# *Greed, Fear, and Rushes*

Andreas Park* University of Toronto

Lones Smith† University of Michigan

January 1, 2009

**Abstract**

We develop a unified and tractable theory of sudden mass movements using a continuum agent timing game. We assume that an underlying payoff-relevant fundamental "ripens", peaks at an optimal "harvest time", and then "rots". These payoffs are multiplied by a hill-shaped quantile rank reward that subsumes "greed" and "fear" — namely, greed for greater rewards that come from outlasting others, and the fear of pre-emption. In this setting, we study the symmetric Nash equilibria.

Three local timing games can occur: a slow war of attrition, a slow pre-emption game, and a sudden pre-emption game, or a "rush". Rushes always exist, and are late with greedy players and early with fearful players. We relate measures of fear and greed, the timing and size of rushes, and the entry rate before and after rushes. Our theory provides an integrated understanding of seemingly unrelated phenomena, showing how unraveling in matching markets, liquidity runs on companies and financial bubbles all are part of the same class of problems.

*JEL Classification:* C73, D81.

*Keywords:* Games of Timing, War of Attrition, Preemption Game, Unraveling.

*Email: `andreas.park@utoronto.ca`; web: http://www.chass.utoronto.ca/∼apark/.

†Email: `lones@umich.edu`; web: http://www.umich.edu/∼lones. Lones thanks the NSF for research support.

# 1 Introduction

Mass rushes are a common feature of many economic settings: the rush of students to find a suitable match in fraternities; the rush of young MDs to find hospitals for specialization-internships; the rush to trade in popping financial bubbles; the rush of bank depositors withdrawing from a bank; and the rush of white flight from a racially-tipping neighborhood. We develop a unified and tractable theory of these sudden mass movements that sheds new light on rushes that may appear disparate. The cited rushes come in many shapes and sizes. Some rushes occur at the most sensible moment, such as fleeing a building during a fire alarm, or a sinking ship when the captain gives the order. But many arise either inefficiently early, or inefficiently late. Some are proportionately quite large, and others quite small. Our theory unearths and explains a co-variation the size and the timing of the rushes.

To meaningfully discuss timing, we obviously must first endow time with economic content. We thus assume that an underlying payoff-relevant fundamental "ripens", peaks at an optimal "harvest time", and then "rots". Our theory explains why some rushes occur before the harvest time, and others after the harvest time. Matching-related rushes, for one, tend to occur inefficiently early: Fraternity rushes happen before the academic year has even started; hiring of federal judicial clerkships famously takes place long before students have completed their second year of law school, and law firms hire interns before students have started their degree (Roth and Xing, 1994). Bank runs, eg., occur inefficiently early — they start on unfounded rumors of an otherwise "healthy" bank's stability. Conversely, a sales-rush that pops a financial bubble occurs long after the harvest time, when fundamentals have peaked. For instance, the dotcom bubble burst on March 10, 2000, while fundamentals "peaked" in 1999.[1]

The strategic side of the model will subsume the two rubrics of "greed" and "fear". The driving force of early rushes is the fear of pre-emption. By contrast, late rushes are best characterized by greed, the lust for greater rewards that come from outlasting others. Of course, these speak to the two underlying timing games in economics. At the heart of our analysis are three types of local timing games: in a *war of attrition*, the passage of time is fundamentally harmful but strategically beneficial, and players slowly enter; the reverse incentives also drive gradual entry in a *slow-entry pre-emption game*; finally, a *rush* is a sudden pre-emption game. We give conditions on primitives that govern the form of timing game, whether rushes are early or late, and large or small. We also study how changes in the primitives affect behavior.

[1]NASDAQ firm profits were higher in 1999 than later years, and market-to-book values peaked in 1999. High market-to-book values indicate either high market values or low book values. See Pastor and Veronesi (2006), Figures 5 (book-to-market values) and 6 (profitability, measured by Return on Equity).

A novel dimension of our timing game is the continuum of players — it is a "large game". This is the standard formalization of the idea from partial equilibrium analysis that everyone is economically small, which well describes the settings we study. In our strategic setting, it means in particular that the equilibrium evolves deterministically. In our game, payoffs multiplicatively depend on the stopping time and on quantile rank. To best capture our economic settings, we assume that people prefer to be neither first nor last — specifically, quantile payoffs are hill-shaped, first rising and then falling.

Our next powerful simplification is to focus on the symmetric Nash equilibria in which players commit to strategies that depend solely on the payoff fundamental. Since everyone uses the same strategy, no player can expect a higher equilibrium payoff than anyone else. In other words, if some player attains a preferred quantile rank, then he sacrifices on the fundamental. In this sense, players formally pay for their quantile rank.

Greed and fear admit simple formalizations in our strategic structure. Players are *fearful* if the quantile rank payoff from pre-empting all other players exceeds the average quantile rank payoff; the difference of these payoffs measures *fear*. Players are *greedy* if the last quantile rank payoff exceeds the average quantile rank payoff; the difference of these payoffs measures *greed*. Loosely, an early rush occurs when players fear being left out, while greed for a better later quantile rank dominates in a late rush. We show that people never play a slow stopping game when fearful, nor an early rush when greedy.

A slow timing game obtains when the marginal costs and benefits of delay coincide. This yields a differential equation that describes the distribution of stopping times in a slow timing game. Quantile payoffs must fall when delay is exogenously beneficial and must rise when delay is exogenously costly — respectively, a pre-emption game and a war of attrition. Since quantile payoffs rise and then fall in the rank, quantile payoffs eventually start to fall as players stop. This clearly cannot coexist with a worsening fundamental payoff. Logically, this means that a positive mass of quantile ranks must play at once — namely, a rush must occur. When might this rush occur?

Payoffs in a rush equal the average of the included quantile rank payoffs.[2] And in equilibrium, this rush payoff coincides with the adjacent payoffs from gradual entry. Yet given our strategic structure, when players are fearful, the earliest pre-emption payoff exceeds any average quantile payoffs in a late rush. In this case, only an early rush is possible. On the other hand, when players are greedy, the quantile rank payoff just after a rush dominates the averaged payoff in an early rush; here, only a late rush can occur.

Having developed the basic equilibrium framework, we illustrate the power of our

[2] There is an analogy in mechanism design to atomic play: there non-monotonic valuation functions get ironed so that for some interval of valuations people receive the good at random. An atom in our game corresponds to an ironed portion of the allocation function in mechanism design.

framework with a comparative static. In a *positive monotone payoff shift*, quantile rank payoffs shift so that they peak later, and later movers get relatively more. For instance, in the matching example this occurs if there is a stronger social penalty for moving early. In a financial bubble, a greater weight on higher relative performance causes such a shift; and in a possible liquidity rush, it is induced by greater deposit insurance.

A positive monotone payoff shift moves the strategic structure from fear to greed. For early rushes, it yields a larger rush volume, and shorter pre-emption game. For late rushes, it yields a smaller rush volume, and longer war of attrition. In brief, the closer to the most efficient time to rush, the larger the rush. We also explore behavior outside the rush. For log-concave time-payoffs (most non-exponential payoff ripen-rot-functions satisfy this assumption), the stopping hazard rate falls during pre-emptive phases and rises during war of attrition-phases. In an asset bubble, this implies that selling should intensify as one approaches the moment of the rush.

In the last section of this paper, we apply our general model in three famous models of rushes. We respectively adapt and subsume $(a)$ the matching insights in Roth and Xing (1994), $(b)$ the bank run model of Diamond and Dybvig (1983), and $(c)$ asset-price bubble, aggravated by relative performance incentives of mutual fund managers, as done by Shleifer and Vishny (1997). We construct toy models for each setting that project into our simple general framework.

Our comparative static applies to each of these three toy models. In particular, 'unraveling" in matching is the progressively early matching that occurs each year (Niederle and Roth (2003) study this for gastroenterology internships). We instead explain it as an equilibrium comparative static in our setting. Due to a strong drop in the supply of eligible applicants, hospitals feared a thin market. This shifted payoffs from high to low quantiles, as it became more attractive to move early (despite the negative social stigma). This caused earlier offers, and we argue smaller rushes.

Our asset bubble toy model must adopt a metaphorical interpretation of price as time. The reason is that our formal model only allows our strategies to depend on the fundamental or time, but not on price — which intuitively makes no sense. Our comparative static then predicts that later bubbles (larger price drops) induce smaller rushes. We verify this surprising prediction using a large data set (1993-2003) on opening sessions for the Dow Jones Industrial Average.

**Related Literature.** Our work builds on the timing games literature. Our setup is essentially a mix of a war of attrition and a pre-emption game — models in the literature can usually be classified as being one or the other.[3] Mixing the two allows us to gain the big-picture insights of common features of situations with economic rushes.

Wars of attrition have been explored in areas as diverse as duopoly exit (Fudenberg and Tirole (1986)), patent races (Fudenberg, Gilbert, Stiglitz, and Tirole (1983)), optimal investment timing (Chamley and Gale (1994) and Gul and Lundholm (1995)), or the adoption of a new technology (Farrell and Saloner (1986) and Choi (1997)). Murto and Välimäki (2006) develop a market exit model with a pure information externality (others' stopping decisions signal the quality of the market). All-pay auctions and all-pay contests have a similar flavor as only the last few/highest bids obtain the price; see Siegel (2007) for a recent insightful paper.

Pre-emption game-like settings have been used to explain market entry decisions (Reinganum (1981a,b), Fudenberg and Tirole (1985), Levin and Peck (2003), and Argenziano and Schmidt-Dengler (2008)) or patent races (Weeds (2002)). Brunnermeier and Morgan (2006) elaborate on clock games, where their analysis is a dynamic extension of the otherwise static global games approach. Hopenhayn and Squintani (2006) study pre-emption games in which players' underlying, privately known state (that determines payoffs) changes stochastically over time; since agents learn (and thus their information improves) over time, this can be seen as a pre-emption game in which payoffs (through better informed decisions) grow over time. Even financial bubbles can be understood as a pre-emption game (for example, Abreu and Brunnermeier (2003), who also assume unobservable actions): Everyone wants to sell before the bubble bursts but, by the same token, stay in as long as the bubble lasts. For recent empirical work on speculative bubbles (as prolonged mispricings) in foreign exchange markets see Brunnermeier, Nagel, and Petersen (2008): the price behavior in markets where traders extensively employ carry trades has features of bubbles and crashes ('Up the stairs, down the elevator'; the phrase is borrowed from Stefan Nagel) that are in line with the predictions of our model. The decisive feature of pre-emption games is that players prefer to act before others.

The remainder of the paper is structured as follows: In Section 2 we outline the details of our model. Section 3 formalizes greed and fear. In Section 4 outlines the conditions for the equilibria, Section 5 considers comparative statics in the payoff structures and describes properties of the adoption rates. In Section 6 we present three miniature models of economic examples that justify our reduced form approach and argue how empirical evidence supports our framework and results.

[3] In more recent work, Park and Smith (2008) study finite player timing games, but they have no payoff growth and thus their model does not admit slow pre-emption games.

## 2 A Simple Reduced Form Model with Timed Rushes

A continuum of mass 1 of identical players engage in a continuous time stopping game starting at $t = 0$. They can stop only once and stopping is irrevocable. Actions are either unobservable or committed to simultaneously before play, so that a player's strategy simply specifies the time he will stop.[4] A mixed strategy is thus a non-decreasing and right-continuous function $Q : [0, \infty) \to [0, 1]$ that measures the cumulative probability that a player has stopped by time $t$. We explore the Nash equilibria since there is one information set. As is typical practice in the timing games literature we ignore asymmetric equilibria for symmetry captures the anonymity of players' roles.

A player's payoff multiplicatively depends on his ordinal stopping quantile rank and his stopping time. Namely, there is a *reward scale factor* $v := [0, 1] \to \mathbb{R}_+$ which is a continuous and differentiable function of a player's stopping quantile. We assume that the very first stopper gets a positive payoff, $v(0) > 0$, and that these factors are "hill-shaped", first rising and then falling (Figure 1).

If players stop at the same time, they receive the average of their rank scale rewards — as if a fair lottery decided the order of the simultaneous stoppers. So if fraction $q$ of players has already stopped, and then fraction $p - q$ stops at exactly the same moment, then their rank reward scale factor is $1/(p - q) \int_q^p v(x)\ dx$. The overall *average rank scale factor* is $\int_0^1 v(x)\ dx$, which, without loss of generality, we normalize to 1.

For the purposes of a comparative statics analysis we sometimes index these rank rewards by $\theta \in \mathbb{R}$. Rank rewards categorized by $\theta$ are then assumed to obey the *monotone ratio property*: $v^\theta(x)/v^\theta(y)$ increases in $\theta$ for $x > y$. This property is often used in information economics as a requirement on signal-densities to express that higher signals imply higher values.

Apart from the quantile rank, there is also a time component of payoffs for which, in the spirit of the Hotelling's (1931) tree-cutting model, we assume that they first 'ripen' and then 'rot'. Specifically, the time payoff factor $\pi(t)$ smoothly rises to $\bar{\pi} := \pi(t_\pi) > 1 = \pi(0)$, peaks at the *harvest time* $t_\pi > 0$, and then falls smoothly. For instance, $\pi(t)$ may be the present value of a fundamental.

In what follows we shall use the following running example to illustrate our findings: $v(x) = -(x - \theta)^2 + (\theta - 1/12)^2 + 13/12$. This function is hill-shaped for $\theta \in (0, 1)$ and satisfies monotone ratio domination in $\theta$ and $x$.[5]

[4]For a sharpened focus on the timing and size of rushes, we consider only purely time-dependent (open loop) strategies. However, time can also be metaphorically interpreted by representing a underlying state variable upon which strategies may be conditioned. We will provide an example in Section 6.

[5]This function is a parabola with maximum rank $\theta$ that integrates to 1 on $[0, 1]$. Monotone ratio domination is synonymous to log-supermodularity; it thus suffices to check if $(\partial^2 \ln(v^\theta(x))/(\partial x \partial \theta) \geq 0$.

# 3 Greed and Fear

In the games that we study, being first is not as attractive as being second, and being last is not as attractive as being second to last. At the same time, in any symmetric equilibrium, a player cannot expect a quantile payoff larger than the average quantile payoff, $1$. This leaves them with the decision of whether or not to engage in the game or to pre-empt or outwait everyone else.

In particular, when people want to stop before the game, then they *fear being pre-empted*, if they want to outwait everyone else, they exhibit *greed*. We measure fear by $F = v(0) - 1$ and we measure greed by $G = v(1) - 1$. This motivates

**Definition (Greed and Fear)** *Players are* fearful *if $F > 0$, and they are* greedy *if $G > 0$.*

Of course people can't be simultaneously greedy and fearful because this would imply that all hill-shaped quantile payoffs are above average.

Monotone ratio shifts in rank rewards intuitively alleviate fear but raise greed because quantile payoffs are shifted towards higher ranks: being a late rank is not as costly any more (fear is reduced), but being early pays relatively less (greed increases).

**Lemma 1 (Payoff Shifts and Greed vs. Fear)** *Fear decreases and greed increases in $\theta$.*

When $v^{\theta'}$ dominates $v^{\theta}$ by monotone ratio domination, we can write $v^{\theta'}$ as $v^{\theta'}(k) = h(k)v^{\theta}(k)$ with $h$ being a function that increases in $k$. Since the sum of quantile payoffs remains constant (and quantile payoffs positive), this implies that $h(0) < 1$. Therefore fear decreases as $v^{\theta'}(0) - 1 < v^{\theta}(0) - 1$. Similarly, $h(1) > 1$ and thus greed increases as $v^{\theta'}(1) - 1 > v^{\theta}(1) - 1$.

In the running example, it is straightforward to check that $F(\theta) = v^{\theta}(0) - 1 = (4/3 - \theta) - 1$ decreases in $\theta$ and $G(\theta) = v^{\theta}(1) - 1 = 1/3 + \theta - 1$ increases in $\theta$. Thus fear is reduced, greed increases. Moreover, since $v(0) = 4/3 - \theta$ and $v(1) = 1/3 + \theta$, it is clear that people are fearful for $\theta < 1/3$ and greedy for $\theta > 2/3$.

# 4 Equilibrium Analysis

The basics of players behavior can be understood with standard tools from economic analysis. Players stop immediately if the marginal cost of delay exceeds the marginal benefit, and they will delay if the marginal benefit of delay exceeds the marginal cost. To actively engage in a symmetric mixed equilibrium marginal benefit and marginal

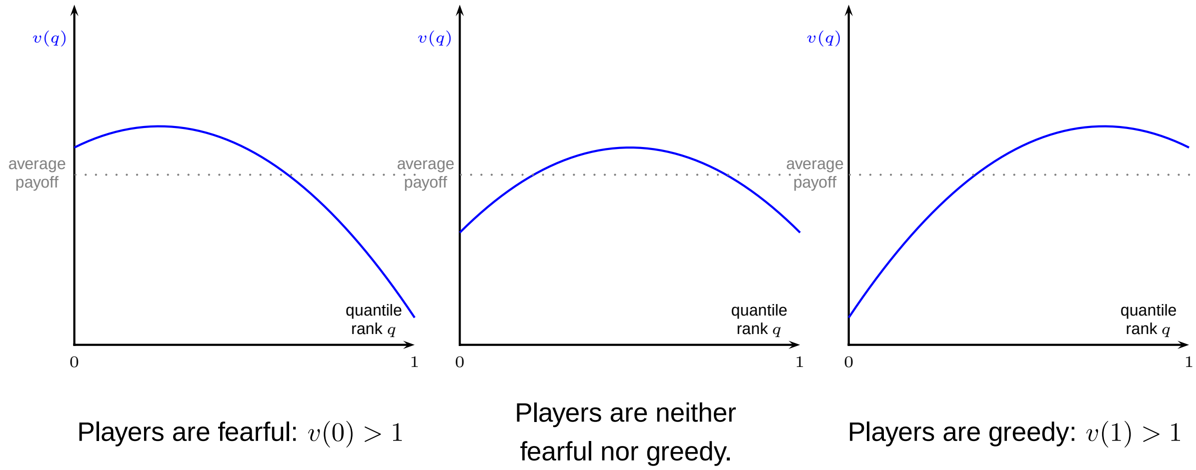


Figure 1: **Illustrations for Fear and Greed for the running example**. The left panel plots an example for rank rewards with fear, where the first payoff is larger than the average. The middle panel plots an example where there is neither fear nor greed, and the right panel plots an example for greed, where the last payoff is higher than the average payoff.

cost must exactly offset each other for the players. Analogous to standard economic problems, this occurs if the elasticity of time and quantile rank payoffs in time is $-1$:

$$\frac{\partial \log v}{\partial \log \pi} = -1. \tag{1}$$

Since actions are unobservable, when stopping at any point in time, a player knows only his time payoff $\pi$, but not his quantile rank. Instead, he has to compute his *expected* quantile rank payoff. Suppose players have each stopped with the common probability $Q(t) = q$,[6] and $Q(t)$ is continuous. Then the fraction of players that have stopped is also $q$ and, therefore, the expected quantile payoff coincides with the actual quantile payoff and is thus $v(q)$ (a formal argument is in the appendix). Applied to (1) we have that in equilibrium

$$\dot{Q}(t)\,\frac{v'(Q(t))/v(Q(t))}{\dot{\pi}(t)/\pi(t)} = -1. \tag{2}$$

Using this equation, we can make some simple observations about the nature of play: As $\dot{Q}$ measures the change in the quantile that has stopped, it is non-negative. Time payoff $\pi$ increases before the harvest time and decreases thereafter, so that there is a marginal delay benefit before the harvest time, and a marginal delay cost thereafter. Consequently, $v'$, which measures the quantile rank or strategic incentive to delay, must be negative for play to occur before the harvest time and positive for play to occur after

[6]In what follows we will use $q$ for realizations of $Q(t)$ and we will use $\dot{Q}$ to denote the derivative of $Q(t)$ (when it exists).

the harvest time — otherwise the quotient cannot be negative.

Next, equation (2) is a differential equation for $\dot{Q}$, the solution of which describes how people have to play in equilibrium so that indeed marginal benefits of delay coincide with marginal costs. For a definitive solution of the differential condition, however, we also need boundary conditions. To obtain these, recall that in a mixed strategy equilibrium, payoffs are constant. So once we know the payoff at one point, we know the equilibrium payoff of the game and close examination of the payoff will yield the right boundary condition. Looking at the elasticity interpretation it is obvious that the key focal point for payoffs is the harvest time. But before we discuss the boundary conditions, we will make a small detour.

We have already established that when time benefits increase, an equilibrium is possible only if quantile payoffs decline; analogously for marginal time delay costs. This yields two possible equilibrium constellations. The usual interpretation of a situation where a strategic incentive to stop opposes an exogenous benefit of delay is that players engage in a *pre-emption game* phase. The opposite case, when there is a strategic incentive to continue and an exogenous cost of delay is interpreted as a *war of attrition* phase.

As $v$ is hill-shaped there is a unique $q_v$ so that $v'$ changes its sign. Now by the above logic, any war of attrition phase transpires in the time interval after the harvest time, $[t_\pi, \infty)$, and any pre-emption game phase before the harvest time, $[0, t_\pi]$. Similarly, any war of attrition phase transpires in the *initial* $q$-interval $[0, q_v]$, and any pre-emption game phase in the *latter* $q$-interval $[q_v, 1]$. Together, these facts at once preclude equilibria having both war of attrition and pre-emption game phases.

Wars of attrition and pre-emption games are, in fact, played in a unique manner, determined by payoffs at the harvest time. These yield unique boundary conditions for the elasticity differential equation. To see this, observe that any war of attrition phase must start at the harvest time, and any pre-emption game phase must end precisely at the harvest time. Why? The equilibrium payoff for a war of attrition phase starting at the harvest time is $v(0)\bar{\pi}$.[7] Now suppose that the war of attrition starts at a later time. Then a player could profitably deviate, by stopping at the harvest time, because he would secure the same expected quantile rank payoff $v(0)$ but at a higher time scale factor. Similarly, if a pre-emption game ended at any time before the harvest time, then a player could profitably delay until the harvest time. He would then secure the same expected quantile payoff $v(1)$ but at a higher time scale factor.

While we now know how and when wars of attrition and pre-emption games will be played, two important questions remain: what happens at the end of the war of attrition phase when $v'$ changes its sign? And how is it possible that there is a pre-emption game

[7] In a mixed strategy equilibrium, payoffs are constant on the support.

phase (so that $q > q_v$) when there is no war of attrition preceding it?

The answer to both questions is that these situations call for a *rush*. In a rush, people stop with positive probability, which is sometimes called an atom. In a mixed strategy equilibrium, players must earn the same payoff at any time in the support. So the payoff from, say, the war of attrition that precedes a rush must coincide with the payoff in the rush. As we cannot have both a pre-emption game and a war of attrition as part of an equilibrium, we can focus on *initial* and *terminal* rushes. The intuitive payoff in an initial rush where people play with probability $q$, denoted by $V_0$, is the average of all initial expected quantile payoffs; similarly, when people have stopped with probability $q$ and then in a terminal rush, the payoffs are the average of all remaining expected quantile rewards, denoted by $V_1$. Formally,

$$V_0(q) = \frac{1}{q} \int_0^q v(x)\, dx \text{ and } V_1(q) = \frac{1}{1-q} \int_q^1 v(x)\, dx. \tag{3}$$

We can now apply standard economics to determine the relation of quantile payoffs $v$ and rush payoffs $V_0$ and $V_1$: $v$ is the marginal payoff of the average $V_0$ and thus $v$ and $V_0$ coincide whenever $V_0$ has a maximum; similarly for $V_1$.

This then answers the question of how we can have a pre-emption game without a preceding war of attrition: there is a rush before the pre-emption game. But not just any rush, there is a unique rush because at only one point do rush payoffs and pre-emption game payoffs coincide. Similarly, when there an equilibrium with a war of attrition, then it won't end with a subsequent pre-emption game but with a rush that is uniquely determined.

Next, the averages $V_0$ and $V_1$ have locally isolated maxima only if they are hill-shaped, and it turns out that their shapes are determined by fear and greed. To see this observe that as $v$ increases, $V_0$ averages small payoffs so that average and marginal can only intersect when the marginal declines. Since $v$ is hill-shaped, there can be at most one such intersection. There will be exactly one, if $V_0$ exceeds $v$ for the complete rush, that is when $q = 1$. The payoff from this complete rush is 1, so if and only if $v(1) < 1$, or, in words, if people are not greedy, then we have the situation when the rush and the pre-emption game payoff can coincide. A similar argument applies to fearfulness for $V_1$. This effectively proves the following lemma.

**Lemma 2 (Average Quantile Payoffs with Fear and Greed)** *$V_0$ is hill-shaped if and only if players are not greedy, $V_1$ is hill-shaped if and only if players are not fearful.*

In the above we have aimed to make an intuitive case for our equilibria, but we now need to be more specific with respect to some technical details. First, we only consider equilibria where play occurs on an interval of time, or possibly just a single focal moment in time. This eliminates a continuum of arbitrary equilibria in which time becomes a focal coordination device that is not otherwise mandated by primitives. Further, the only focal point in time that we consider is the harvest time $t_\pi$.

To simplify the exposition, we also assume that the most that a player can obtain by out-waiting all others, $\bar{\pi}v(1)$, exceeds what he could maximally obtain when all players start the game by rushing with probability $q$; this payoff is $\max_q V_0(q)$. This requirement holds if the harvest time payoff premium is large enough relative to quantile payoffs. Equally well, this inequality is met provided the last entrant is not overly penalized, since that precludes a gradual pre-emption game.[8] Together, these conditions ensure that there is a meaningful incentive to delay.[9]

We can now turn to outline the equilibria of this setup. Apart from the two classes outlined before —those with war of attrition phase followed by a terminal rush and those with a pre-emption game phase preceded by an initial rush— there is a third class that involves a complete atom at the harvest time.

**Theorem (Equilibria with Rushes, Pre-emption Games and Wars of Attrition)**

*(1)* EQUILIBRIUM WITH A LATE RUSH

- $(a)$ *If players are fearful, then there is no equilibrium with a war of attrition.*
- $(b)$ *If players are not fearful, then there exists a unique equilibrium with a war of attrition that is followed by a rush.*

*(2)* EQUILIBRIUM WITH AN EARLY RUSH

- $(a)$ *If players are greedy, then there is no equilibrium with a pre-emption game.*
- $(b)$ *If players are not greedy, then there exists a unique equilibrium with a pre-emption game that is preceded by a rush.*

*(3)* THE EFFICIENT RUSH

*A rush at the harvest time is an equilibrium only if players are neither greedy nor fearful.*

The text leading to this theorem substitutes a proof: Lemma 2 establishes that there is no terminal atom if people are fearful, similarly for greedy people and initial atoms.

A complete rush at the harvest time pays 1 and can thus be an equilibrium only if it does not pay for a player to play before (when he gets $v(0)$) or after this rush (when he gets $v(1)$). Thus either greediness or fearfulness preclude such an equilibrium.

[8] As an example consider $\pi(t) = (t+1)e^{-rt}$ with $r = 10\%$. Applied to the running example, the condition always holds: since $\bar{\pi} \approx 4$ so that $v(1)\bar{\pi} \geq 4/3$ for all $\theta \in [0,1]$ whereas the maximum of $V_0$ is $\frac{3}{4}(\theta - \frac{2}{3})^2 + 1 \leq 4/3$ for $\theta \in [0,1]$.

[9] This assumption is by no means necessary to obtain our results, but the description becomes more cumbersome; details of the equilibrium analysis for this case are posted on our website.

In the context of our framework, the Efficient Rush is the efficient benchmark. For instance in matching, it arises when all companies make offers at the time when all or the optimal amount of information has been revealed and when no-one pre-empts this most efficient scenario. In bank runs, the efficient rush arises when all investors withdraw their deposits at the expiration of a long-term investment.

The big insight is that when players are fearful, the only equilibrium is one with an early rush, and when players are greedy, the only equilibrium is one with a late rush. Put differently, when greed dominates fear, people take too much risk and delay beyond their optimal stopping point. When fear dominates greed, people forego payoffs and rush too early.

# 5 The Size of Rushes and the Length of Play

To obtain a better understanding of the impact of quantile payoffs and to generate testable implications, we now consider monotone ratio shifts in the quantile payoff structure. While quantile payoffs themselves are often unobserved, one can generate proxy variables that capture certain qualitative features of quantile payoffs. For instance, in liquidity runs, deposit insurance provides a lower bound for high-order quantile payoffs, the strength of the relation between fund and investors proxies the extent to which withdrawal causes a penalty for early movers. We find

**Proposition 1 (Payoff Shifts)** *When $\theta$ rises, in the late rush equilibrium the war of attrition phase lasts longer and the terminal rush shrinks, while in the early rush equilibrium the pre-emption game phase starts later and the initial rush grows.*

With hill-shaped rank structures, a monotone shift favors later movers, for instance, by shifting the hill-top to the right. Players are then more willing to hold out so that war of attrition phases last longer. The shift towards later ranks has the reverse effect on pre-emption game phases: the rush preceding the pre-emptive phase grows and also the survivor payoff grows. In combination, the elasticity of decreasing rank rewards shrinks so that the exit rate increases. This speeds up the pre-emptive phase. Figure 4 is a general schematic that illustrates both the length of the gradual entry phases respectively, and the size of the rushes. Figures 2 and 3 illustrate the comparative static for the running example, the explicit solutions are in the appendix.

Our final result allows us to distinguish wars of attrition from pre-emption games by studying players' **Exit Rates.** These can be measured by a strategy's density $\dot{Q}(t)$: While

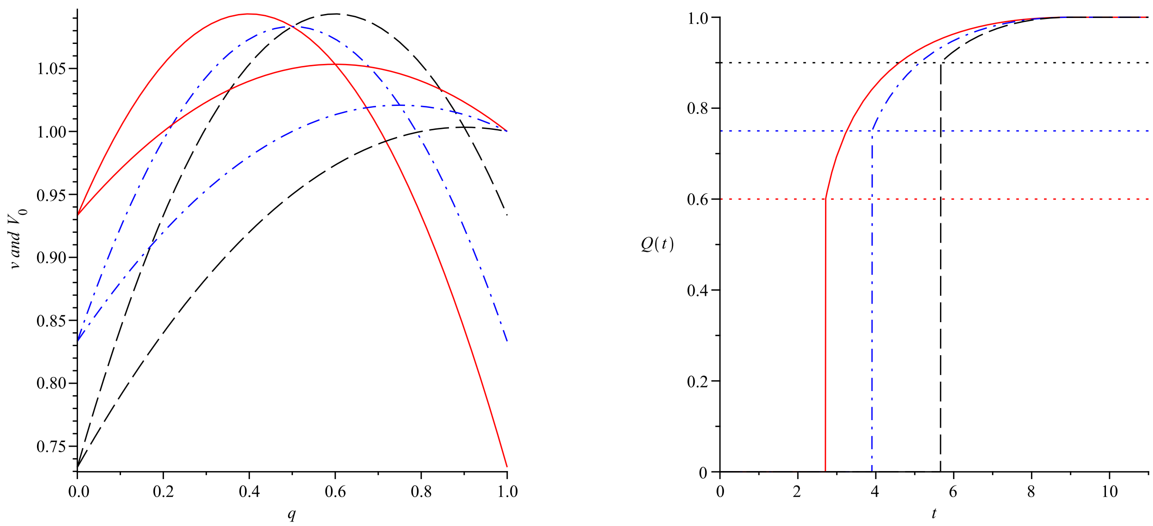


Figure 2: **Rank-Shifts and their impact on Equilibrium Play: The comparative static illustrated for early rushes.** The figure is based on the running example (see the appendix for explicit solutions) for $\theta = 2/5$ (red or and solid line), $\theta = 1/2$ (blue and dash-dotted line) and $\theta = 3/5$ (black and dashed line). The LEFT PANEL plots the quantile payoff functions and the quantile payoffs from late rushes. As can be seen in the left panel, as $\theta$ increases, the tops of the hill for both $v$ and $V_0$ move to the right. To determine equilibrium strategies explicitly, we assume that $\pi(t) = (t+1)e^{-t/10}$ so that $t_\pi = 9$. The RIGHT PANEL then plots the equilibrium strategy $Q$ for the early rush. The dotted horizontal lines signify $\arg\max_q V_0(q)$ for the respective $\theta$'s. Entry begins at the time when $Q(t) \equiv \arg\max_q V_0(q)$ with a rush of size $\arg\max_q V_0(q)$ and continues until time $t_\pi = 9$ (whence $Q(t) = 1$). The larger is $\theta$, the later in time is the rush, and the larger is the rush.

a player's strategy measures the accumulated probability of stopping, its density captures the frequency of stopping decisions. If these densities differ between pre-emption games and wars of attrition, then there are testable implications of our equilibrium predictions — increasing or decreasing rates should be detectable in suitably constructed panel data of comparable rush-situations.

To obtain conclusive results we assume that quantile rewards are concave[10] and that time payoffs $\pi$ are log-concave. The latter is a very weak premise and hardly rules out candidate functions. For instance, if $\pi(t) = \gamma(t) \cdot e^{-rt}$ (so that $\pi$ is the discounted value of a fundamental $\gamma$), then log-concavity holds as long as $\gamma$ is not exponential. In the running example, quantile rewards are concave (because it is a parabola).

**Proposition 2 (Exit Rates)** *Assume a hill-shaped and concave quantile reward function $v$ and a log-concave time-payoff $\pi$. Then the adoption rate $\dot{Q}$ is increasing in time $t$ during a war of*

[10]We have already argued that when rank rewards are hill shaped, so are expected rank rewards $v$. Thus the shape of $v$ is preserved under expectations. Curvature is also preserved, for the same reason: $v'$ a polynomial with coefficients given by the first differences of $v$; and thus $v''$ is a polynomial with coefficients given by the first differences of the first differences. When $v$ is concave, these double-first differences are negative, and so will be $v''$.

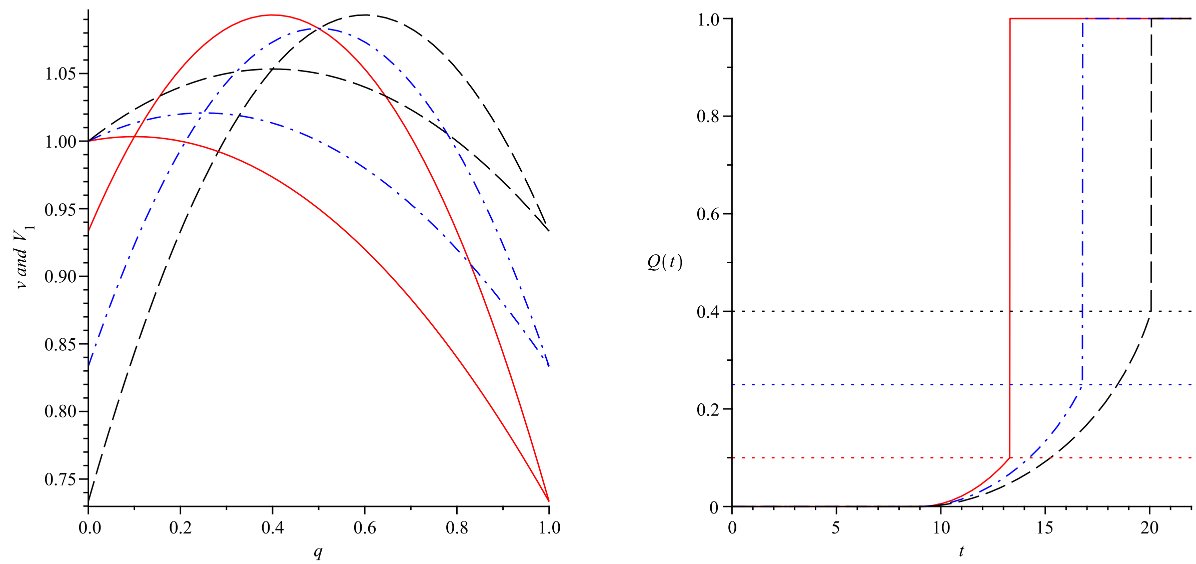


Figure 3: **Rank-Shifts and their impact on Equilibrium Play: The comparative static illustrated for late rushes.** The figure is based on the running example for $\theta = 2/5$ (red or and solid line), $\theta = 1/2$ (blue and dash-dotted line) and $\theta = 3/5$ (black and dashed line). The LEFT PANEL plots the quantile payoff functions and the quantile payoffs from late rushes. As can be seen in the left panel, as $\theta$ increases, the tops of the hill for both $v$ and $V_1$ move to the right. To determine equilibrium strategies explicitly, we assume that $\pi(t) = (t+1)e^{-t/10}$ so that $t_\pi = 9$. The RIGHT PANEL then plots the equilibrium strategy $Q$ for the late rush. The dotted horizontal lines signify $\arg\max_q V_1(q)$ for the respective $\theta$'s. Entry begins at $t_\pi = 9$ and continues until the time when $Q(t) \equiv \arg\max_q V_1(q)$, at which point there is rush that ends the game. The larger is $\theta$, the later in time is the rush, and the smaller is the rush.

*attrition phase and decreasing during a pre-emption game phase.*

Intuitively, wars of attrition intensify whereas pre-emption games taper off. For instance in a bubble, before the rush the selling pace quickens, as fund activity intensifies. This was indeed observed by Griffin, Harris, and Topaloglu (2006) for the dot-com bubble.

# 6 Some Economic Rushes with our Reduced Form

For a variety of economic settings, we will now construct the sparest plausible model that yields our reduced form for payoffs. We justify in this section that rushes occurred, and that the economics matched the assumptions of a harvest time in time payoffs, and a hill-shaped rank order. We then make predictions based on our major propositions for these economic settings using these representations.

The ideas that underly these models are not new but often expressed in the literature, albeit mostly casually. The three main examples that we discuss are an asset bubble, a liquidity run, and job market matching. Our models are based on the reasoning that we outlined in the introduction: "I know that I am in a bubble, and I know that it will

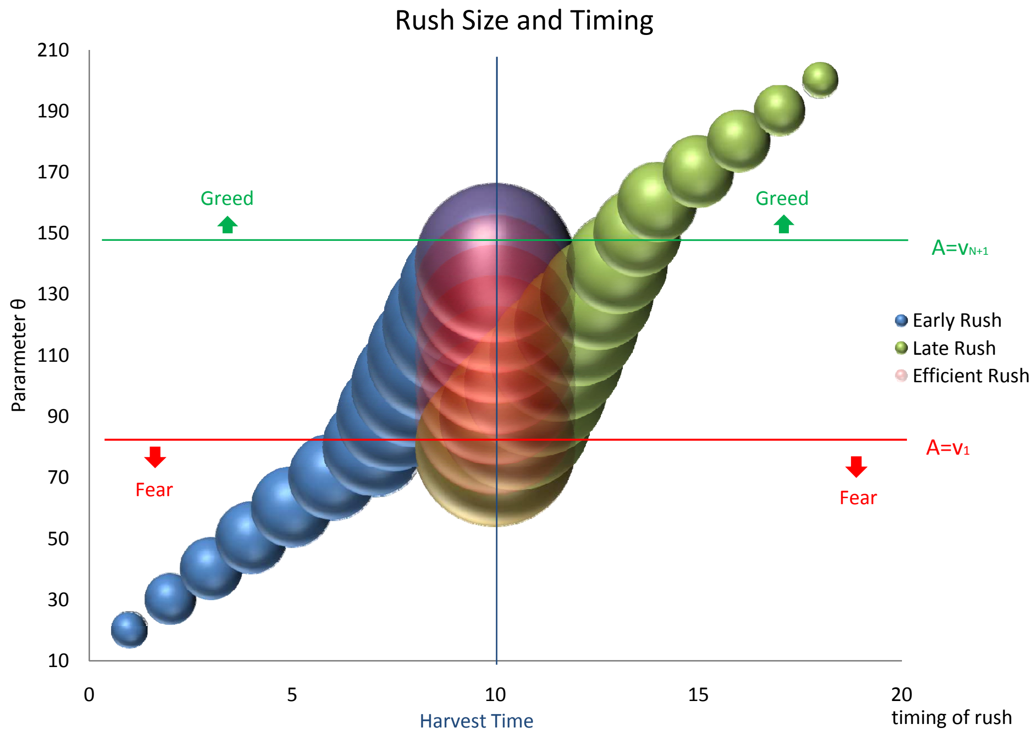


Figure 4: **General Illustration of Proposition 1.** The figure illustrates simultaneously the size of rushes, the timing of rushes and the length of gradual entry. The larger a rush, the larger the dot in the above figure; the further is the dot from the harvest time, the longer is the gradual entry phase. Next, the larger parameter $\theta$, the larger is the early rush, and the closer it is to the harvest time. For late rushes, the relation is the reverse: the larger $\theta$, the smaller the rush and the further it is from the harvest time. Moreover, the figure also illustrates that when there is fear, there is no late rush and likewise, when there is greed, then there is no early rush. Absent fear and greed, the efficient rush at the harvest time is also an equilibrium.

collapse when everyone sells. So I want to get out before that happens. But suppose I sell out early and a competitor stays in a little longer and gains a lot more than I. Wouldn't that bother me a great deal?"; "This company may go bust, so I want to pull my funds before it happens. But suppose it becomes known that I was the first to pull my funds, effectively pulling the rug under their feet. Would I get punished by other borrowers? Would this exclude me from future business?"; and "I would like to get the best candidate for the job and I don't want to wait until all the good ones have been hired by others. But if I start hiring before everyone else, would my colleagues sneer at me at the next convention?" Based on these thoughts, we build three toy models.[11]

[11]A technical disclaimer: In our main analysis we normalized the average quantile payoff to 1 to simplify the exposition. The payoff structures in the examples that we develop now will not be normalized, again to simplify the exposition. The bounds for greed and fear that we discuss below are obtained by comparing $v(0)$ and $v(1)$ to the average payoff, $\int_0^1 v(q)dq$, instead of 1.

## 6.1 The Rush to Sell in a Bubble

OVERVIEW. On the most basic level, financial bubbles are a pre-emption game in the spirit of Blanchard and Watson (1982): prices are rising and people must decide when to exit the bubble. At the same time, as more people exit, the bubble becomes more likely to burst. Translated into the language of this paper, in a pure pre-emption game quantile payoffs are strictly declining so that people are fearful — the resulting early rush precludes bubbles from arising in the first place.

Yet the strict pre-emption game perspective ignores the fact that financial market participants see their payoffs not in absolute but relative terms. Most of the market trading activity stems from institutional investors[12] who act on behalf of others and who are usually paid according to how they do *relative* to their peers.[13] As a consequence, leaving a growing bubble is costly because those peers who stay in a little longer and get out just early enough to avoid the crash will do better.[14] Comparing payoffs to one's peers is, of course, not restricted to institutional investors — the 'keeping up with the Joneses' behavioral phenomenon is commonly observed in many contexts, including in financial markets.

We assert that fund managers take a planning perspective so that they commit ex ante to act at a specific price — thus time here is a metaphor for if prices grow over time then the commitment to sell at a specific time is synonymous to committing to sell at a specific price.

The harvest time is the price level so that from an ex ante point of view any further price gains are outweighed by the possibility of a crash. As we argue, what keeps people in the market is that they greedily want to outsmart their peers. The mixed strategies over 'exit times' that we employ in most of the paper have the interpretation of a gradual selling or 'unwinding' of a position.

A TOY MODEL. The bubble bursts after a fraction of $q$ investors has sold with probability $q \cdot \rho$, $\rho \in (0, 1]$. Variable $\rho$ denotes the resiliency of the market to sustain the bubble even if all strategic fund managers exit. The larger $r$, the lower the impact of the fund managers' sales in popping the bubble. Thus the bubble still exists after $q$ have stopped

[12]Dasgupta and Prat (2007) report: "On the New York Stock Exchange the percentage of outstanding corporate equity held by institutional investors increased from 7.2% in 1950 to 49.8% in 2002 (NYSE Factbook 2003). For OECD countries as a whole, institutional ownership now accounts for around 30% of corporate equity; see Nielsen (2003). Allen (2001) presents persuasive arguments for the importance of financial institutions to asset pricing."

[13]See Berk and Green (2004) or Shleifer and Vishny (1997). Technically, mutual fund managers cannot be paid directly according to the returns that they generate, but instead get paid relative to the funds that they manage. But when a fund does well relative to its peers, then the fund usually experiences inflows of cash. Consequently past performance does have a strong influence on future payoffs.

[14]For a longer discussion of this argument see Brunnermeier and Nagel (2004).

with probability $1 - q\rho$. Normalized, the first stopper gets the selling price if the bubble did not burst and 0 if it did burst. If it did not burst, then later ranks get higher compensation (through increased fund inflows) of $1 + \theta q$ of the selling price, where parameter $\theta \geq 0$ measures how responsive markets are to relative performance.

We assume that the selling price $p(t)$ increases monotonically and less than exponentially over time until the bubble bursts whence it drops to 0. Moreover, the bubble bursts for exogenous reasons as the price grows with probability $1 - e^{-r \cdot p(t)}$. To see the rationale for this exogenous bursting, consider the following excerpt from an article by Fareed Zakaria in Newsweek:

> "'Leverage' is the fancy Wall Street word for debt. It's at the heart of the current crisis. Warren Buffett explained the problem in his inimitable way on "The Charlie Rose Show." "Leverage," he said, "is the only way a smart guy can go broke…You do smart things, you eventually get very rich. If you do smart things and use leverage and you do one wrong thing along the way, it could wipe you out, because anything times zero is zero. But it's reinforcing when the people around you are doing it successfully, you're doing it successfully, and it's a lot like Cinderella at the ball. The guys look better all the time, the music sounds better, it's more and more fun, you think, 'Why the hell should I leave at a quarter to 12? I'll leave at two minutes to 12.' But the trouble is, there are no clocks on the wall. And everybody thinks they're going to leave at two minutes to 12." *Newsweek*, October 20, 2008.

Translated into this paper, investors do not know when it's midnight, but the chance of having passed the two minutes to midnight bound rises over time. Consequently, the bubble is still around at time $t$ when fraction $q$ has stopped with probability $(1 - q\rho) \times e^{-r \cdot p(t)}$. The payoff that one gets when the bubble has not burst is $(1 + \theta q) \cdot p(t)$. In combination, the payoff from stopping at time $t$ as quantile $q$ is

$$\underbrace{(1 - q\rho) \times (1 + \theta q)}_{v(q)} \times \underbrace{e^{-r \cdot (t)} p(t)}_{\pi(t)}.$$

For $\theta \in (\rho, \rho/(1 - 2\rho))$, the combined term $v(q) = (1 - q\rho)(1 + \theta q)$ is hill-shaped in $q$. For $\theta < 3\rho/(3 - 2\rho)$, people are fearful, and for $\theta > 3\rho/(3 - 4\rho)$ people are greedy. The time-related component $e^{-r \cdot p(t)} p(t)$ is hill-shaped in the price (or $t$).

Our analysis showed that if people are fearful then a bubble does not arise. But if people are greedy, then they will ride the bubble and a late rush occurs. In this toy model, greed is caused by relative payoff compensation. Brunnermeier and Nagel (2004) confirm this idea showing that hedge funds, who are subject to such relative compensation, rode bubbles for individual tech stocks.

Our first comparative static, translated into an environment with a financial bubble states that as relative payoff compensation becomes more pronounced, greed increases and thus people ride the bubble for longer. At the same time, even though people get greedier, they also slowly unwind their positions. So while the rush point is reached later in time, the rush itself is smaller because people have already begun to sell their holdings. This notion was tentatively confirmed by Griffin, Harris, and Topaloglu (2006) who document how institutional investors sold their positions prior to the bursting of the tech-bubble. Moreover, as the bubble has grown for longer, when it bursts, the price drop is larger. Our second comparative static predicts that as the bubble grows, people exit the market at an increasing rate.

Empirical evidence for first comparative static would thus show that prices drops and selling volume in a rush are negatively related. We have not found a fitting analysis in the literature, and thus assembled some data for the companies in the Dow Jones Industrial Average. While it is straightforward to determine price drops, it is more difficult to determine a rush that fits our formulation. A rush to sell would appear in the trading volume so one may be tempted to merely use daily volume. Yet this data is 'contaminated': if there is a rush that triggers a price drop, then market participants will learn about it. This may lead to more selling which biases the size of the rush. Our analysis, however, concerns the initial, triggering rush.

As a proxy for this rush we use the selling volume that obtains during the opening session. This provides a controlled environment in which market participants would not know of a rush until the specialist posts the results of the opening session. This data is therefore not contaminated by the reaction to a rush.[15]

Specifically, we obtained data for the NYSE listed DJIA companies as of May 2008; Microsoft and Intel are traded on Nasdaq where there is no formal opening session with trades. They were thus omitted from the data. Our data spans the period from 01/1993–12/2002. Data on closing prices, dividends, and stock splits is from CRSP; returns are computed close-to-close and volume and price-changes were adjusted for dividends and splits. The opening volume is the size of the largest transaction on the NYSE in the first 30 seconds of trading; this number is obtained from the TAQ database.

[15]In the opening session at NYSE the specialist collects orders and sets the price that equates demand and supply. The details of the matching rules are rather involved and the price setting choices are not fully transparent; for instance, the specialist may take positions to ensure that prices changes are not too extreme, s/he may take positions if 'liquidity on one side of the market has dried up' and so on. The imbalance of orders was not posted —until recently— while the opening session was still under way. Consequently, we have a controlled environment possible.

For these companies, we first determined all days when there was a price drop relative to the preceding trading day's closing prices. We then determined whether the opening volume is buyer or seller dominated by applying a procedure akin to so-called Lee and Ready (1991) 'tick-tests' that are commonly employed in the financial market microstructure literature,[16] and then we used the data for seller-dominated volume only. To be able to compare stocks, we scaled each stock's volume by its average open volume.

Figure 5 plots seller-initiated volume against price drops. Our theory predicts that even small selling volume at the open can trigger a large price drop during the remainder of the trading day and that large trading volume can go along with only small price drops. Thus to support our theory the figure should display a positive relation between the rush measure and price declines.

The figure plots all combinations of selling-volume and price declines, and most of the data points are concentrated around the origin. This indicates that usually moderate selling volume at the open coincides with moderate price declines. Our results, on the other hand, intuitively apply to the tails, suggesting that large price drops be coupled with low volume or large volume be coupled with small price drops.

Visually, the graph indeed suggest that there may be an increasing, possibly concave relation among the tails. A formal analysis confirms this: Excluding cases where simultaneously returns are not too negative and rushes small (i.e. excluding a rectangular at the origin), a regression of returns on rushes always yields a positive slope coefficient that is significant at the 99% level. For instance, if we look only at returns smaller than $-5\%$ and rushes larger than 5, then the coefficient on the rush is .0031 with a t-statistic of 18.8. For returns smaller than $-1.5\%$ and rushes larger than 2, the coefficient is .00067 with a t-statistic of 8.73. Almost all variations yield the same outcomes concerning the sign of the slope, as do quantile regressions between the tails. Further, even when including the points around the origin, fitting a negative exponential function to mimic a concave relation (i.e. running a regression on log-price-drops) yields a significant (99% level), increasing relation.

## 6.2 Liquidity Runs

OVERVIEW. Bank runs were a commonly observed phenomenon in the early part of the 20th century in the U.S. (and famously portrayed in movie pictures such as 'It's a wonderful life'); but banking panics or crises continue to arise in developed and developing

[16]The procedure is employed to distinguish buyer- and seller-initiated trades and it goes as follows: if the opening price is higher that previous day's closing price, then there must have been more buyers and thus the opening transaction is classified as buyer-dominated; the reverse for seller-dominated. In our analysis of selling-rushes we care for the seller-dominated transactions.

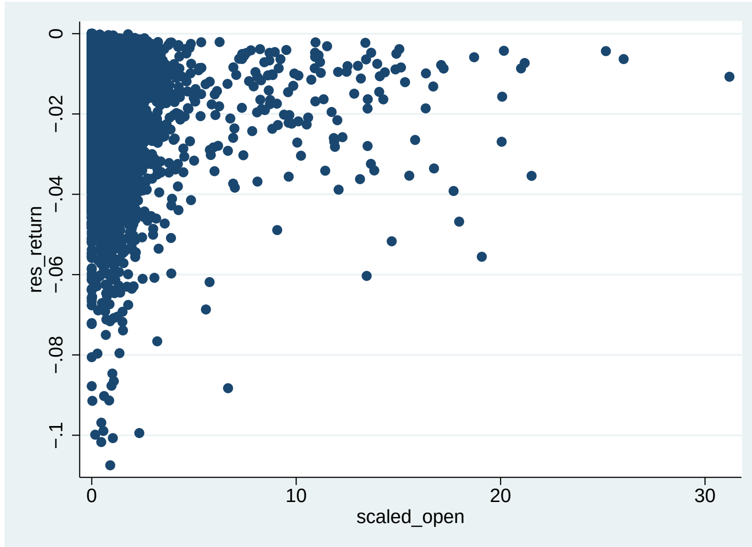


Figure 5: **Data Plot**: Returns are on the vertical axis, our selling rush-measure is on the horizontal axis. Specifically, the selling-rush measure is the selling volume at the open (buying vs. selling being identified by a Lee-Ready 'tick-test') scaled by the stock's average opening volume. The returns are computed as the percentage price change from the rush day's open to close; the results remain unaffected if we use the close-to-close returns.

countries alike.[17] Recent examples are the 2007 run on Northern Rock in the UK (despite deposit insurance), or the run on unregulated investment vehicles such as hedge funds in the wake of the U.S. sub-prime mortgage crisis. A more general phenomenon are liquidity runs which may arise when companies find themselves in financial trouble. In such a situation, lenders may refuse to roll-over short-term loans.

Most theoretical models in the literature employ two period models, e.g. Diamond and Dybvig (1983) or Allen and Gale (1998). While inspired by Diamond-Dybvig, our setting is inherently multistage. In the classic bank run literature, a bank run is the bad one of several equilibria. Our analysis argues when runs are the only equilibrium and it explains under which conditions a run would not happen. Specifically, whether or not we see a run is determined by whether or not there is greed or fear: small payoffs for high quantiles are the result of fear (that the company collapses); small payoffs in the low quantiles are due to greed (the penalty for running the bank first is high, so people may hold out). If fear dominates, then we have a rush; if greed dominates, then we have no rush. Beow we develop a simple model that captures these features.

While the bursting of a bubble is intrinsically a late rush (otherwise, there would be no bubble to begin with), liquidity rushes can be late or early.

[17]Finland (1991-93), Norway (1988-92), Japan (1992-present), Spain (1977-85), and Sweden (1991), several Asian countries (1997-98), Russia (1998), Brazil (1999), Turkey (2000) and (2001), Argentina (2001), and England (2007; 'Northern Rock').

An early liquidity run occurs when people withdraw their deposits long before the project matures, when supposedly the yield is largest. The harvest time there corresponds to the time when the project expires, so a late rush (after harvest time) is not a possibility. An early rush may not occur by design alone. If people are greedy because, for instance, there is a very high explicit or implicit penalty on people who make the first withdrawal, then an early rush will not occur and instead all investors delay until the expiration date (which can be interpreted as an efficient 'rush' at the harvest time).

Yet liquidity runs can also be late: the first warning signs for the recent Subprime crisis were visible in early 2007 when the spread between high- and low-rated mortgage backed securities increased strongly. In August 2007 the noise became even louder with Bear Stearns closing two in-house hedge funds. Yet the full market reaction emerged only in the Summer of 2008 when several institutions ran out of cash. One can hypothesize that there were strong pressures not to pull funds early. Our model then predicts excessive delay.

A TOY MODEL. An investment fund expires at time $T$ (cost-free withdrawals are possible at or after $T$), although the fund may collapse for exogenous reasons up until time $t$ with exponential probability. Payoffs at time $T$ are, for simplicity, normalized to 1. Early withdrawals carry two penalties: First, we assume that early withdrawls pay only fraction $(t+1)/(T+1)$ of the terminal payout $T$. Second, there is a reputational penalty as early stoppers are excluded from future projects with chance $1-q\rho$, where $\rho$ measures the intensity of the reaction. For instance, in the case of a successful turnaround Bear Stearns may shun those short-term lenders that called their loans earliest. So while the investor obtains funds $(t+1)/(T+1)$, he would like to get these reinvested at some later stage and can do so only with probability $\rho q$.

For simplicity we assume that the NPV of future projects is normalized to double the funds that the investor currently has deposited there, $F$. Then

$$\mathsf{Pr}(\text{reinvest}) \cdot 2 \cdot F + (1 - \mathsf{Pr}(\text{reinvest})) \cdot F = q\rho \cdot 2 \cdot F + (1 - q\rho) \cdot F = (1 + q\rho)\, F.$$

Of course, the funds here are $F = (t+1)/(T+1)$.

Being a late withdrawer, however, comes at a cost because the institution may run out of cash and collapse. This occurs with probability $q/\theta$ where parameter $\theta$ measures the liquidity of the institution (be it by cash deposit requirements, by the quality of the backing parent financial institution or by the nature of the investments). If the institution collapses, payoffs are normalized to 0. Positive payoffs are thus obtained only with probability $1 - q/\theta$. Then the total payoff from stopping as quantile $q$ at time $t < T$ is

$$\underbrace{\left(1 - \frac{q}{\theta}\right)(1 + \rho q)}_{v(q)} \; \underbrace{\frac{t+1}{T+1}\, e^{-rt}}_{\pi(t)}.$$

The time-scaling factor $\pi(t)$ payoff peaks at time $t = (1 - r)/r$. So if $r < 1/(T + 1)$, then the harvest time is the terminal time $T$, otherwise $t_\pi < T$.

Quantile payoffs consist of a rank-increasing component, given by the future business consideration, and a rank-decreasing component, given by the probability that the institution implodes due to a lack of liquidity.

The quantile payoffs peak at $q^* = (1 - \rho\theta)/2\rho$ for $\theta \in (\rho^{-1} - 2, \rho^{-1})$; for $\theta \geq \rho^{-1}$ function $v$ is monotonically increasing, for $\theta \leq \rho^{-1} - 2$ it is monotonically decreasing. Comparing first and last quantile payoffs to the average, investors are fearful if $\theta > \rho^{-1} - 4/3$ and they are greedy for $\theta < \rho^{-1} - 2/3$.

Applied to our comparative static, large values of $\theta$ imply that a backing investment bank keeps a fund afloat even if a large fraction of its investors has decided to withdraw. If the implicit promise is credible enough so that greediness is fostered, then an early liquidity run is not an equilibrium. Deposit insurance has a similar effect in raising high quantile payoffs; repeat interactions and the threat of losing a long-term partnership lower early payoffs ($\rho$ increases) through the penalty for the early abandonment. If, however, the backing financial institution is weak so that $\theta$ is small, then players are fearful, thus engaging in an early run. To prevent an early rush entirely one needs a sufficient penalty for the early quantile — deposit insurance with a cap, as is currently the case, may not be sufficient to eliminate fear entirely.

Our second comparative static predicts that after the run, the rate of withdrawals is first large but then tapers off. Thus an institution that has the funds to survive the early run, may observe a high withdrawal rate soon after the run. But the rate will decline so a linear prediction as to the speed may cause more unease than necessary and calls for government protection will be premature.

## 6.3 The Rush to Match

OVERVIEW. Al Roth's many works on matching markets contain several ideas that loosely connect to our framework. Pairwise matching markets are a well-studied area of economics with rushes.[18] For instance, college fraternity recruiting is most associated

[18] See Roth and Xing (1994).

with a scramble to match — hence the moniker "rushes".[19] Matching markets share the two key features that we identify: First, there is a harvest time, since early or late matching generally involves inefficiencies. A high school basketball player might not be properly trained for the NBA. Or the third year medical school student might not yet know his speciality field preferences. On the other hand, if they match too late, then they fail to exploit their skills. Second, there are social punishments for early matching in many models,[20] so that being among the first is not ideal. But an increasingly thin market awaits the latest entrants. There is obviously a myriad of possible models; we will provide one, and show how it reduces to our strategic form.

A TOY MODEL. We will now formalize the ideas implicitly expressed in Roth's work on matching. There is a mass of 1 of high types; if hired, they yield payoff 1. When stopping, firms search the pool of applicants for a high type. Their success rate depends on a search technology $\rho \in (0,1)$ (the higher the better) and the mass of high quality workers in the pool. If they cannot secure a high type worker, they settle for a low type worker who yields a lower payoff that we normalize without loss of generality to 0. The mass of high types after fraction $q$ of firms has stopped changes linearly at a rate which depends on the mass of high types present, $A(q)$, and the technology $q$. Thus

$$A(0) = 1 \text{ and } A'(q) = -\rho A(q) \quad \Rightarrow \quad A(q) = e^{-\rho q}.$$

A company's total utility of hiring a high quality worker is $\pi(t) \cdot (1 + q\theta)$, where $\pi(t)$ measure the increasing-decreasing time payoff (with a harvest time) and $(1 + q\theta)$ is the social penalty of early quantile movers. This gives rise to the total payoff of

$$\pi(t) \underbrace{(1 + q\theta)e^{-\rho q}}_{v(q)}.$$

The quantile dependent component of this payoff, $v(q)$, is maximal at $q^* = \rho^{-1} - \theta^{-1}$. Consequently, the quantile payoff is hill-shaped for $\theta \in (\rho, \rho/(1-\rho))$. Next, comparing the payoffs for $q = 0$ and $q = 1$ with the average quantile payoff reveals that people are fearful for

$$\theta < \frac{\rho(1 - \rho - e^{-\rho})}{e^{-\rho}(1 + \rho) - 1},$$

[19]Mongell and Roth (1991) observed that the name rushes owed to the unraveling tendency of such markets, as they moved earlier and earlier, not necessarily because of the speed.

[20]Avery, Jolls, Posner, and Roth (2007) describe the latest rules imposed on the market of federal judicial law clerks. While no explicit penalties are mentioned, there is social pressure to adhere to the rule. At the same time, about one quarter of the judges interviewed admitted to having made offers before the official starting time.

and they are greedy for

$$\theta > \frac{\rho\,(1 - e^{-\rho}(1+\rho))}{e^{-\rho}(1+\rho) - 1 + e^{-\rho}\rho^2}.$$

So let us relate this toy-model and our analysis to some real findings. Unraveling in matching is generally an example of an early rush and our analysis helps understand when early rushes do and do not occur: an early rush is only *not* an equilibrium if payoffs are sufficiently back-loaded so that people are greedy, i.e. if there is a reason for people to delay relative to others. For instance, the market has to be sufficiently thick even for late movers ($\rho$ would be small); alternatively, a sufficient penalty for early movers can cause higher quantiles to receive relatively larger payoffs ($\theta$ is large). Our second comparative static predicts that, following the rush, the rate at which companies enter the market is first large and then tapers off.

A famous example for unraveling in matching is the market for gastroenterology internships, as described in Niederle and Roth (2003), McKinney, Niederle, and Roth (2005) and Niederle, Proctor, and Roth (2006).

The market formerly operated under a centralized matching mechanism which broke down (for the first time) in 1996. At that time, the market experienced a negative supply shock of positions. Prospective candidates, in turn, anticipated this, did not apply and thus there was an even greater reduction in applicants than positions. This lowered the total expected payoff for hospitals. While there was still a negative stigma of making early offers and breaking the system, fear increased so that the efficient outcome —a 'rush at the harvest time' in our model— was no longer possible. Instead, people played an early rush equilibrium.

Over the following years, a dynamic developed that is captured by our comparative static: The negative stigma of early offers declined increased fear further, pushing the rush forward in time.

While we have no data ourselves, Niederle, Proctor, and Roth (2006) (which addresses the second market breakdown in 2005) offers some insightful graphs that illustrate both the dynamic effect and the existence of a rush. Figure 1 in their paper (p. 219) illustrates how entry moved earlier in time and was stretched out over a longer time span, as predicted by our Proposition 1. Their Figure 3 (p. 220) illustrates the early rush that occurs in September.

In the matching examples outlined by Roth et al., there is arguably a harvest time, which is either the date of the centralized match, or it is the graduation date of the students. One way or another, there is no possibility for a late match. The contribution of our model here is to outline that the efficient match at the harvest time is the unique outcome only if people are greedy.

Yet matches can also be late, at least in pop-culture: movies are full of boy-man characters who fear the long-term commitment of a marriage-match. If they are one of the first to match, then, relative to their friends, they can no longer go to parties and fool around as their friends still can. If they delay for too long, then both the best women are taken and there aren't that many friends left to party with. Arguably here, if the peer pressure to stay unattached is large, then we observe a late rush where people delay excessively long.

## Conclusion

"There is a general rush for the new found *Dorado*."
— Morning Courier & New-York Enquirer (December 11, 1848)

We have introduced and explored a tractable class of games that generates endogenously timed rushes of variable size. While we argue that it succinctly subsumes a wide array of economic rushes, it also precludes some. For instance, we assume that it is the quantile order per se that matters. While this includes traffic examples like arrival at a parking structure that often hits capacity, it excludes an analysis of rush hour proper, where the intertemporal density of cars matters.[21]

We will now discuss some prominent rushes that have been observed over time and discuss how they may or may not fit into our framework.

GOLD AND MINERAL RUSHES. We have shied away from gold rushes and other mineral rushes for unrelated reasons. As Wikipedia notes, "Early gold rushers in California of 1848 got the biggest prizes: Some of these 'forty-eighters,' as these very earliest gold-seekers were also sometimes called, were able to collect large amounts of easily accessible gold—in some cases, thousands of dollars worth each day." Of course, some later movers became rich by selling shovels and other equipment, but it is difficult to make firm statements about the ex ante payoff structure that motivated the timing choices based on hindsight.

LAND RUSHES, AND OTHER ASSET GRABS. In a land rush people are allowed to stake a claim on a piece of land in a pre-defined area beginning at a pre-specified time; most famously this occurred in the 1889 Oklahoma Land Rush,[22] and in other rushes in Oklahoma in 1891, 1892, 1893, and 1895. While the beginning of the rush was officially announced (it would be the harvest time in our framework) by a cannon's shot, many

[21] Additionally, for rush hour, any rank rewards would be U-shaped.

[22] See Bohanon and Coelho (1998) for a brief historical account or the movie "Far and Away" (1992) for a cineastic depiction.

people tried to occupy the land in the days before the rush. These so-called "Sooners" would sneak onto the land and stake a claim before the rush started. This was not without risk for the army would patrol the area and remove 'sooners' from the land. So the choice was: should one be a sooner or a boomer (those that follow the cannon's 'boom!')? And if one is a sooner — when's the best time to tiptoe onto the land? Only if the punishment for 'sooning' is very high, then the payoff structure would satisfy the property of greediness so that all delay until the harvest (cannon-shot) time.

# A Appendix: Omitted Proofs

## A.1 Expected Quantile Payoffs

Function $v(x)$ denotes the payoff that a player obtains if a fraction $x$ of players has stopped. When players play a symmetric mixed strategy $Q(t)$, we claim that when $Q(t) = q$ the expected payoff is $v(q)$. To see this suppose first that there are $N+1$ players and that a player obtains payoff $v(k/N)$ when $k \in [0, N]$ players have stopped. Then the expected payoff is

$$\mathsf{E}[v|q] = \sum_{k=0}^{N} \binom{N}{k} q^k (1-q)^{N-k} v(k/N). \tag{4}$$

By Weierstrass' approximation theorem, the continuous function on $v : [0,1] \to \mathbb{R}_+$ can be approximated by a polynomial. Moreover, the approximating polynomial can be expressed using the Bernstein polynomials because these form a basis for the polynomials (see, for instance, Milovanovic, Mitrinovic, and Rassias (1994)). Formally, for any $\epsilon$ there is an $N$ such that

$$|v(q) - \sum_{k=0}^{N} \binom{N}{k} q^k (1-q)^{N-k} v(k/N)| < \epsilon. \tag{5}$$

Consequently, $\mathsf{E}[v|q] \to_{N\to\infty} v(q)$.

## A.2 Explicit details for the Running Example

The running example has functional form $v(x) = -(x-\theta)^2 + (\theta - 1/12)^2 + 13/12$. This function is a parabola with maximum rank $\theta$ that integrates to $1$ on $[0,1]$, and it is log-supermodular because $(\partial^2 \ln(v^\theta(x))/(\partial x \partial\theta) \geq 0$. We can further compute payoffs from initial and terminal rushes

$$V_0(q) = -\frac{1}{3}\left(q - \frac{3\theta}{2}\right)^2 + \frac{3}{4}\left(\theta - \frac{2}{3}\right)^2 + 1, \;\; V_1(q) = -\frac{1}{3}\left(q - \frac{3\theta+1}{2}\right)^2 + \frac{3}{4}\left(\theta - \frac{1}{3}\right)^2 + 1.$$

It is straightforward to check that $v(q) = V_0(q)$ for $q = 3\theta/2$ and $v(q) = V_1(q)$ for $q = (3\theta - 1)/2$. Solving the differential equation (2) explicitly we find that for mixed strategy play in the war of attrition and pre-emption game respectively is

$$Q(t) = \theta - \sqrt{\theta^2 + \left(\frac{4}{3} - \theta\right)\left(1 - \frac{e^{r(t+1)-1}}{(t+1)r}\right)}, \quad Q(t) = \theta + \sqrt{\theta^2 + \left(\frac{4}{3} - \theta\right) - \left(\frac{1}{3} + \theta\right)\frac{e^{r(t+1)-1}}{(t+1)r}}.$$

Figures 2 and 3 then illustrate the comparative static for the running example.

## A.3 Comparative Statics on Rushes: Proof of Proposition 1

Since monotone ratio domination applies, $v^{\theta'}(q) = h(q)v^{\theta}(q)$ for some increasing function $h$. To show that $R$ in *RP* increases, it suffices to show that at $q^{\theta} = \arg\max V_0^{\theta}(q)$, $V_0^{\theta'}$ increases. Dropping superscript $\theta$ on $q$, this is straightforward because

$$\begin{aligned} q^2 \cdot (V_0^{\theta'}(q))' &= qv^{\theta'}(q) - qV_0^{\theta'}(q) = v^{\theta}(q)h(q)q - \int_0^q v^{\theta}(x)h(x)\,dx \\ &= \int_0^q [h(q)v^{\theta}(q) - v^{\theta}(x)h(x)]\,dx \\ &= h(q)\underbrace{\int_0^q [v^{\theta}(q) - v^{\theta}(x)]\,dx}_{=0 \text{ at } q=q^{\theta}} + \int_0^q v^{\theta}(x)\underbrace{[h(q) - h(x)]}_{>0}\,dx > 0. \end{aligned}$$

The first term in the last line of the above is zero because at $q^{\theta}$, $v^{\theta} = V_0^{\theta}$, the second term is positive because $h$ is increasing. Thus since $V_0^{\theta'}$ increases at $q^{\theta}$, it peaks for a larger value of $q$. A similar argument applies to $\arg\max V_1^{\theta}$.

Next, to show that the gradual entry game is played for longer, it suffices to show that the exit rate $\dot{Q}^{\theta'}$ is lower than $\dot{Q}^{\theta}$. Consider the case of a $W$ phase. Then $\dot{\pi} < 0$ for all $t > t_{\pi}$, so that $-\dot{\pi}/\pi > 0$. Thus

$$\dot{Q}^{\theta} = -\frac{\dot{\pi}}{\pi}\frac{v_{\theta}(q)}{v'_{\theta}(q)} > -\frac{\dot{\pi}}{\pi}\frac{v_{\theta'}(q)}{v'_{\theta'}(q)} = \dot{Q}^{\theta'} \quad \Leftrightarrow \quad \frac{v_{\theta}(q)}{v'_{\theta}(q)} > \frac{v_{\theta'}(q)}{v'_{\theta'}(q)}. \tag{6}$$

Using that $v^{\theta'}(q) = v^{\theta}(q)h(q)$, we can simplify and rearrange

$$\frac{v^{\theta}(q)}{(v^{\theta}(q))'} > \frac{v^{\theta}(q)h(q)}{(v^{\theta}(q))'h(q) + h'(q)v^{\theta}(q)} \quad \Leftrightarrow \quad (v^{\theta}(q))^2 h'(q) > 0.$$

Then at any $t > t_{\pi}$, $Q^{\theta}(t) > Q^{\theta'}(t)$, thus the war of attrition lasts longer with $v^{\theta'}$. A symmetric argument shows that a $P$ phases in *RP* would last shorter. □

## A.4 Exit Rates: Proof of Proposition 2

Observe that

$$\ddot{Q} = -\frac{\ddot{\pi}\pi - \dot{\pi}^2}{\pi^2}\frac{v}{v'} - \frac{\dot{\pi}}{\pi}\,\dot{Q}\left(1 - \frac{v''v}{v'^2}\right) = \frac{v}{v'}\left(\frac{\dot{\pi}^2 - \ddot{\pi}\pi}{\pi^2} + \left(\frac{\dot{\pi}}{\pi}\right)^2\left(1 - \frac{v''v}{v'^2}\right)\right).$$

Since $v$ is concave, $v'' < 0$. In a war of attrition-phase, $v' > 0$, so that the last term is larger 1. Thus if $-\frac{\ddot{\pi}\pi - \dot{\pi}^2}{\pi^2} + \left(\frac{\dot{\pi}}{\pi}\right)^2 > 0$, then $\ddot{Q} > 0$. But this holds by log-concavity of $\pi$. Similarly, for a slow pre-emption game-phase, $v' < 0$, and thus the sign of $\ddot{Q}$ is reversed.

# References


ABREU, D., AND M. BRUNNERMEIER (2003): "Bubbles and Crashes," *Econometrica*, 71(1), 173–204.

ALLEN, F. (2001): "Do Financial Institutions Matter?," *Journal of Finance*, 56(4), 1165–1175.

ALLEN, F., AND D. GALE (1998): "Optimal Financial Crises," *Journal of Finance*, 53(4), 1245–1284.

ARGENZIANO, R., AND P. SCHMIDT-DENGLER (2008): "N-Player Preemption Games," mineo, University of Essex and LSE.

AVERY, C., C. JOLLS, R. A. POSNER, AND A. E. ROTH (2007): "The New Market for Federal Judicial Law Clerks," *The University of Chicago Law Review*, 74(2), 447–486.

BERK, J. B., AND R. C. GREEN (2004): "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112, 1269–1295.

BLANCHARD, O. J., AND M. W. WATSON (1982): "Bubbles, Rational Expectations and Financial Markets," *Crises in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*, p. Lexington.

BOHANON, C. E., AND P. R. P. COELHO (1998): "The Costs of Free Land: The Oklahoma Land Rushes," *The Journal of Real Estate Finance and Economics*, 16(2), 205–221.

BRUNNERMEIER, M., AND J. MORGAN (2006): "Clock Games: Theory and Experiments," mimeo, Princeton University.

BRUNNERMEIER, M., S. NAGEL, AND L. PETERSEN (2008): "Carry Trades and Currency Crashes," Working paper, Stanford University.

BRUNNERMEIER, M. K., AND S. NAGEL (2004): "Hedge Funds and the Technology Bubble," *The Journal of Finance*, 59(5), 2013–2040.

CHAMLEY, C., AND D. GALE (1994): "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, 62, 1065–1085.

CHOI, J. P. (1997): "Herd Behavior, the "Penguin Effect," and the Supression of Informational Diffusion: an Analysis of Informational Externalities and Payoff Interdependencies," *RAND Journal of Economics*, pp. 407–425.

DASGUPTA, A., AND A. PRAT (2007): "Information Aggregation in Financial Markets with Career Concerns," Working paper, London School of Economics.

DIAMOND, D. W., AND P. H. DYBVIG (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91(3), 401–419.

FARRELL, J., AND G. SALONER (1986): "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *American Economic Review*, 76(5), 940–955.

FUDENBERG, D., R. GILBERT, J. STIGLITZ, AND J. TIROLE (1983): "Preemption, Leapfrogging, and Competition in Patent Races," *European Economic Review*, 22, 3–31.

FUDENBERG, D., AND J. TIROLE (1985): "Preemption and Rent Equalization in the Adoption of New Technology," *Review of Economic Studies*, 52, 383–402.

——— (1986): "A Theory of Exit in Duopoly," *Econometrica*, 54, 943–960.

GRIFFIN, J., J. HARRIS, AND S. TOPALOGLU (2006): "Who drove and burst the tech bubble?," Working paper, Queens University.

GUL, F., AND R. LUNDHOLM (1995): "Endogenous Timing and the Clustering of Agents' Decisions," *Journal of Political Economy*, 103(5), 1039–1066.

HOPENHAYN, H., AND F. SQUINTANI (2006): "Preemption Games with Private Information," mimeo, UCLA.

HOTELLING, H. (1931): "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39(2), 137–175.

LEE, C. M. C., AND M. J. READY (1991): "Inferring Trade Direction from Intraday Data," *The Journal of Finance*, 46(2), 733–746.

LEVIN, D., AND J. PECK (2003): "To Grab for the Market or to Bide One's Time: A Dynamic Model of Entry," *RAND Journal of Economics*, 34, 536–556.

LJUNGQVIST, A., AND W. J. WILHELM (2003): "IPO pricing in the dot-com bubble," *Journal of Finance*, 58, 723–752.

MCKINNEY, C. N., M. NIEDERLE, AND A. E. ROTH (2005): "The collapse of a medical labor clearinghouse (and why such failures are rare)," *American Economic Review*, 95(3), 878–889.

MILOVANOVIC, G., D. MITRINOVIC, AND T. RASSIAS (1994): *Topics in Polynomials: Extremal Problems, Inequalities, Zeros*. World Scientific.

MONGELL, S., AND A. E. ROTH (1991): "Sorority Rush as a Two-Sided Matching Mechanism," *The American Economic Review*, 81(3), 441–464.

MURTO, P., AND J. VÄLIMÄKI (2006): "Learning in a model with exit," Working paper, Helsinki School of Economics.

NIEDERLE, M., D. D. PROCTOR, AND A. E. ROTH (2006): "What Will Be Needed for the New Gastroenterology Fellowship Match to Succeed?," *Gastroenterology*, 130(1), 218–224.

NIEDERLE, M., AND A. E. ROTH (2003): "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of Political Economy*, 111(6), 1342–1352.

NIELSEN, K. (2003): "Institutional Investors and the Market for Corporate Equity," working paper, University of Copenhagen.

PARK, A., AND L. SMITH (2008): "Caller Number Five and other timing games," *Theoretical Economics*, 3(2), 231256.

PASTOR, L., AND P. VERONESI (2006): "Was there a Nasdaq bubble in the late 1990s?," *Journal of Financial Economics*, 81(1), 61–100.

REINGANUM, J. (1981a): "Market Structure and the Diffusion of New Technology," *Bell Journal of Economics*, 12, 618–624.

——— (1981b): "On the Diffusion of a New Technology: A Game-Theoretic Approach," *Review of Economic Studies*, 43, 395–406.

ROTH, A. E., AND X. XING (1994): "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84(4), 992–1044.

SHLEIFER, A., AND R. W. VISHNY (1997): "The Limits of Arbitrage," *Journal of Finance*, 52(1), 35–55.

SIEGEL, R. (2007): "All-Pay Contests," mimeo, Stanford GSB.

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