payoff relevant conditions (i.e. under different realizations of the stochastic variable). Under complete information and simultaneous moves, such heterogeneity disappears and multiple equilibria reappear. The current analysis relies on informational heterogeneity which plays a similar role, but allows for moves to be simultaneous.

Last, the analysis of the paper shows that the global games analysis carries over naturally to dynamic settings which have 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. For example, dynamic games of speculative attacks where a central bank’s reserves are eroded over time could be accommodated within the presented framework.28 As mentioned earlier, the current model belongs to the class of quasisupermodular games. Giannitsarou and Toxvaerd (2002) show that within a class of dynamic supermodular games, similar techniques to the ones developed here can be applied to establish uniqueness of perfect Bayesian equilibria. Possible applications of the techniques developed there include dynamic search games, arms races, extraction of common property exhaustible resources and dynamic models of multisector economies with players switching between sectors.

Appendices

A. Proof of Lemmata 2 and 3
Consider the space of bounded functions \( \Omega \) on \( Y_t \times X_t \times \mathbb{R} \) and define the operator \( M : \Omega[Y_t \times X_t \times \mathbb{R}] \to \Omega[Y_t \times X_t \times \mathbb{R}] \) 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 \( \Omega[Y_t \times X_t \times \mathbb{R}] \) with the sup-norm. With this norm, the space \( \Omega \) is a Banach space. Let \( V(z_t, x_t, \theta_t) > \tilde{V}(z_t, x_t, \theta_t) \) for all \( (z_t, x_t, \theta_t) \). Then

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

28 I thank Hyun Song Shin for this suggestion.
Thus the mapping $M$ satisfies monotonicity. Next, let $a > 0$. Thus

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

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

$$\leq MV(z_t, x_t, \theta_t) + \delta a$$

and the mapping $M$ satisfies discounting. Therefore, by Blackwell’s sufficiency conditions, $M$ is a contraction mapping (with modulus $\delta$) on $\Omega$. 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 $(z_t, x_t, \theta_t)$ (see e.g. Stokey and Lucas with Prescott, 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 $\theta_t$, 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 $\theta_t$. Also, for $\theta_t > \theta$, $V(z_t, x_t, \theta_t)$ 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 $\theta_t$. 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 $\theta_t < \theta$ 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 $\theta_t \geq \theta$ and consider an increase in $\theta_t$. Both the value of raiding and that of waiting will increase. A simple argument shows that for sufficiently high $\theta_t$, the value of raiding overtakes that of waiting. Assume that for all $\theta_t$

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

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 $\delta \in [0, 1]$. Assumption A4, i.e. the assumption that

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

is strictly increasing in $\theta_t$, 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.29 In conclusion, for each sequence $z^t$ there exists a unique finite $\tilde{\theta}(z^t) \in [0, \infty]$ such that

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

Since $V(z_t, x_t, \theta_t)$ is weakly increasing in $x_t$, so is $\tilde{\theta}(z^t)$. Similarly, since $V(z_t, x_t, \theta_t)$ is weakly decreasing in $z_t$, $\tilde{\theta}(z^t)$ is weakly decreasing in $z_t$. This also holds for any future measure of

29To see this, define $\Lambda(z_t, x_t, \theta_t) = V(z_t, x_t, \theta_t) - R(z_t, x_t, \theta_t)$. Straightforward manipulation yields $\Lambda(z_t, x_t, \theta_t) = \max \{0, -R(z_t, x_t, \theta_t) + \delta E[\Lambda(z_{t+1}, x_{t+1}, \theta_{t+1})] + \delta E[\Lambda(z_t, x_t, \theta_t)] \}$. Under Assumption A3-A4, this defines a mapping from $\Omega$ into itself which is strictly increasing in $\theta_t$. Furthermore, this mapping is a contraction with a unique fixed point. Last, note that since $W(z_t, x_t, \theta_t) = \delta V(z_t, x_t, \theta_t)$, it follows that $\Lambda(z_t, x_t, \theta_t) = \max \{0, \Delta(z_t, x_t, \theta_t)\}$. 

raiders, and thus $\tilde{\theta}(z^t)$ is also weakly decreasing in $z^t$. The first-best trigger $\bar{\theta}(z^t)$ is just the value of $\tilde{\theta}(z^t)$ for the sequence $z^t$ with $z_s = 0$ for all $s \geq t$. Last, note that for all $\theta_t > \tilde{\theta}$ 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 $\theta_t$ and weakly increasing in $x_t$. 

**B. Proof of Lemma 8**

Given a signal $s$, the density of the posterior about $\theta$ is given by Bayes’ rule as

$$f_{\theta|s}(\theta | s) = \frac{g(\theta) f_s(s | \theta)}{f_s(s)}$$

where $f_s$ and $f_{s|\theta}$ are the densities of $s$ and $s|\theta$ respectively. Now note that, since $s$ is the sum of $\theta$ and $\sigma \varepsilon$, its density will be given by the convolution of the densities of their densities, i.e. $g$ and $f_{\sigma \varepsilon}$. Also, noting that $F_{\sigma \varepsilon}(\omega) = F(\omega / \sigma)$ (thus $f_{\sigma \varepsilon}(\omega) = \sigma^{-1} f(\omega / \sigma)$), the unconditional density of $s$ is given by

$$f_s(s) = \sigma^{-1} \int_{-\infty}^{\infty} g(\theta) f\left(\frac{s - \theta}{\sigma}\right) d\theta$$

Moreover, to find $f_{s|\theta}$ note that

$$F_{s|\theta}(\varsigma | \theta) = \Pr(s \leq \varsigma | \theta) = \Pr\left(\varepsilon \leq \frac{\varsigma - \theta}{\sigma} \bigg| \theta\right) = F\left(\frac{\varsigma - \theta}{\sigma}\right) \Rightarrow$$

$$f_{s|\theta}(\varsigma | \theta) = \frac{d}{ds} F_{s|\theta}(\varsigma | \theta) = \sigma^{-1} f\left(\frac{\varsigma - \theta}{\sigma}\right)$$

because $\theta$ and $\varepsilon$ are independent. Thus, substituting (6) and (8) in (5), the density of $\theta$ given $s$ is:

$$f_{\theta|s}(\theta | s) = \frac{g(\theta) f\left(\frac{s - \theta}{\sigma}\right)}{\int_{-\infty}^{\infty} g(\theta) f\left(\frac{s - \theta}{\sigma}\right) d\theta}$$

Therefore the posterior distribution is

$$F_{\theta|s}(\theta | s) = \int_{-\infty}^{\theta} f_{\theta|s}(\theta | s) d\theta$$

$$= \frac{\int_{-\infty}^{\theta} g(\theta) f\left(\frac{s - \theta}{\sigma}\right) d\theta}{\int_{-\infty}^{\infty} g(\theta) f\left(\frac{s - \theta}{\sigma}\right) d\theta}$$

The second part of equality (2) follows by performing the transformation $u = \sigma^{-1}(s - \theta)$.
C. Proof of Lemma 9

To prove Lemma 9, 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 \( s_t = s^*_t \). Second, it is shown that for lower signals waiting is optimal, while for higher signals raiding is optimal.

First rewrite \( \Delta^*_\sigma(s, k) \), by changing variables using \( z = y \left[ 1 - F \left( \frac{s - \theta}{\sigma} \right) \right] \):

\[
\Delta^*_\sigma(s, k) = \int_0^y \Delta(z, x, k) y^{-1} dz
\]

For \( k = s \),

\[
\Delta^*_\sigma(s, s) = \int_0^y \Delta(z, x, s) y^{-1} dz
\]

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 \( s^* \) that solves

\[
\int_0^y \Delta(z, x, s^*) y^{-1} dz = 0
\]

Thus that there is exactly one cutoff signal \( s^* \) 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 \( s > s^* \) and waits whenever \( s < s^* \). In order to do this, recall that the game displays action single crossing (follows from Lemma 3), state monotonicity (follows from Lemma 2) and that the noise distribution has the monotone likelihood ratio property (Assumption A7).

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

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

\[
= \int_{-\infty}^{\infty} \Delta \left( y \left[ 1 - F(-m) \right], x, s \right) f \left( \frac{s - k}{\sigma} - m \right) dm
\]

by changing variables so that \( m = \sigma^{-1}(\theta - k) \). Now rewrite the above expression as

\[
\Delta^*_\sigma(s, k) = \tilde{\Delta}(s, k, s')
\]

\[
\equiv \int_{-\infty}^{\infty} \gamma(m, s') \varphi(s, m) dm
\]

where

\[
\gamma(m, s') = \Delta \left( 1 - F(-m), x, s' \right)
\]

\[
\varphi(s, m) = f \left( \frac{s - k}{\sigma} - m \right)
\]
Because of the monotone likelihood ratio property, \( \tilde{\Delta}(., k, s') \) 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

\[
\tilde{\Delta}(s, k, s') < 0 \quad \text{if} \quad s > s^*(k, s')
\]

\[
\tilde{\Delta}(s, k, s') > 0 \quad \text{if} \quad s < s^*(k, s')
\]

By state monotonicity, \( \tilde{\Delta}(s, k, s') \) is strictly decreasing in \( s' \). Now let \( s > s' \) and suppose that

\[
\tilde{\Delta}(s, k, s) = 0
\]

It follows that

\[
\tilde{\Delta}(s', k, s) > \tilde{\Delta}(s', k, s) > \tilde{\Delta}(s, k, s) = 0
\]

where the first inequality comes from state monotonicity and the second comes from the action single crossing property. A symmetric argument holds for \( s < s' \). This implies that there exists a best response function \( \beta : \mathbb{R} \to \mathbb{R} \) such that

\[
\Delta^*_\sigma(s, k) < 0 \quad \text{if} \quad s > \beta(k)
\]

\[
\Delta^*_\sigma(s, k) = 0 \quad \text{if} \quad s = \beta(k)
\]

\[
\Delta^*_\sigma(s, k) > 0 \quad \text{if} \quad s < \beta(k)
\]

But there exists a unique \( s^* \) that solves

\[
\Delta^*_\sigma(s^*, s^*) = 0
\]

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

\[
a_k(s) = \begin{cases} 
1 & \text{if} \quad s > s^* \\
0 & \text{if} \quad s < s^*
\end{cases}
\]

This proves Lemma 9.

\[\blacksquare\]

D. Proof of Lemma 10

It was shown in Lemma 9 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.

Recall that \( \Delta_\sigma(s, k) \) is

\[
\Delta_\sigma(s, k) = E_{\theta|s} \left[ \Delta \left( y \left[ 1 - F \left( \frac{k_t - \theta}{\sigma} \right) \right], x_t, \theta_t \right) \right]
\]

\[
= \int_{-\infty}^{\infty} \Delta \left( y \left[ 1 - F \left( \frac{k_t - \theta}{\sigma} \right) \right], x_t, \theta \right) dF_{\theta|s}(\theta|s)
\]
where \( F_{\theta|s} (\theta|s) \) is the posterior distribution. To do a change of variables using \( z = y \left[ 1 - F \left( \frac{s - \theta}{\sigma} \right) \right] \), first note that
\[
\theta = k - \sigma F^{-1} \left( \frac{y - z}{y} \right)
\]
Next, let \( \Psi_{\sigma}(z; s, k) \) be the posterior distribution evaluated at this \( \theta \). From Lemma 8 it follows that
\[
\Psi_{\sigma}(z; s, k) = F_{\theta|s} \left( k - \sigma F^{-1} \left( \frac{y - z}{y} \right) \right) = \frac{\int_{-\infty}^{\infty} g(s - \sigma u) f(u) du}{\int_{-\infty}^{\infty} g(s - \sigma u) f(u) du}
\]
Therefore (9) becomes,
\[
\Delta_{\sigma}(s, k) = \int_{0}^{y} \Delta \left( z, x, k - \sigma F^{-1} \left( \frac{y - z}{y} \right) \right) d\Psi_{\sigma}(z; s, k)
\]
Recall from the proof of Lemma 9 that for the associated game
\[
\Delta^*_{\sigma}(s, k) = \int_{0}^{y} \Delta (z, x, k) y^{-1} dz = \int_{0}^{y} \Delta (z, x, s) d\Psi^*_{\sigma}(z; s, k)
\]
where \( \Psi^*_{\sigma}(z; s, k) = F_{\theta|s}^\sigma \left( k - \sigma F^{-1} \left( \frac{y - z}{y} \right) \right) = 1 - F \left( \frac{z - k}{\sigma} + F^{-1} \left( \frac{y - z}{y} \right) \right) \). Thus \( \Psi^*_{\sigma}(z; s, s) = \frac{y - z}{y} \), which is the distribution function of the uniform distribution on \([0, y]\).

Returning to \( \Psi_{\sigma}(z; s, k) \), note that for some small \( |\xi| \)
\[
\Psi_{\sigma}(z; s, s - \sigma \xi) = \frac{\int_{\xi + F^{-1}(\frac{y - z}{y})}^{\infty} g(s - \sigma u) f(u) du}{\int_{-\infty}^{\infty} g(s - \sigma u) f(u) du} \to 1 - F \left( \xi + F^{-1} \left( \frac{y - z}{y} \right) \right) = \Psi^*_{\sigma}(z; s, s - \sigma \xi)
\]
Therefore, \( \Delta_{\sigma}(s, s - \sigma \xi) \to \Delta^*_{\sigma}(s, s - \sigma \xi) \) continuously as \( \sigma \to 0 \).

What remains to be shown is that \( \Delta_{\sigma}(s, s - \sigma \xi) \to \Delta^*_{\sigma}(s, s - \sigma \xi) \) uniformly as \( \sigma \to 0 \).
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 \( s \) and \( \bar{s} \) such that for all \( k: \Delta_{\sigma}(s, k) > 0 \) for \( s < \bar{s} \) and \( \Delta_{\sigma}(s, k) < 0 \) for \( s > \bar{s} \). This follows from the existence of dominance regions \([-\infty, \bar{s}] \) and \([\bar{s}, \infty] \), where there is a unique optimal action. One can thus pick any pair \( s \) and \( \bar{s} \) such that \( \bar{s} < \bar{s} \) and \( \bar{s} > \bar{s} \), and restrict attention to the compact interval \( S \equiv [s, \bar{s}] \). Since \( S \) is compact and the second argument of the \( \Delta_{\sigma} \) function is continuous with respect to \( s \) (i.e., the function \( s - \sigma \xi \)), the set \( K \equiv [s - \sigma \xi, \bar{s} - \sigma \xi] \) is also compact. Hence \( \Delta_{\sigma}(s, k) \) takes values in a compact set. Next, define the sup-norm (or uniform convergence norm)
\[
\|\Delta\| = \sup_{s, k} \{|\Delta(s, k)|\}
\]
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
\[
\forall \varepsilon_1 > 0, \exists \delta_1 \mid |s - s'| < \delta_1 \Rightarrow |\Delta_{\sigma}(s, k) - \Delta_{\sigma}(s', k)| < \varepsilon_1 \forall k
\]
\[
\forall \varepsilon_2 > 0, \exists \delta_2 \mid |k - k'| < \delta_2 \Rightarrow |\Delta_{\sigma}(s, k) - \Delta_{\sigma}(s, k')| < \varepsilon_2 \forall s
\]
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
\[
|\Delta_\sigma(s, k) - \Delta_\sigma(s', k')| = |\Delta_\sigma(s, k) - \Delta_\sigma(s', k) + \Delta_\sigma(s', k) - \Delta_\sigma(s', k')| \\
\leq |\Delta_\sigma(s, k) - \Delta_\sigma(s', k)| + |\Delta_\sigma(s', k) - \Delta_\sigma(s', k')| \\
\leq \varepsilon_1 + \varepsilon_2 \equiv \varepsilon
\]
and continuity with respect to the Euclidean metric follows. Denoting by \(C(S \times K)\) the space of continuous functions on \(S \times K\), it follows that \(\Delta_\sigma(s, k) \in C(S \times K)\). But showing uniform convergence is equivalent to showing that as \(\sigma \to 0\),
\[
||\Delta_\sigma - \Delta_\sigma^*|| = \sup_{s, k} \{|\Delta_\sigma(s, k) - \Delta_\sigma^*(s, k)|\} \to 0
\]
with respect to the sup-norm. By substituting for the relevant functions and taking limits, the result follows ■

References


