Market Valuation and Merger Waves

MATTHEW RHODES-KROPF and S. VISWANATHAN*

ABSTRACT

Does valuation affect mergers? Data suggest that periods of stock merger activity are correlated with high market valuations. The naïve explanation that overvalued bidders wish to use stock is incomplete because targets should not be eager to accept stock. However, we show that potential market value deviations from fundamental values on both sides of the transaction can rationally lead to a correlation between stock merger activity and market valuation. Merger waves and waves of cash and stock purchases can be rationally driven by periods of over- and undervaluation of the stock market. Thus, valuation fundamentally impacts mergers.

ONE OF THE PUZZLES IN FINANCE is why there are periods when mergers are plentiful and other periods when merger activity is much lower. For example, in the period 1963–1964 there were 3,311 total acquisition announcements, while in 1968–1969 there were 10,569 acquisition announcements. Similarly in both the period from 1979 to 1980 and from 1990 to 1991 there were approximately 4,000 acquisition announcements while the late 1980s and late 1990s were much more active, with 9,278 announcements in 1999 alone (see Mergerstat Review 2001). These periods of high activity seem to be correlated with high market valuations, as shown by Maksimovic and Phillips (2001) and Jovanovic and Rousseau (2001). For example, 1998–2000 saw over $1.5 trillion in announced deals per year while 2001, after the market correction, saw half as much. Furthermore, casual observation suggests that firms tend to use stock in these high activity/high stock market periods as an "acquisition currency." In 1990 the percentage of stock as a fraction of total deal value was only 24%, while by 1998 the use of stock peaked at 68% of total deal value![1] Further, Martin (1996) shows that firms that use stock in acquisitions have lower book-to-market ratios than those that use cash. Stock deals were especially common in the high-flying high-technology sector where most takeovers involved securities. From 1996 to 2000 the computer software, supplies, and services industry

* Matthew Rhodes-Kropf and S. Viswanathan are with Columbia University and Duke University, respectively. We thank Michael Bradley, Espen Eckbo, Larry Glosten, Pete Kyle, Augustin Landier, Costis Maglaras, Jan Mahrt-Smith, Mit Mehta, Michael Riordan, David T. Robinson, Tano Santos, and Daniel Wolfenzon for useful discussions. The anonymous referee's comments have improved the paper considerably. We thank participants at the American Finance Association, Columbia University, Duke University, London Business School, National Bureau of Economic Research, Oxford University, Rutgers University, and Stanford University workshops for insightful comments. All errors are our own.

[1] Source: JP Morgan M&A Research, Thomson Financial Securities Data Company Inc. based on announced transactions.

group accounted for 16.5% of all transactions, and ex post this industry is widely regarded as having been overvalued (see Mergerstat Review 2002). The classic example is America Online's acquisition of Time Warner in a stock-for-stock deal. While the market value of American Online fell on announcement of the merger, the general view today is that America Online got an excellent deal as its stock was overvalued (ex post the stock has fallen from $73.75 the day before the announcement to $27.28 on March 11, 2002). The technology boom also saw companies like Cisco use stock aggressively as a way to undertake mergers.

The inability of financial theory to explain merger waves is noted by Brealey and Myers (1996) in their classic textbook *Principles of Corporate Finance*. In their concluding chapter, "What We Do and Do Not Know about Finance," they pose the question, "How can we explain merger waves?" and they cite the need for "better theories to help explain these bubbles of financial activity" (p. 997). We propose that private information on both sides can lead rationally to increased stock merger activity that is correlated with market valuation.

Mergers involving securities are inherently different from cash takeovers as they involve a valuation problem. The target is offered shares in the bidding firm at some exchange ratio. Since the target firm receives shares, they are concerned about whether the valuation of these bidder shares is appropriate. Furthermore, the valuation of bidder shares often changes in response to the announcement of the takeover itself. Valuation of bids is sometimes so contentious that courts are used to help determine if the highest bid was accepted.[2] Clearly, valuation is of great practical concern in takeovers and it is difficult to determine the true value of an offer. We build on the idea that targets attempt to value offers with limited information.

Our approach is based on a rational model of stock mergers. In our model, managers of bidding firms have private information about the stand-alone value of their firms and the potential value of merging with a target firm. Managers of targets have private information about the stand-alone value of their company. Both bidders and targets have market values that may not reflect the true value of their companies. Furthermore, the possible misvaluations have two components—a firm-specific component and a market-wide component. One might think of these as mispriced factors shared by the bidders and targets and mispriced factors that are not shared. In equilibrium, stock bids reflect the expected level of synergies between the firms. However, the target has limited information about the components of the misvaluation, and therefore has difficulty assessing the synergies.

The target management, on observing the bidders' fractional offers, decides whether to accept or reject a bid. The rational target knows whether their own firm is overvalued or undervalued, so they are not easily fooled, but they cannot determine whether this misvaluation is a market effect, a sector effect, or a

[2] See Kaplan (1993) for a discussion of the Paramount-Viacom merger where valuation issues were important. Hietala, Kaplan, and Robinson (2003) also analyze the Paramount–Viacom merger in detail.

firm effect. Fiduciary responsibility requires the target management to accept any offer that, given management's information, yields more than the stand-alone value. Hence, target management's decision is based on its assessment of synergies given the bids and management's own private information. A positive assessment of the offered synergies results in acceptance of the bid (and vice versa).[3] It is this assessment of synergies that is critical in our model.

Since the target's information and the bidder's bid are both positively related to the market-wide component of the misvaluation, the target attempts to filter out the market- (or sector-) wide misvaluation effect. The target correctly adjusts the bids for potential market overvaluation, but being a Bayesian updater, he puts some weight on high synergies as well. When the market-wide overvaluation is high, the estimation error associated with the synergy is high too, so the offer is more likely to be accepted. Thus, when the market is overvalued, then the target is more likely to overestimate the synergies *even though he can see that his own price is affected by the same overvaluation* because he still underestimates the shared component of the misvaluation.

This is because an overvalued target expects that some of this overvaluation is due to a market-wide effect and some is due to a firm-specific effect. The rational target is, of course, right on average. However, the more the market is overvalued, the larger is the target's expectation of his firm-specific misvaluation because he cannot tell which effect is causing his own misvaluation. Therefore, the target filters out of the bids too little of the market-wide effect when the market is overvalued and too much when then market is undervalued (as mentioned, getting it correct on average). Therefore, the bids tend to look better when the market is overvalued. The target is not irrational; he can simply do no better given his information.

The opposite effect occurs when the target is overvalued because of firm-specific reasons. Since the target cannot tell the difference between firm-specific and market-wide effects, the target expects that some of this overvaluation is due to each effect. Therefore, the more overvalued he is, the larger is his expectation of the market-wide effect. Thus, when the target is relatively overvalued, he filters too great a market-wide effect out of the bids, making the bids look low.

For example, a manager of a software firm has information about expected software sales, but he does not know if his information relates only to his software, or to software sales in general, or to most goods in the economy. Furthermore, at the peak of an expansion, the software manager may get signals of slowing software demand before market participants. Managers will see the direct impact on their own firms, but since they do not possess the information of all of the other firms, they will not be able to conclude that a recession is imminent. They are unable to distinguish what part of their information is shared or specific only to them. Therefore, around those times when an actual

[3] Support for positive synergies has been found by every major study. See Andrade, Mitchell, and Stafford (2001) for a summary of the literature. Recently, Hou, Olsson, and Robinson (2000) find that synergies are positive using a long-run approach.

recession is approaching, targets will be more likely to accept offers from overvalued bidders. This effect occurs because macroeconomic information is first held in pieces by individual participants in the economy.

In general, if misvaluation may be related across firms, then the target's assessment of synergies will be positively related to the bid but negatively related to his own reservation value (the target's true stand-alone value). Thus, the target perceives a bid to be high if the synergies are high, the bidding firm's stock is overvalued, or the target is *relatively* undervalued. Therefore, our theory predicts that mergers are more likely to occur in overvalued markets or sectors, and relatively undervalued targets are more likely to sell.

We have not assumed that synergies are higher in boom times. Nor have we assumed that some managers are willing to sell their firm for less than it is worth. Nor is it the case that some managers have limited rationality. Instead there is a simple explanation: The target is concerned that instead of synergies, bidders have overvalued stock or that the target is relatively undervalued. Thus, the target uses all available information to get an expectation about the offered value. The target is on average correct and thus increases his firm's value by accepting those offers that exceed his reservation price in expectation. However, if the market or sector is overvalued, then the target is more likely to overestimate the synergies, and if only the target is overvalued, then he is more likely to underestimate the synergies.

Thus, our theory is a Myers and Majluf (1984) setup such that overvalued bidders make high stock bids. The stock merger market does not collapse because some bidders have positive synergies. In addition, the target (buyer of the stock) has some noisy information about the bidder's (who is selling stock) valuation. This leads to mistakes that are correlated with valuation.

Throughout this analysis we allow multiple bidders. Multiple bidders provide the target with more information about synergies. This is because the target to some extent can filter the common misvaluation of the bidders. We show that the target's assessment of any bid decreases if the bids of other bidders increase; that is, high bids by other bidders signal an increased likelihood of a high market-wide misvaluation. The key limitation of the information from other bids is that synergies can also be correlated across firms.

The possibility of correlated synergies also provides intuition about how merger waves can last in the face of a rational market. With each merger, the market updates prices, lowering the probability of a second merger. This effect could rapidly eliminate misvaluation and end waves. However, with each new merger, the market increases its expectation of the probability that the synergies of all firms are high. Thus, subsequent mergers lead to smaller price revisions. Therefore, a merger wave that occurs when markets become overvalued may not end until the market realizes the true value of the synergies of the early mergers. Thus, waves of financial activity will occur in overvalued markets and end with a market crash when participants learn information about synergies that leads them to question the gains from the entire sequence of mergers. Hence, it is not the case that mergers predict market crashes, but it is rather that market crashes are preceded by mergers.

We extend our model to consider the possibility of cash bids. Cash bids will not be affected by misvaluation. Thus, since management is more likely to overestimate the value of stock bids in overvalued markets, waves of activity will also coincide with a higher fraction of deals completed using stock. This result is consistent with Andrade et al. (2001), who find a much larger positive announcement effect on the target for cash offers and a less negative effect on the acquirer. We suggest that this does not imply that cash mergers are better than stock mergers, but rather that cash mergers are more likely to occur in undervalued markets. So, the rational market updates and increases stock prices.

Our paper differs from the other approaches to merger waves. Jovanovic and Rousseau (2001, 2002) build on Gort (1969) and provide complete information models of merger waves that are based on technological change and the Q-theory. Mergers correspond to the purchase of used capital and merger waves occur when there is reallocation across sectors. Consequently, high Q-firms buy low Q-firms. We suggest it may also be that overvalued firms buy undervalued firms. Shleifer and Vishny (2003) provide a more behavioral story of merger waves in which it is common knowledge that the market is mispriced but will correct itself in the long run. When bidders are overvalued, merger waves with stock occur because some managers care only for the short-run market price (which does not adjust for the overvaluation of bidders) while others care for the long-run value (they are essentially issuing cheap equity to get something valuable). Gorton, Kahl, and Rosen (2000) suggest that mergers are a defensive mechanism by managers who do not wish to be taken over by others. Thus, inefficient merger waves occur when it looks like an efficient merger wave may be arriving shortly (such as just after a technological innovation). Holmstrom and Kaplan (2001) summarize the research on takeovers and argue that corporate governance issues led to the merger waves of the 1980s and 1990s. Toxvaerd (2002) proposes a theory of merger waves based on a dynamic preemption game. Persons and Warther (1997) provide a symmetric information model of financial innovation where the value of the innovations is positively correlated and there is learning. Hence, successful adoption by other agents increases one's estimate of the innovation, leading to more adoption; that is, clustering occurs. Further, a sequence of unsuccessful adoptions leads to a shutdown of the financial innovation. Although this is not a theory of merger waves, it is a theory of sporadic activity.

Our theory shows that merger waves can occur solely because of valuation issues. However, we want to emphasize that our theory does not imply that the desire to merge could not be caused by innovation, deregulation, or corporate governance issues, etc. Rather, we suggest that valuation impacts mergers and merger waves regardless of the underlying motivation for the mergers. Furthermore, we demonstrate why any merger may involve cash versus securities in a rational framework.

This paper is organized as follows. Section I contains the general model. Section II demonstrates the equilibrium and considers how firms bid, how the target chooses the winner, and the winner's payment. Section III examines mergers and the target's reservation price. Section IV shows how mergers can occur in waves. Section V explores the possibility that some bidders can bid with cash. Section VI concludes. In other work, Rhodes-Kropf, Robinson, and Viswanathan (2004) consider the empirical implications of these ideas.

I. The Model

The basic model of a merger is a second-price auction (the specifics of the auction are discussed in Sec. II). Intuitively, the competition among firms for an acquisition target is similar to the competition among bidders in an English (oral ascending bid) auction. However, an English auction is strategically equivalent to a second-price auction, and a second-price auction offers much greater tractability.[4] We assume n risk-neutral firms with synergistic values for a target firm bid in the auction (where $N = \{1, \ldots, n\}$ represents the set of n bidders). The risk-neutral target firm, firm T, considers the bids and decides whether to accept an offer. After the auction, the market reacts (as anticipated by the bidders and target). Then, in the last period, the value of all firms, including the joint firm (if the merger has occurred) is realized.

A bidding firm, firm i, has a private value V_i for firm T. This is the true value of firm T, X_T, multiplied by a factor that represents the synergy $(1 + s_i)$,

$$V_i = X_T(1 + s_i).$$

The synergy, $s_i > -1$, but may be positive or negative, so V_i may be greater than or less than the target's stand-alone value, X_T. Thus, merging the firms could add value (positive s_i) or destroy value (negative s_i). However, the bidding firm does not know the true value of the target, X_T or the synergy, s_i. Instead the firm only knows its own value for the target as a merger partner, V_i. All participants in the auction believe that the synergies and thus the merger values are independently and identically distributed and drawn from the distribution $F_s(s)$. Thus, this is an independent private value auction. However, this assumption does not preclude a known common component in the bidders' values. Therefore, we can think of the synergies as containing a common and firm-specific component, $(1 + s_i) = (1 + \lambda)(1 + \omega_i)$. Section III will elaborate on this idea.

The bidding firm also has private information about the value of their own assets, X, where firm i has the value $X_i > 0$.[5] The target and the market do not know X_i; however, both see the current market value of the firm, $M_i > 0$.

[4] It is well known that with cash bids, the oral ascending bid auction is strategically equivalent to the second-price auction in which the highest bidder wins and pays the bid of the second highest bidder. In our more complicated setting, where bidders bid a fraction of the joint firm that they will give to the owners of the target firm, similar reasoning will establish that an open ascending bid auction is strategically equivalent to an appropriately designed second-price auction in which the winning bidder must pay a fraction that depends only on the bids of the other bidders (see equation (2) and the paragraph that follows it). Our model employs the second-price auction because it is considerably more tractable and eases understanding. The insights and results of our paper hold true under the oral ascending auction as well.

[5] The variable X_i could alternatively represent the bidder's beliefs about their true value.

The market value of the firm, M_i does not necessarily equal X_i because it is possible that the market has misvalued the assets. We assume that there are two forms of misvaluation: market-wide (shared) misvaluation and firm-specific (not shared) misvaluation. Thus,

$$X_i = M_i(1 - \varepsilon_i)(1 - \rho),$$

where ρ represents the market-wide misvaluation and is the same for every firm, and ε_i is the firm-specific misvaluation.[6] Thus, if ρ or ε_i are positive, then the market is overvalued, and if they are negative then the market is undervalued. The firm-specific misvaluation $\varepsilon_i < 1$ is drawn i.i.d with zero mean from $F_\varepsilon(\varepsilon)$ and market-wide misvaluation ρ is drawn with zero mean from $F_\rho(\rho)$.[7,8] The variables ρ and ε_i are independent. Therefore, $E[X_i \mid M_i] = M_i$; that is, on average the market correctly prices the firms.

Target firms also have a market value, M_T. This market value is known to the bidding firms. The target, however, also knows the true stand-alone value of his assets, X_T. This value is different from the market value because of the same two forms of misvaluation that affect bidders: market-wide misvaluation, ρ, and firm-specific misvaluation, ε_T. Thus,

$$X_T = M_T(1 - \varepsilon_T)(1 - \rho),$$

where ρ is the same common component that affects the bidders, and ε_T is specific to the target. Since the target, firm T, has a stand-alone value, this value functions as a reserve price. Thus, the target may not accept any offer.

We can think of ρ and ε in different ways. We may think a manager has information about how well his own product will sell. The manager's superior information provides him an estimate of true value that differs from the market. However, the manager does not know how much of his information also relates to other firms. For example, when a manager of a shoe store chain determines that shoes will not sell well, he does not know if this relates only to shoes, or if soft goods in general will not sell well, or if most goods in the economy will not sell well. Managers understand how their information impacts their own firm, but managers are not macroeconomic experts and therefore do not understand what part of their information relates to other firms. In this case ρ can represent the information that relates to all firms in the economy, while ε relates only to a particular firm.

We can also think of ρ as mispriced factors that are shared by the target and bidders, and consider ε as mispriced factors that are not shared. For example, many factors may have been overpriced in the late 1990s, but it would seem

[6] Corollary 4 considers the idea that bidders come from the same sector but from a different sector than the target. In this case, $(1 - \varepsilon_i) = (1 - \psi_b)(1 - \phi_i)$, where ψ_b is the sector misvaluation shared by the bidders and ϕ_i is the completely firm-specific error.

[7] The assumption that the errors are drawn identically is not needed for the major results of the paper, but eases exposition.

[8] Samuelson (1987) suggests the inclusion of a firm-specific error in a comment on Hansen (1985), who has a model of mergers with no misvaluation.

that high-tech and internet factors were even more overpriced. Since some firms load more on some factors than others, firms have differential misvaluation.

Although each firm knows if they are undervalued or overvalued, they do not know ρ or ε_i. Bidders also have information through M_i, X_i, V_i, and M_T about the target's and every other bidder's true value, but they do not know the true value of any other player. Therefore, they cannot risklessly arbitrage a market-wide misvaluation by trading other firms' stock. We assume that some form of limited arbitrage allows equilibrium misvaluation. Management is simply not the marginal investor. This is true even in their own stock. We assume that adverse selection and SEC insider information rules prevent managers from buying and selling their own undervalued or overvalued stock in large enough quantities to restore efficient pricing.

Since the firms know V with certainty, they are not concerned with the target's superior information about X_T. If the true value of the synergy depended on the target's information, then the bidding firm would need to worry about whether the target accepts the offer; that is, the target will accept if the offer is too high. This type of adverse selection is the focus of interesting papers by Fishman (1989), Eckbo, Giammarino, and Heinkel (1990), and Hansen (1987), but is not considered here. Berkovitch and Narayanan (1990) also assume that only acquirers have private information about the synergy, although misvaluation plays no role in their theory.

To begin we assume that all firms must bid using only their own equity.[9] However, in Section V we consider the possibility of cash bids. The assumption of limited cash is justifiable if raising cash is costly. We assume that since firms are misvalued, the lemons problem prevents firms from selling their own stock for cash. This same problem does not collapse the merger market because some firms have large synergies and the target firm has information superior to the market. Section V considers cash bids, and shows why there may be waves of stock or cash mergers.

An equity bid consists of an offer of fraction α_i of the joint firm. After the auction, if the bid is accepted, the total firm value will be $X_i + V_i$. However, since the target and the market know neither X_i nor V_i, they will rationally value any offer at $E[\alpha_i(X_i + V_i) \mid \Phi_T]$ and $E[\alpha_i(X_i + V_i) \mid \Phi_M]$, where Φ_T and Φ_M represent the target's and market's (differing) information set. The contents of this information set will be considered further below. Throughout the paper, the players will rationally Bayesian update the value of any offer and the market will update and condition values on all available information. Thus, even though bids may be misvalued, in expectation all prices will reflect expected values. Figure 1 provides an overview of the model.

After the target sees the bids, the target must decide which bid to accept, if any. We assume that fiduciary responsibility rules require the target to accept

[9] Rhodes-Kropf and Viswanathan (2000) formalize many of the problems of bidding with securities, and Rhodes-Kropf and Viswanathan (2002) extend this idea to auctions financed by securities markets. The problems raised in this work do not exist here because bidders have valuable assets.

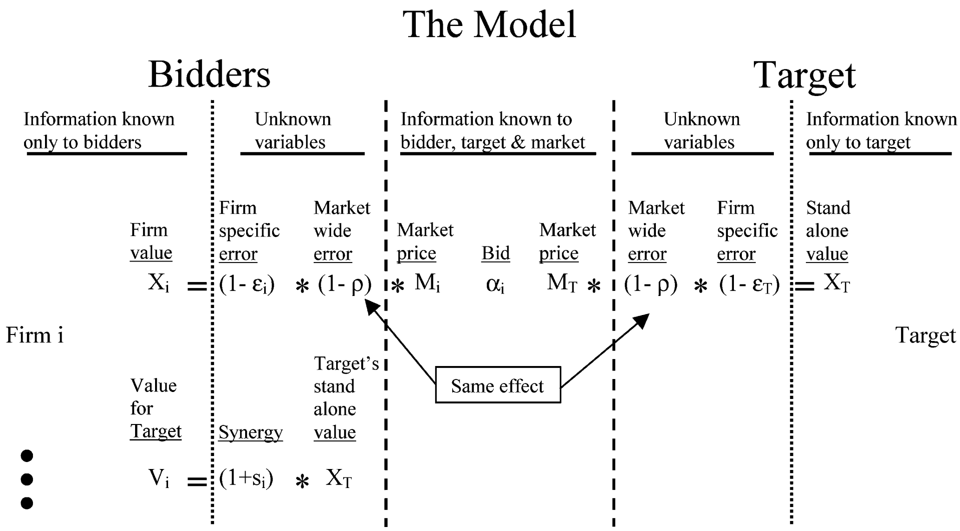

Figure 1. Model with information partitions.

only the offer with the highest expected value.[10] Thus, the target must reject (accept) any offer with long-run expected value less (greater) than X_T. Thus, we are assuming that managers are long-run value maximizers.[11]

II. Equity Auction with Misvalued Stock

In a second-price auction with equity bids, firms bid by stating a fraction α of the joint firm that they will give to the owners of the target firm. The high bid is clearly not necessarily from the firm that states the highest fraction, α. A firm with very substantial assets (such as IBM) must bid a lower α than a low-asset firm (such as a local computer vendor), even if their values for firm T are the same. Furthermore, the target's incomplete information about the true value of the bidder's assets, X_i, means that the rule used to rank the bids will be a function of the bidder's stock market value, M_i, but not X_i. The target also has information about his own misvaluation and information from the other bids. Thus, the scoring rule is a function of the target's private information and all the firms' bids and stock prices,

$$Z_i \equiv g(\alpha_i, M_i, \Phi_T) \equiv g(\alpha_i, M_i, \alpha_j, M_j \forall j \neq i, M_T, X_T), \tag{1}$$

where the highest Z wins the auction. Since all errors are drawn identically, and all synergies are drawn identically, we assume that every bidder is scored in a symmetric fashion.

[10] This assumption rules out expected revenue-enhancing rules (when $F(\cdot)$ is not regular) that require the seller to commit to accepting lower valued offers to encourage better types to bid higher.

[11] In a previous version, we also considered short-run managers who accepted any offer with current market value above the current value of their stock, M_T. The addition of short-run managers adds noise, but does not result in waves.

In a second-price auction the winning firm pays a fraction $\widehat{\alpha}_1$ that equals the lowest fraction they could have bid and just tied the second highest bidder,

$$Z_2 \equiv g(\alpha_2, M_2, \hat{\alpha}_1, M_1, \Phi_{T-}) = g(\hat{\alpha}_1, M_1, \alpha_2, M_2, \Phi_{T-}), \tag{2}$$

where $\Phi_{T-} = \alpha_j, M_j \;\; \forall j \neq 1\, or\, 2, M_T, X_T$. Note that the highest bidder's true willingness to pay is not a part of equation (2). Thus, the fraction the winner must pay does not depend on what they are willing to bid (or the information revealed by what they are willing to bid), but rather it depends on what the second highest bidder bid. We can see the strategic equivalence to the oral ascending auction since equation (2) is the point at which the second highest bidder would have dropped out of the bidding in an ascending bid auction.

The remainder of this section will determine the equilibrium in the auction and the necessary assumptions. The reader who is less interested in the formation of the equilibrium is invited to skip to Section III.

To examine the equilibrium we will first consider the bids, then the probability of winning, then the reservation price, and then the expected payment.

A. Bids

To determine who wins the auction and what they pay, we must first determine how bidders choose to bid. When the target is not accurately informed about the asset value of the bidding firms, the natural inclination is to think that firms who are overvalued by the market bid a much larger fraction α than firms that are correctly valued. Surprisingly, this section will show that under a large class of reasonable scoring rules, this is not the case. In fact, all bidders will bid the true largest fraction that they would ever be willing to pay, $\frac{V_i}{X_i + V_i}$.

In order to determine the equilibrium bids, we must put some more structure on the scoring function. The following lemma shows the condition that is both necessary and sufficient for bidders to bid the truth, $\frac{V_i}{X_i + V_i}$.

LEMMA 1: *In a second-price auction, if $g(\cdot)$ is continuous in every bidder's $\alpha, g(0, M_i, \Phi_T) - g(\alpha, M_j, \Phi_T) < 0 \, \forall \, \alpha > 0$ and if whenever two bidders achieve the same score*

$$\begin{aligned} g(\alpha_i, M_i, \Phi_T) &\equiv g(\alpha_i, M_i, \alpha_j, M_j, \alpha_\kappa, M_\kappa \forall \kappa \neq j, M_T, X_T) \\ &= g(\alpha_j, M_j, \alpha_i, M_i, \alpha_\kappa, M_\kappa \forall \kappa \neq j, M_T, X_T) \\ &\equiv g(\alpha_j, M_j, \Phi_T), \end{aligned} \tag{3}$$

then

$$\frac{\partial g}{\partial \alpha_i}(\alpha_i, M_i, \Phi_T) > \frac{\partial g}{\partial \alpha_i}(\alpha_j, M_j, \Phi_T) \quad \forall \alpha, M, \tag{4}$$

then it is a dominant strategy for bidders to bid

$$\alpha_i = \frac{V_i}{X_i + V_i}. \tag{5}$$

Proof: See the Appendix.

The intuition is that raising the bid increases the probability that a firm wins. However, firms are not willing to bid so that if they win they pay more than V. And while they will happily pay less than V, their payment is determined by the second highest bidder. Therefore, they will not stop increasing their bid unless their payment would exceed V, *or* raising their bid lowers the chance that they win. It might seem that if we simply assumed $\frac{\partial g}{\partial \alpha} > 0$, then raising the bid would always increase the chance of winning (this is true if X_i is known). However, firms have to consider the possibility that increasing their bid affects the scores of other bidding firms by altering the target's information about the market-wide misvaluation. The only relevant consideration is between two potential firms with the same score. Continuity of the scoring function ensures that as α increases, bidders tie before they beat another bidder. Either tied firm only has the incentive to raise his bid as long as doing so increases his own score more than the other firm's score (or decreases his own score less than the other firm's score). If this is true, then firms will always gain by increasing the fraction they bid until they reach $\alpha_i = \frac{V_i}{X_i + V_i}$. To the reader unfamiliar with the second-price auction, it may seem odd that bidders bid the truth. Keep in mind that the winning bidder will only be charged the fraction that he would have had to say to just beat the second highest bidder. Thus, it is as though the winner is just beating the second highest bidder in an English auction.

Note that many odd scoring rules satisfy condition (4). Some targets may be tempted to rank bids as is often done in newspapers, where the target's and buyer's market values are added together and multiplied by α. This scoring rule satisfies condition (4) but it ignores a great deal of the target's information. The next section will focus on the scoring rule that chooses the highest bid.

B. Choosing the Winner

Thus, firms bid the truth even though the market has misvalued their assets, as long as the conditions in Lemma 1 hold. We now focus on how the target will choose the winner, that is, pick the equilibrium scoring rule.

Even though each firm bids $\alpha_i = \frac{V_i}{X_i + V_i}$, the target cannot determine V_i because the target does not know X_i. Therefore, the target must award the firm to the highest score, $Z_1 = g(\alpha_1, M_1, \Phi_T)$, which may not be the firm with the highest value. As we will see in a moment, this is true even when the target uses all available information.

Ex ante, the market's best estimate of the true asset value of a firm is the firm's market value. The target, however, has information about his own misvaluation. Since part of the misvaluation is the same for every firm, the target has a better estimate of X_i than M_i. Even before the auction, the target's estimate of any firm's value is $E[X_i \mid M_i, X_T, M_T]$. Thus, if the target is overvalued, he assumes (correctly on average) that part of this is due to a market-wide effect and part is due to a firm-specific effect.[12]

[12] Remember, we assumed some limits to arbitrage. Thus, managers do not have the ability to complete enough risky arbitrage trades to ensure that the market is correctly valued.

After the bids, the target updates his expectation of X_i. The target must decide the probability that the firm is overvalued versus the probability that the firm has a large synergy. For example, let bidder firm 1 have a market value of $M_1 = \$100$. If he bids $\alpha_1 = 20\%$, then it might be the case that he values firm T at least \$25. Or it might be the case that the true value of his assets is less than \$100. If his assets are worth $X_1 = \$80$, then he only needs to value firm T at \$20 in order to be willing to bid 20%. Thus, the question the target must ask is what is the probability that the bidder firm has a high synergy ($V_1 \geq \$25$) or that the bidder firm is overvalued ($M_1 > X_1$)? If the probability of a high V_1 is low, then it is more likely that $M_1 > X_1$, and the target will revise down his expectation of X_1, and therefore he will also expect lower synergies.

Fiduciary responsibility requires the target to accept only the highest offer. This assumption tells us that the largest score should be assigned to the offer with the highest expected value. Therefore, the only equilibrium scoring rule is any monotonic transformation of

$$Z_i = g(\alpha_i, M_i, \Phi_T) = E\left[\alpha_i * (V_i + X_i) \mid \alpha_j, M_j \; \forall j, X_T, M_T\right]. \tag{6}$$

It may seem that the target should not be concerned with the true value of the offer as equation (6) suggests, but rather the target should consider only the offer with the highest stock value. With Bayesian updating, we will show in a moment that these two scoring rules yield the same outcome, that is, they each rank the same bid as the highest bid.

Lemma 1 tells us that as long as condition (4) holds, then the firms will bid the truth, $\alpha_i = \frac{V_i}{X_i+V_i}$. If the firms bid the truth then the rational updating scoring rule, equation (6), becomes

$$Z_i = g_i(\alpha_i, \Phi_T) = E\left[\alpha_i \left(V_i + \frac{1-\alpha_i}{\alpha_i} V_i\right) \mid \alpha_i, \Phi_T\right] = E[V_i \mid \alpha_i, \Phi_T].$$

Therefore, since bidders are bidding the truth, the target is attempting to choose the firm with the highest positive synergy value. It would seem that all of the information that improves the accuracy of the expectation of the synergy improves the scoring rule. However, the following lemma shows that this is not the case, and $\frac{\alpha_i}{1-\alpha_i} M_i$ is sufficient to rank the bids.

LEMMA 2: *If the conditions in Lemma 1 hold for scoring rule (6), then*

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] = E\left[V_i \,\middle|\, \frac{\alpha_i}{1-\alpha_i} M_i \; \forall i, X_T, M_T\right] \tag{7}$$

and $\frac{\alpha_i}{1-\alpha_i} M_i$ *is sufficient information to rank the bids.*

Proof: See the Appendix.

This tells us that for the expectation of V_i, $\frac{\alpha_i}{1-\alpha_i} M_i$ is a sufficient statistic for α_i and M_i. Although M_i and α_i do not add information above $\frac{\alpha_i}{1-\alpha_i} M_i$, this does not mean that the best estimate of V_i is $\frac{\alpha_i}{1-\alpha_i} M_i$; it is not. However, the target cannot tell the difference between a low V_i with an overvalued stock (positive ε or ρ) and a high V_i with an undervalued stock. So, the target gives the same score to each.

Lemma 2 tells us that any scoring rule needs only to be a monotonic transformation of $\frac{\alpha}{1-\alpha}M$. Although this may not seem intuitively obvious, it is easy to understand why this works. Without misvalued stock, the target would be able to award the firm to the highest offer. However, with misvaluation, the target may not be able to tell the difference between a firm with overvalued stock and a firm with truly high synergies, but an undervalued stock. With any offer by the firm, $\frac{\alpha}{1-\alpha}M$ equals the ratio of the offered synergies and the error in the stock price. Since everything is drawn i.i.d, this ratio provides all available information about the rank of the bid. Thus, the fact that $\frac{\alpha}{1-\alpha}M$ is a sufficient statistic just tells us formally that the target cannot tell the difference between a high bid and overvalued stock and instead he knows only the ratio of the value and the error.

Thus, without synergies, adverse selection would collapse the merger market. It may seem that condition (4) (which yields truthful bidding) is assuming away adverse selection; it is not. Adverse selection arises because at any given score, the target must compare the probability that the synergies are high with the probability that the errors are high. If the expected synergies at a given score are negative, then the market collapses as the target will not accept a bid with a negative value (the next section will further examine the target's decision). Condition (4) says that on the margin, raising α increases the expectation of V_i (even if the increase is only to a less negative number). Subsection II.D will show that condition (4) is similar to but weaker than the standard assumption in auction theory of affiliation.

We asked earlier why the target cares about the true value of the offer. If the target can immediately sell their stock, then shouldn't the target accept the highest stock offer rather than worry about the true value? Yes. However, the following corollary tells us that the market's ranking of the bids is the same as the target's management.

COROLLARY 1: *The market's ranking of the bids is the same as the target's ranking of the bids.*

Proof: This is a direct consequence of Lemma 2, which showed that $\frac{\alpha_i}{1-\alpha_i}M_i$ is sufficient information to rank the bids. This conclusion from Lemma 2 is still true when the market evaluates the bids, except that the market's scoring rule is

$$= E\left[V_i \,\middle|\, \frac{\alpha_i}{1-\alpha_i}M_i \;\forall i, M_T\right]. \tag{8}$$

Q.E.D.

We will see in a moment that the target's greater information does allow him to get a more accurate expectation of the value of the bid, but Corollary 1 shows that it does not affect the order of the bids.

After the target ranks the bids, he must choose whether to accept any of the offers. We assume that the seller commits to a reservation price R that equals the lowest expected value the seller would be willing to accept. Therefore, a merger will only occur if a firm bids such that their score is greater than R. We assume that if only one bidder bids above R, then a merger still occurs, and the target charges the bidder the fraction they just would have had to bid to achieve a score of R. This is essentially the assumption that when there is only one bidder, then all of the bargaining power resides with the bidder. Our results are not qualitatively affected by any assumption such that the probability of a merger occurring is increasing if the reservation price decreases or the bidder's willingness to pay increases.

This section has shown that the only relevant information available to the market is $\frac{\alpha_i}{1-\alpha_i} M_i \forall i$, and the only additional useful information held by the target is X_T. This information will help us determine what the bidder must actually pay and it will help us show that the scoring rule does satisfy condition (4) as we assumed.

C. The Firm's Payment

The Z-score tells us which firm will win the auction, but it does not tell us what they pay. If at least two bids are above the target's reservation price, then the winner must pay a fraction $\hat{\alpha}$ that equals the lowest fraction they could have bid and just tied the second highest bidder,

$$Z_2 \equiv g(\alpha_2, M_2, \hat{\alpha}_1, M_1, \Phi_{T-}) = g(\hat{\alpha}_1, M_1, \alpha_2, M_2, \Phi_{T-}), \tag{9}$$

where $\Phi_{T-} = \alpha_j, M_j \;\; \forall j \neq 1 \, or \, 2, M_T, X_T$ and the subscripts 1 and 2 represent the bidders who received the highest and second highest Z-scores. This results in a simple definition of $\hat{\alpha}_1$.

LEMMA 3: *If at least two bids are above the target's reservation price, then the winning bidder pays*

$$\hat{\alpha}_1 = \frac{V_2(1-\varepsilon_1)}{X_1(1-\varepsilon_2) + V_2(1-\varepsilon_1)}. \tag{10}$$

Proof: See the Appendix.

Therefore, the winning firm's payment decreases if it is overvalued and increases if the second highest bidder is overvalued. Furthermore, it is easy to show that firm-specific misvaluation lowers the target's expected revenue.[13]

[13] Ex ante the expectations of both errors are zero. However, given that bidder 1 wins the auction $E[\varepsilon_1 \mid g_1 > g_2] > E[\varepsilon_2 \mid g_1 > g_2]$. Therefore,

$$\frac{V_2(1-\varepsilon_1)}{X_1(1-\varepsilon_2) + V_2(1-\varepsilon_1)} < \frac{V_2}{X_1 + V_2},$$

and bidder 1's expected payment is lower than if there were no misvaluation.

Equation (10) makes the following corollary to Lemma 3 easy to understand.

COROLLARY 2: *Market-wide misvaluation does not affect the equilibrium fraction that any firm is willing to offer, and therefore does not alter which firm offers the highest bid nor the amount they pay.*

Proof: Lemma 1 showed that the bids could be ranked by $\frac{\alpha}{1-\alpha}M$. Substituting for M and α shows that $\frac{\alpha_i}{1-\alpha_i}M_i = \frac{X_T(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$. Thus, the score of every bid is affected by $(1-\rho)$ in the same way and the rank is preserved. Furthermore, Lemma 3 shows that the bid fraction required to make the highest offer is unaffected by $(1-\rho)$. Q.E.D.

Market-wide misvaluation has no effect on rank because offers are compared to each other. However, we will see that the market-wide misvaluation has a large effect on overall acquisition activity. This is because the target's evaluation of whether to accept *any* bid is affected by market-wide misvaluation.

If there is only one bid above the reservation price, R, then as noted above, we assume that a merger still occurs and the target charges the bidder the fraction they just would have had to bid to achieve a score of R. Thus, the target only accepts a bid that has an expected value above R and then charges a lower fraction $\hat{\alpha}_1$. Section III will demonstrate the impact of the target's decision rule.

Overall, possible firm-specific misvaluation alters the firms' payments and who wins, but for given stock valuations, market-wide misvaluation does not alter how a fully rational target ranks the bids. If every firm is currently overvalued or undervalued by some amount, then the fraction that any firm is willing to offer is unaffected. However, we will see in a moment that both types of misvaluation affect whether the target will accept any bid at all.

D. Affiliation

Lemma 1 demonstrated the weakest condition that ensures firms bid the truth. We have assumed thus far that the target's scoring rule, equation (6), satisfies this condition. For the remainder of the paper we make the following reasonable primitive assumption about the distributions of the random variables to ensure that condition (4) holds and bidders do indeed bid the truth.

ASSUMPTION 1: *The random variables* $\log(1-\rho), \log(1-\varepsilon_T), \log(1-\varepsilon_i)$, *and* $\log(1+s_i)$ *for all i have log-concave densities.*[14]

The usual assumption in auction theory is that bidder values are affiliated. Since the firm synergy values are independent, it would seem that affiliation

[14] We can make the assumptions on the log of these variables because each variable has a distribution over the positive real line.

is trivially satisfied. However, because of market-wide misvaluation, the sufficient statistics used by the target are not independent. In fact, the target learns something about the synergies from looking at his own misvaluation and from looking at all of the bids. Lemma 2 shows that if firms bid the truth, then in equilibrium

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] = X_T E\left[(1+s_i) \left| \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)} \forall i, (1-\varepsilon_T)(1-\rho)\right.\right] \quad (11)$$

or equivalently

$$= X_T E\left[(1+s_i) \left| \frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}, \frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)} \forall j \neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right.\right]. \quad (12)$$

However, as the following lemma shows, the log-concavity assumption ensures that the sufficient statistics are affiliated with $(1+s_i)$.

LEMMA 4: *Under Assumption 1 the random variables* $1+s_i$, $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$, $\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)} \forall j \neq i$, *and* $\frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$ *are affiliated.*

Proof: See the Appendix.

Affiliation essentially means that the expectation of $(1+s_i)$ increases with any of the sufficient statistics (see Milgrom and Weber (1982) for a formal definition of affiliation). Log-concavity is a standard although not completely trivial assumption. Examples of log-concave densities include the multivariate beta, Direchlet, exponential, gamma, Laplace, normal, uniform, Weibull, and Wishart distributions. For example, if $(1-\varepsilon_i)$, $(1+s_i)$, $(1-\varepsilon_T)$, and $(1-\rho)$ are distributed lognormally, then the log of each variable has a log-concave density.

Lemma 4 can be used to show that condition (4) holds, and will allow us to prove a number of interesting theorems.[15] Milgrom and Weber (1982) and others have shown that in general, little can be said without some form of the affiliation property.

[15] Affiliation ensures that

$$\frac{\partial g_i}{\partial \alpha_i}(\alpha_i, \Phi_T) > 0 \quad \forall \alpha, M,$$

since each statistic in the expectation of $(1+s_i)$ increases with α_i. Thus, $g_i(0, \Phi_T) - g_j(\alpha, \Phi_T) < 0 \; \forall \alpha > 0$. Affiliation also ensures that

$$\frac{\partial g_j}{\partial \alpha_i}(\alpha_j, \Phi_T) < 0 \quad \forall \alpha, M,$$

since the expectation of $(1+s_j)$ decreases if α_i increases. We see that this assumption is stronger than condition (4). The requirements of Lemma 1 are satisfied.

III. Mergers

We have now determined the equilibrium bid and the scoring rule. This section will show how the target's reservation price combined with misvaluation leads to increased merger activity in overvalued markets.

Since the target has a stand-alone value, X_T, the target is unwilling to (and has a fiduciary responsibility not to) accept any offer that delivers less than X_T. The scoring rule examined above demonstrated that the highest expected value offer is from the firm that bids the highest $\frac{\alpha}{1-\alpha}M$. However, the highest offer may have an expected value less than the target's stand-alone value. This can occur when expected misvaluations are large relative to expected synergies. In this case the target will refuse the offer and no merger will occur. Thus, X_T is the target's reservation price and the target's acceptance rule is simply to accept any offer such that[16,17]

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] > X_T. \tag{13}$$

This simple rule will cause merger waves, but we begin by focusing on a single merger.

THEOREM 1: *Stock mergers are more likely to occur in overvalued markets than in undervalued markets.*

Proof: Subsection II.D shows that Lemma 2 allows the target's acceptance rule, equation (13), to be rewritten as

$$E\left[(1+s_i)\,\middle|\, \frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}, \frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j \neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right] > 1. \tag{14}$$

The only term that depends on the market-wide misvaluation is $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$, which is increasing in ρ. Lemma 4 ensures that the expectation of the synergy is increasing in $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$. Therefore, the more overvalued the market (the larger ρ is), the more likely it is that a bid exceeds the reservation price, and thus a merger occurs (and vice versa for undervalued markets). Q.E.D.

It is not the case that synergies are higher in boom times. It is not the case that some managers are willing to sell their firms for less than they are worth. Nor is it the case that some managers have limited rationality. Instead there is a simple explanation: The target is concerned that any bidder has overvalued stock rather than a high synergy. Thus, the target uses all available information to get an expectation of the offered value. The target is on average correct

[16] An earlier version of this paper also considered a short-run manager who accepts offers only if $E[V_i \mid \alpha_i, M_i, \Phi_M] > M_T$. Short-run managers are just as likely to value an offer above the current market price in bad times as in good. Thus, the addition of short-run managers does not cause (nor is it necessary for) merger waves.

[17] Since this is a second-price auction, the target accepts any offer with a value above his reservation value (given the true bid) but will then *charge* them a fraction $\hat{\alpha}$ that is the smallest bid the bidder could have made and just been accepted (see Sec. II.C).

and thus increases his firm's value by accepting those offers that exceed his reservation price in expectation. However, if the market is overvalued, then the target is more likely to overestimate the synergies, *even though he can see that his own price is affected by the same overvaluation* because he still underestimates the market-wide misvaluation. The target makes the correct adjustment for potential market overvaluation once it receives a high stock offer, but being a Bayesian updater, it puts some weight on high synergies as well. When the market-wide overvaluation is high, the estimation error associated with the synergy is high too, so the offer is more likely to be accepted. Therefore, the target accepts more mergers in overvalued markets and accepts less in undervalued markets.

Fully rational firms simply make what turns out to be mistakes in evaluating offers. That is, their decision turns out to be wrong ex post even though it was correct ex ante. Although firms have an ability to understand the impact of their information on their own firms, they cannot determine what part of their information relates to all firms. For example, at the peak of an expansion, firms are likely to receive signals of slowing demand before market participants. Managers will see the direct impact on their own firms, but since they do not possess the information of all the other firms, they will not be able to conclude that a recession is imminent. Any time macroeconomic news is held in pieces by individual participants in the economy, those participants will be unable to determine if their information relates to everyone or only to them.

This does not imply that the target looses money by accepting an offer. In an overvalued market, the target can expect his own stock to fall. Thus, accepting a merger proposal with a positive synergy will reduce the impact on the target when the market corrects. We will also see in a moment that this does not imply that the market has an arbitrage opportunity; the prices will correctly react to news of a merger.

A. Who Merges?

Before we consider the market reaction, there are a few corollaries that elaborate on Theorem 1.

COROLLARY 3: *On average, overvalued firms or firms with large synergies win takeover battles and undervalued targets are purchased.*

Proof: Each term in equation (14) increases if firm i has a larger firm-specific misvaluation, ε_i, or if firm i has a larger synergy, s_i. Therefore, firms with greater firm-specific misvaluation or synergies are more likely to be over the reservation price. Furthermore, $\frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$ increases if the target specific error, ε_T, decreases. Therefore, targets who have smaller firm-specific misvaluation (ε_T is smaller) are more likely to accept an offer. Q.E.D.

This corollary elucidates the type of errors that are likely to occur. If the bidding firm has a large firm-specific overvaluation, then it is more likely to win, because the target cannot distinguish between a large synergy, s_i, and a large firm-specific error, ε_i. This is easily seen because each relevant statistic in equation (14) is a function of $\frac{(1+s_i)}{(1-\varepsilon_i)}$, so large s_i and large ε_i have the same effect on every statistic.

The smaller a target's firm-specific misvaluation, ε_T, the larger their estimate of every bidder's synergy. This is because the target only knows the total error, $(1-\rho)(1-\varepsilon_T)$. For a given total error, if ε_T is smaller, then ρ must be larger. The target knows that the larger the market-wide component, the greater all of the bids will look. So, the target would reduce the expected value of an offer more if he knew that the market-wide component was larger. However, the target does not know he has a small ε_T. Therefore, he underestimates ρ and does not reduce the expectation of the offers by as much as he should. His expectation is, therefore, more likely to be above the reservation price. In general, the smaller a target's ε_T, the more likely it is that a merger occurs.

If we allow for the possibility that bidders come from a different sector than the target, then bidders may share a component of misvaluation with each other, but not with the target. This simple extension will allow us to consider cross-sector versus within-sector mergers. We assume for a moment that $(1-\varepsilon_i) = (1-\psi_b)(1-\phi_i)$, where ψ_b is the sector misvaluation shared by the bidders and ϕ_i is the completely firm-specific error, and the target's error $(1-\varepsilon_T)$ may or may not share the sector misvaluation.

COROLLARY 4: *Within-sector stock mergers are more likely to occur in overvalued sectors than in undervalued sectors. Furthermore, on average, overvalued sectors will purchase firms in relatively undervalued sectors.*

Proof: See the Appendix.

This corollary provides an additional mechanism by which market/factor/sector overvaluation leads to mergers. If bidders are in an overvalued sector and targets are in an undervalued sector (or bidders load more on a factor that is overvalued), then targets will confuse high synergies with high sector valuation of bidders and accept mergers. This effect would exist even if target managers had no private information. This corollary may explain the purchase during the 1990s by internet or telecom firms of firms with hard assets. Such mergers included the acquisition of Frontier Telephone (a long-distance provider) by Global Crossing and the acquisition of Time Warner by AOL. This corollary also provides an explanation of diversifying mergers in which firms in a more overvalued industry buy firms in a less overvalued industry (the market realizes these differential overvaluations ex post).

B. The Effect of the Losing Bidders

The following corollary shows that sector overvaluation is less likely to confuse targets when there is more than one bidder from a sector.

COROLLARY 5: *The larger the bids of the losing bidders, the lower the probability of a merger occurring.*

Proof: In equation (14) the conditioning variables $\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)} \forall j \neq i$ all decrease if the bid of another firm $\frac{(1+s_j)}{(1-\varepsilon_j)(1-\rho)}$ increases. Affiliation ensures that this decreases the expectation of V_i and therefore decreases the probability of a merger. Q.E.D.

The bids of the losing firms are relevant to the target because they provide information about shared misvaluation. If all of the bids are high, then the target suspects that this is because all the bidders are overvalued, or the target is undervalued. Therefore, he lowers his estimate of the synergies from the winning firm. Thus, more competing firms provide more information and increase the accuracy of the target. However, the following corollary shows that when the synergies have a common component, $(1+s_i) = (1+\lambda)(1+\omega_i)$, then there is a limit to the information that can be learned from competing bids.

COROLLARY 6: *If the synergies have a common component, then the bids of the losing firms are less informative about the synergies.*

Proof: See the Appendix.

Although increased competition reduces the information asymmetry and therefore the effects of market-wide misvaluation, if the synergies have a common component, then there is a limit to the information that can be gleaned from the competing bids.

The intuition for Theorem 1 and Corollaries 3 to 6 is that although the target is rational and thus correct on average, the noise in the model leads to different types of mistakes by the target. The target sees $\frac{\alpha_i}{1-\alpha_i} M_i = \frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$ from the bidder. If either the market-wide or firm-specific error, ρ or ϵ_i, is larger, then $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$ is larger, but the target does not know if a larger $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$ is due to a larger synergy, $(1+s_i)$, or a larger error, ρ or ε_i. Thus, if $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$ is larger, then the target assumes that the synergy, $(1+s_i)$, is somewhat larger and that the misvaluation, $(1-\varepsilon_i)(1-\rho)$, is somewhat smaller (affiliation ensures that this is true). The target also uses all of his information from the other bids and his own misvaluation to try to determine if the increase is due to a shared effect. But the target's information is noisy and confounded by correlated synergies. Thus, an increase in s_i, ρ, ε_i, or $(1-\varepsilon_T)$ all increase the expectation of the synergy.

All of these results tell us that a simple lack of information can lead us to find exactly what our intuition would expect: Merger activity increases in overvalued markets and sectors, and overvalued firms buy undervalued firms.

C. Price Reaction

The fact that a merger occurs provides information about the true value of the target and the bidding firm. In a rational setting, participants recognize how the market will react to news of the merger. For simplicity we will assume that the market learns of the merger after the auction.[18] Thus, we examine the change in market prices on the announcement day.

COROLLARY 7: *On the announcement of a stock merger the target's and acquirer's market price could rise or fall. If the target's reservation price does not bind then the market price of the second highest bidder falls.*

Proof: See the Appendix.

A transaction is more likely if $\rho > 0, \varepsilon_1 > 0, s_i > 0$, and $\varepsilon_T < 0$. Therefore, conditional on an accepted offer, the expectations of ρ, ε_1, s_i, and $-\varepsilon_T$ are greater than zero. Prices must adjust until the prices once again equal the expectations of true value. For example, $E[\rho \mid merger] > 0$ pushes all prices down.

If a takeover is rebuffed then the target's price could fall if $E[\varepsilon_T \mid no\ merger] > 0$ is the largest effect. Furthermore, the bidders' stock prices should rise since the expectation of ρ and ε_1 is less than zero when the bids are refused.

Thus, it is easy to see why empirical work finds that the winning firm's stock price falls and the target's stock price rises on a takeover announcement.[19] This simply suggests that the market expects the winning firm to be overvalued, the target to be undervalued, and expects the synergies to be small or that competition gave most of the synergies to the target. It is also interesting to note that the losing bidder should have a permanent negative change in his stock price, and bidders in failed acquisitions should have a positive stock price change.

Therefore, taken together, the resulting data could make it appear that takeovers destroy value. However, all of our stock movements are the result of rational updating. Firms are attempting to create synergies in an environment with limited and asymmetric information. The stocks move not because any firm is destroying value by merging, but because in the attempt to create value, they are revealing information about what their price should have been.

IV. Merger Waves

Now that we understand individual mergers and market reactions, we can develop an understanding of merger waves. To do this we consider a sequence of potential mergers.

The model begins at time zero with prices and realizations of each variable. Then at each of m sequential time periods, an auction occurs for an acquisition candidate. Each firm has only one chance to merge. The market does not know

[18] To model how the market values change throughout the auction is a paper in and of itself. At each point in the auction the bidder's market value depends on what he bids, what others bid, and the differing probabilities about who will win! Thus, we consider the reaction post announcement.

[19] See McCardle and Viswanathan (1994) for an industrial organization model of price reactions around mergers.

the identity of the participants in the merger contest until after it is over.[20] After each auction, market prices react to the observed information. We assume that the event *merger* or *no merger* and all bids are observed at the end of each period.[21] At the end of m periods, true values are revealed.

In order to discuss merger waves, we must define a wave. The expected probability of a merger occurring in the first period depends on the distributions of the errors and the synergies. If every variable received a new realization each period (drawn from the same distribution each period), then the expected probability of a merger occurring in every period would be the same. We call this the unconditional expected probability of a merger. We will use this expected probability as the benchmark for a wave.

DEFINITION 1: *A merger wave is defined as a sequence of time periods (two or more) in which the probability of a merger occurring is above the unconditional expected probability of a merger.*[22]

Thus, a merger wave will be caused by realizations of the errors and synergies, and merger waves will be affected by market reactions.

THEOREM 2: *A high enough realization of the market-wide misvaluation, ρ, will cause a merger wave, even though given a merger in the first period, the market reduces prices until in expectation, there is no market-wide misvaluation left in prices.*

Proof: See the Appendix.

Theorem 1 showed that mergers occur more often the more the market is overvalued. Thus, a large enough realization of ρ will begin a merger wave. However, the merger wave will occur only if the market *stays* overvalued. However, the rational market updates correctly. Corollary 7 shows us that after a first-period merger, prices move until $(1-\rho)/E[1-\rho \mid \textit{merger}]$ is the common mispricing. This reduces the probability of the second merger since $E[1-\rho \mid \textit{merger}] < 1$. Although the $E[(1-\rho)/E[1-\rho \mid \textit{merger}]] = 1$, the realization may not be one. If $(1-\rho)/E[1-\rho \mid \textit{merger}]$ is lower, then another merger is more likely to occur, and it is lower if the misvaluation in the first period is higher. Therefore, merger waves can be caused by an overvalued market.

However, after the first merger contest, prices will react and alter the errors. Thus, markets might adjust prices and rapidly end waves. We do not think this is an accurate characterization of waves. Synergies are most likely correlated

[20] As noted above, it would take an entire paper to consider how market prices and thus information and bids change with the market's perception of the probability of winning and the synergies.

[21] All results hold if we assume that the highest bid is not observed or that no bid is observed.

[22] Therefore, when looking at historical data we will never know if a time period with a large number of mergers was a merger wave or just a high number of positive realizations. However, the more mergers we see, the more likely the time period is a wave.

across mergers, in which case, merger waves can also be caused by high realizations of the common synergy, λ. As the next theorem shows, this possibility will make the market unable to fully self-correct and end waves even when the waves are caused by overvaluation.

THEOREM 3: *The higher the expectation of the common synergy component, the less the market learns about market-wide misvaluation from a merger. The potential of a common synergy component extends the life of a merger wave that is caused by misvaluation.*

Proof: Whether or not there is a common component to synergies, Corollary 7 shows us that after a merger $E[\rho \mid merger] > 0$, so prices must decrease until $(1-\rho)/E[1-\rho \mid merger]$ is the common mispricing. When there is a common component to synergies, then each merger increases the expectation of λ (see the Appendix). The higher the expectation of λ, the lower the expectation of the market-wide misvaluation conditional on another merger, $E[(1-\rho)/E[1-\rho \mid merger] \mid merger]$ (see the Appendix). Therefore, each new merger decreases all prices at a decreasing rate. If there is no common synergy component, then the countereffect will not exist, and prices will decrease much faster. Q.E.D.

Overvaluation that causes a merger wave may not be fully corrected by a market that rationally updates. Each subsequent merger signals less and less information about market-wide overvaluation, as the market increases its expectation about a common synergy factor. Hence, the first merger leads to a significant downward revision in the market index, but subsequent mergers do not move the index as much. Therefore, a merger wave may be followed by a market crash when the participants learn information about the synergies that leads them to question the gains from the entire sequence of mergers.

On average, after a merger, firms are correctly priced. Hence, we should not expect a wave or any ex post drift in prices. In fact, a second merger is less likely than the first because of the market correction, but if the market is still overvalued, then a merger is more likely than it would be in expectation unconditionally. Thus, this is not a theory of clustering in the statistical sense. Rather, mergers occur at the same time because they are correlated with market overvaluation.

We could also model the stock prices, synergies, and errors as following a random walk. For example, we could assume that at the end of each period, a value is drawn that is added to the synergy and whatever error is currently left in the prices. This would increase the noise in the model and allow merger waves to last longer, but would not qualitatively change the results. However it is interesting to consider the model in the light of this realistic additional noise. Prices could randomly become better or worse at any time but mergers or their lack would cause prices to mean revert. However, as Theorem 3 made clear, they would not immediately mean revert, particularly with this additional noise. Thus, as fundamentals move away from market prices, possibly combined with a positive shock to the common synergy factor, then merger activity may increase. Since the market is unsure if the increase in activity is due to common synergies or due to overvaluation, prices could drift further and further from fundamentals, leading to more and more activity. However, we would slowly expect prices to mean revert. In this case, we should expect ex post downward drift after mergers. Eventually, the market will learn of the true synergies, and if they are lower than expected, the market will crash, ending the wave. Although a wave is not expected on average, in a long data set we should be unsurprised to find periods where prices drift upward and merger activity increases, followed by a market correction that ends the wave.

V. Equity Versus Cash

Up until this point in the paper, we have only allowed firms to bid with their own stock. If the firm knows that its stock is undervalued, then it may prefer to switch to a cash bid. We assume that only some firms have access to cash. We will show that when markets are overvalued, mergers are more likely to occur and those that occur are more likely to use stock. When markets are undervalued, mergers are less likely to occur and those that occur use cash.

For simplicity (and since balance sheets are public information), we assume that it is common knowledge which firms have access to cash. Managers are rational. Thus, if managers receive a stock offer they perceive as worth accepting from a bidder who has access to cash, they will simply request a similar amount in cash and remove the lemons (those with overvalued stock). With costless access to cash, there is no reason for bidders not to comply unless the true value of their offer is less than the perceived equity value. Therefore, in equilibrium, targets will only accept cash bids from firms that have costless access to cash.

Since only some firms can access cash, the market for stock mergers does not disappear. Rather, those firms with cash always use cash and those firms without access must use stock.

THEOREM 4: *If the target only accepts offers with an expected value greater than the target's true value, X_T, but not all firms have access to cash, then, (1) mergers are more likely to occur in overvalued markets than in undervalued markets, and (2) the method of payment will include a greater fraction of stock deals in overvalued markets than in undervalued markets.*

Proof: Any accepted bid must be perceived to be greater than X_T by the management of the target. Cash bidders will bid up to V_i, which equals $X_T(1+s_i)$. Therefore, for a cash bid to be accepted,

$$X_T(1+s_i) > X_T, \tag{15}$$

or $s_i > 0$.[23] Management's cash acceptance rule is unaffected by misvaluation because they know their true value and thus ask only for positive synergies.

[23] This rule allows managers to accept a cash bid with a value below their current stock price. While this is optimal, it does not seem realistic. It is more reasonable to assume that stockholders

Theorem 1 tells us that the manager is more likely to perceive that a bid is greater than X_T if the market is overvalued, $\rho > 0$. Therefore, when the market is overvalued, management perceives stock bids to be more valuable but management's perception of cash bids is unaltered. Thus, stock bids are more likely to win in an overvalued market and cash bids make up a higher fraction of completed deals in an undervalued market. Q.E.D.

This theorem demonstrates why it is rational for more mergers to occur in stock when the market is overvalued and in cash when the market is undervalued. This is not as obvious as it sounds, because we are not simply saying that bidders with overvalued stock would like to bid with stock. They would, but why would targets accept? Our point is that in any rational model, the participants will choose every action correctly on average. Therefore, the target will correctly reject stock offers that are not valuable enough *on average*. However, the target will make ex post mistakes. The mistakes are correlated with market-wide misvaluation.

Therefore, not only should we see waves of stock mergers in overvalued markets, but in undervalued markets we should see less activity, and that activity should be in cash. This result is consistent with Andrade et al. (2001), who find a larger positive announcement effect on the target for cash offers and a less negative effect on the acquirer. We suggest that this does not imply that cash mergers are better than stock mergers, but rather that cash mergers are more likely to occur in undervalued markets. So, the rational market updates and increases stock prices.

VI. Conclusion

The evidence that waves occur is clear. That as of yet the explanation of waves is incomplete is also clear. There are a number of reasons why any given wave of mergers could occur. For example, deregulation could release pent-up demand, or a new technology could require the redeployment of assets. However, we believe that these reasons do not tell the whole story. Furthermore, these ideas tell us nothing about why the medium of exchange is stock or cash. In this paper, we lay out a valuation effect that is important and we show that this effect can cause a wave even without deregulation or innovation.

Our idea is that even fully rational participants make mistakes, that is, their decision turns out to be the wrong decision ex post even though it was correct ex ante. We focus on how these mistakes could be correlated with specific types of misvaluation. When the market is overvalued, the target rationally reduces the expected value of a given stock offer, and thus, the target values the offer correctly *on average*. However, the target is more likely to overvalue the offer the greater the market overvaluation is *even though the target's own stock is affected by the same market overvaluation*. Thus, market overvaluation increases the

will sue (and win) if the value of the offer is less than the current stock price. Eliminating this kind of merger only magnifies our result, because cash offers would be rejected more often when the market was overvalued.

chance that a merger occurs. Therefore, a wave can occur due to misvaluation even if there is no underlying reason for mergers. Furthermore, waves can be halted by undervaluation even if assets truly should be redeployed. Thus, the impacts of misvaluation are significant.

Misvaluation also influences the medium of exchange. We believe that in most cases, for a stock merger to occur, the target's management must expect the deal to increase value. Managers make errors when evaluating stock offers (although they get it right on average) but not when evaluating cash offers. Therefore, the medium of exchange will contain a higher fraction of stock offers when the market is overvalued and completed deals are more likely to be in cash in undervalued markets. Furthermore, markets will react more positively to news of a cash merger than to an equity merger.

We believe that valuation, or rather misvaluation, has a fundamental impact on all mergers. Valuation affects not only the likelihood that the merger occurs but also the medium of exchange. We show how merger waves and waves of cash and stock purchases can be driven by periods of overvaluation and undervaluation of the stock market.

Appendix

Proof of Lemma 1: In our second-price mechanism, the winning bidder must pay the smallest fraction they could have bid and just achieved the score of the second highest bidder. Consider an α_i such that

$$g(\alpha_i, M_i, \Phi_T) - \max_{\forall j} g(\alpha_j, M_j, \Phi_T) > 0. \tag{A1}$$

Over the interval $[0, \alpha_i]$, $g(\alpha_j, M_j, \Phi_T)\,\forall j$ are continuous functions of α_i on a compact set. Therefore, $g(\alpha_j, M_j, \Phi_T)\forall j$ are of bounded variation, and $\max_{\forall j} g(\alpha_j, M_j, \Phi_T)$ is continuous in α_i. Thus, since

$$g(0, M_i, \Phi_T) - \max_{\forall j} g(\alpha_j, M_j, \Phi_T) < 0, \tag{A2}$$

there must exist at least one point where

$$g(\hat{\alpha}, M_i, \Phi_T) - \max_{\forall j} g(\alpha_j, M_j, \Phi_T) = 0. \tag{A3}$$

At any $\hat{\alpha}$ where this is true, we know

$$\frac{\partial g}{\partial \hat{\alpha}}(\hat{\alpha}, M_i, \Phi_T) > \frac{\partial g}{\partial \hat{\alpha}}(\alpha_j, M_j, \Phi_T) \quad \forall j. \tag{A4}$$

Therefore, $\exists$ a unique $\hat{\alpha}$ (call this alpha $\hat{\alpha}_i$) such that

$$g(\hat{\alpha}_i, M_i, \Phi_T) - \max_{\forall j} g(\alpha_j, M_j, \Phi_T) = 0. \tag{A5}$$

Note that $0 \le \hat{\alpha}_i \le \alpha_i$.

Bidder i maximizes

$$\max_{\alpha_i} E\left[\{V_i - \hat{\alpha}_i(X_i + V_i)\}1_{\{\max_{\forall j}[g(\alpha_j, M_j, \Phi_T)] \leq g(\alpha_i, M_i, \Phi_T)\}} \mid X_i, V_i\right], \tag{A6}$$

or

$$\max_{\alpha_i} E\left[\left\{\frac{V_i}{X_i + V_i} - \hat{\alpha}_i\right\}1_{\{\max_{\forall j}[g(\alpha_j, M_j, \Phi_T)] \leq g(\alpha_i, M_i, \Phi_T)\}} \,\Big|\, X_i, V_i\right](X_i + V_i). \tag{A7}$$

Using the definition of $\hat{\alpha}_i$ from above, this can be rewritten as

$$\max_{\alpha_i} E\left[\left\{\frac{V_i}{X_i + V_i} - \hat{\alpha}_i\right\} 1_{\{\hat{\alpha}_i \leq \alpha_i\}} \Big| X_i, V_i\right](X_i + V_i). \tag{A8}$$

If $\alpha_i > \frac{V_i}{X_i+V_i}$ then reducing α_i to $\frac{V_i}{X_i+V_i}$ only eliminates cases where $\frac{V_i}{X_i+V_i} - \hat{\alpha}_i < 0$ (remember $\hat{\alpha}_i$ is unique and does not change with α_i). Similarly, if $\alpha_i < \frac{V_i}{X_i+V_i}$ then increasing α_i to $\frac{V_i}{X_i+V_i}$ only adds cases where $\frac{V_i}{X_i+V_i} - \hat{\alpha}_i > 0$. Therefore, it is a dominant strategy to set $\alpha_i = \frac{V_i}{X_i+V_i}$. Q.E.D.

Proof of Lemma 2:

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] = E\left[X_T(1 + s_i) \mid \alpha_i, \Phi_T\right] \tag{A9}$$

$$= X_T E\left[(1 + s_i) \,\Big|\, \frac{\alpha_i}{1 - \alpha_i}, M_i\, \forall i, X_T, M_T\right] \tag{A10}$$

$$= X_T E\left[(1 + s_i) \,\Big|\, \frac{\alpha_i}{1 - \alpha_i}\frac{M_i}{M_T}, M_i\, \forall i, \frac{X_T}{M_T}, M_T\right]. \tag{A11}$$

If Lemma 1 holds, then $\alpha_i = \frac{V_i}{X_i+V_i}$, so

$$\frac{\alpha_i}{1 - \alpha_i}\frac{M_i}{M_T} = \frac{(1 + s_i)(1 - \varepsilon_T)}{(1 - \varepsilon_i)}, \tag{A12}$$

$$\frac{X_T}{M_T} = (1 - \varepsilon_T)(1 - \rho). \tag{A13}$$

Therefore,

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] = X_T E\left[(1 + s_i) \,\Big|\, \frac{(1 + s_i)(1 - \varepsilon_T)}{(1 - \varepsilon_i)}, M_i \forall i, (1 - \varepsilon_T)(1 - \rho), M_T\right]. \tag{A14}$$

Furthermore, $\frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$, $(1 + s_i)$, and $(1 - \varepsilon_T)(1 - \rho)$ are independent of M_i and M_T. Therefore,

$$E\left[V_i \mid \alpha_i, M_i, \Phi_T\right] = X_T E\left[(1 + s_i) \,\Big|\, \frac{(1 + s_i)(1 - \varepsilon_T)}{(1 - \varepsilon_i)} \forall i, (1 - \varepsilon_T)(1 - \rho)\right] \tag{A15}$$

$$= X_T E\left[(1+s_i)\,\middle|\, \frac{\alpha_i}{1-\alpha_i}\frac{M_i}{M_T}\forall i, \frac{X_T}{M_T}\right] \tag{A16}$$

$$= E\left[V_i \,\middle|\, \frac{\alpha_i}{1-\alpha_i}M_i\,\forall i, X_T, M_T\right]. \tag{A17}$$

Note that X_T and M_T do not depend on the bidder. Therefore, the only bidder-specific information relevant for the expectation of V_i is $\frac{\alpha_i}{1-\alpha_i}M_i$. Since every variable is drawn identically, the bidders are scored symmetrically, and $\frac{\alpha_i}{1-\alpha_i}M_i$ can be used to rank the bids. Q.E.D.

Proof of Lemma 3: The second highest Z-score is

$$g(\alpha_2, M_2, \hat{\alpha}_1, M_1, \Phi_{T-}) = E\left[V_2 \mid \alpha_2, M_2, \hat{\alpha}_1, M_1, \Phi_{T-}\right]. \tag{A18}$$

However, Lemma 2 demonstrated that the only relevant bidder-specific information for the score is $\frac{\alpha}{1-\alpha}M$. Therefore,

$$g(\alpha_2, M_2, \hat{\alpha}_1, M_1, \Phi_{T-}) = E\left[V_2 \,\middle|\, \frac{\alpha_2}{1-\alpha_2}M_2, \frac{\hat{\alpha}_1}{1-\hat{\alpha}_1}M_1, \frac{\alpha_j}{1-\alpha_j}M_j \forall j \neq 1\, or\, 2, X_T, M_T\right]. \tag{A19}$$

So equation (9) can be written as

$$Z_2 \equiv E\left[V_2 \,\middle|\, \frac{\alpha_2}{1-\alpha_2}M_2, \frac{\hat{\alpha}_1}{1-\hat{\alpha}_1}M_1, \frac{\alpha_j}{1-\alpha_j}M_j \forall j \neq 1\, or\, 2, X_T, M_T\right]$$
$$= E\left[V_1 \,\middle|\, \frac{\hat{\alpha}_1}{1-\hat{\alpha}_1}M_1, \frac{\alpha_2}{1-\alpha_2}M_2, \frac{\alpha_j}{1-\alpha_j}M_j \forall j \neq 1\, or\, 2, X_T, M_T\right]. \tag{A20}$$

Since everything is drawn identically, the scoring rule is symmetric, and these scores will equate when

$$\frac{\hat{\alpha}_1}{1-\hat{\alpha}_1}M_1 = \frac{\alpha_2}{1-\alpha_2}M_2. \tag{A21}$$

Rearranging shows that the highest bidder must pay

$$\hat{\alpha}_1 = \frac{\frac{\alpha_2}{1-\alpha_2}M_2}{M_1 + \frac{\alpha_2}{1-\alpha_2}M_2}. \tag{A22}$$

Substituting $\frac{\alpha_2}{1-\alpha_2} = \frac{V_2}{X_2}$ and $X_i = M_i(1-\varepsilon_i)(1-\rho)$ into equation (A22) shows

$$\hat{\alpha}_1 = \frac{\frac{V_2}{X_2}M_2}{M_1 + \frac{V_2}{X_2}M_2} = \frac{\frac{V_2}{(1-\varepsilon_2)(1-\rho)}}{\frac{X_1}{(1-\varepsilon_1)(1-\rho)} + \frac{V_2}{(1-\varepsilon_2)(1-\rho)}}$$
$$= \frac{V_2(1-\varepsilon_1)}{X_1(1-\varepsilon_2) + V_2(1-\varepsilon_1)}. \qquad \text{(A23)}$$

Q.E.D.

Proof of Lemma 4: To prove that the relevant variables are affiliated, we use the log-concavity assumption from Section II.D. Before we use this assumption we first define log-concavity.

DEFINITION: *A random variable x has log-concave density $f(x)$ if $\log f(x)$ is concave.*

The assumption of log-concavity is standard in economic problems where inference is involved. Caplin and Nalebuff (1991) discuss the origins and implications of this idea and list its applications in economics. Distributions with log-concave densities include the multivariate beta, Direchlet, exponential, gamma, Laplace, normal, uniform, Weibull, and Wishart distributions. Log-concavity implies that the distribution is unimodal.

We note two implications of log-concavity.

IMPLICATION 1: *If x has log-concavity density, then so does $-x$.*

IMPLICATION 2: *If x and y have log-concave densities, so does $x + y$.*

Now define $S_i = \log(1+s_i)$, $u_i = -\log(1-\varepsilon_i)$, $u_T = -\log(1-\varepsilon_T)$, and $m = \log(1-\rho)$. Then

$$\begin{aligned} & f(S_i, u_i, S_i + u_i - m, S_i + u_i - (S_j + u_j)\forall j \neq i, S_i + u_i - u_T) \\ & = g(S_i, u_i) f(S_i + u_i - m, S_i + u_i - (S_j + u_j)\forall j \neq i, S_i + u_i - u_T \mid S_i, u_i) \\ & = h(S_i)\, l(u_i) f(S_i + u_i - m \mid S_i, u_i) f(S_i + u_i - (S_j + u_j)\forall j \neq i \mid S_i, u_i) \\ & \quad \times f(S_i + u_i - u_T \mid S_i, u_i). \end{aligned} \qquad \text{(A24)}$$

Consider the term $f(S_i + u_i - m | S_i, u_i)$, let $t = S_i + u_i - m$, $S_i = x$, $u_i = y$, and note that

$$f(t \mid x, y) = f(t \mid x + y) = f_{-m}(t - x - y). \qquad \text{(A25)}$$

Since m has a log-concave density, so does $-m$. Let $t' > t, x' > x$ and $y' > y$. The monotone-likelihood property requires that

$$\begin{aligned} & f(t \mid x, y) f(t' \mid x', y') > f(t \mid x', y') f(t' \mid x, y) \\ & \quad \Leftrightarrow \log f_{-m}(t' - x' - y') - \log f_{-m}(t - x' - y') \\ & \quad > \log f_{-m}(t' - x - y) - \log f_{-m}(t - x - y) \\ & \quad \Leftrightarrow \frac{d \log f_{-m}(t - x' - y')}{dt} > \frac{d \log f_{-m}(t - x - y)}{dt}, \end{aligned} \tag{A26}$$

which is true from the log-concavity of $-m$ as $t - x' - y' < t - x - y$.

Similar arguments prove the monotone-likelihood property for $f(S_i + u_i - (S_j + u_j) \forall j \neq i \mid S_i, u_i)$ and $f(S_i + u_i - u_T \mid S_i, u_i)$ (we use here the fact that the sum of log-concave densities is log-concave). Using both part (i) and part (ii) of Milgrom and Weber's (1982) Theorem 1, it follows that S_i, $u_i, S_i + u_i - m, S_i + u_i - (S_j + u_j)$, and $S_i + u_i + u_T$ are affiliated. Since subsets of affiliated variables are also affiliated (see Theorem 4 in Milgrom and Weber), it follows that $S_i, S_i + u_i - m, S_i + u_i - (S_j + u_j)$, and $S_i + u_i - u_T$ are affiliated. Then, noting that $s_i = e^{S_i} - 1$, therefore, Theorems 3 and 5 in Milgrom and Weber immediately imply Theorem 1 in our paper (we note that the information in the variables that we condition on is not changed if we take logs). Q.E.D.

Proof of Corollary 4: If the bidders and the target share a sector misvaluation, then Theorem 1 tells us that mergers are more likely to happen in that sector since ρ subsumes any shared misvaluation. With the assumptions that $(1 - \varepsilon_i) = (1 - \psi_b)(1 - \phi_i)$, and that the target does not share the sector misvaluation, $(1 - \varepsilon_T) = (1 - \psi_T)(1 - \phi_T)$, then the conditioning variables $\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$, and $\frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$ become $\frac{(1+s_i)}{(1-\psi_b)(1-\phi_i)(1-\rho)}$ and $\frac{(1+s_i)(1-\psi_T)(1-\phi_T)}{(1-\psi_b)(1-\phi_i)}$. It is straightforward to show that these variables are still affiliated; assuming, of course, that the variables $\log(1 - \phi)$ and $\log(1 - \psi)$ have log-concave densities. Thus, since $\frac{(1+s_i)}{(1-\psi_b)(1-\phi_i)(1-\rho)}$ and $\frac{(1+s_i)(1-\psi_T)(1-\phi_T)}{(1-\psi_b)(1-\phi_i)}$ are increasing in ψ_b, bidders from the same overvalued sector as the target are more likely to win than bidders from other sectors. Furthermore, $\frac{(1+s_i)(1-\psi_T)(1-\phi_T)}{(1-\psi_b)(1-\phi_i)}$ is decreasing in ψ_T, therefore targets from relatively undervalued sectors are more likely to accept offers. Q.E.D.

Proof of Corollary 6: If $(1 + s_i) = (1 + \lambda)(1 + \omega_i)$, then in equation (14) the conditioning variables $\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)} \forall j \neq i$ all become $\frac{(1+\omega_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+\omega_j)} \forall j \neq i$. It is again straightforward to show that these variables are still affiliated as long as we assume that the variables $\log(1 + \lambda)$ and $\log(1 + \omega_i)$ have log-concave densities. Thus, the target is able to learn from the other bids about $\frac{1-\rho}{1+\lambda}$, but cannot tell the difference between a high market-wide synergy and a high market-wide overvaluation. Q.E.D.

Proof of Corollary 7: Note that acceptance of the merger implies that

$$E\left[(1+s_i)\left|\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}, \frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right.\right] > 1. \quad \text{(A27)}$$

Hence, Lemma 4 implies that acceptance occurs if and only if

$$\frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)} > c\left(\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right) = C$$
$$\Longleftrightarrow S_i + u_i - m - \log\ C > 0, \quad \text{(A28)}$$

where $c(\cdot)$ is a nonincreasing function of $\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i$, and $\frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$, and S_i, u_i, and m are defined above. Hence,

$$m < S_i + u_i - \log C \Longleftrightarrow \rho = 1 - e^m > 1 - e^{S_i+u_i-\log C}. \quad \text{(A29)}$$

From this it directly follows that

$$E\left[\rho \mid \rho > 1 - e^{S_i+u_i-\log C}, \frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right] > E[\rho] = 0. \quad \text{(A30)}$$

Conditional on an accepted offer, it can be similarly shown that $E[\varepsilon_1 \mid \textit{merger}] > 0, E[\varepsilon_T \mid \textit{merger}] < 0$, $E[\varepsilon_2 \mid \textit{merger}] > 0$, $E[s_i \mid \textit{merger}] > 0$, $E[\lambda \mid \textit{merger}] > 0$ therefore, the proofs are omitted.

Market prices must adjust until they are once again equal to the expectations of true value. The expression $E[\rho \mid \textit{merger}] > 0$ pushes all prices down until $(1-\rho)/E[1-\rho \mid \textit{merger}]$ is the common mispricing that is not corrected. $E[\varepsilon_1 \mid \textit{merger}] > 0, E[\varepsilon_2 \mid \textit{merger}] > 0$, and $E[\varepsilon_T \mid \textit{merger}] < 0$ have similar effects on the bidders' and target's firm-specific mispricing. And, $E[s_i \mid \textit{merger}] > 0$ pushes both the winning bidder's price and the target's price up. Q.E.D.

Proof of Theorem 2: Theorem 1 showed that a large enough ρ will ensure that the probability of a merger in the first period is above the ex ante expected probability. From Theorem 1 we know that a merger occurs in the first period if

$$E\left[(1+s_i) \mid a_1, b_1, c_1\right] > 1, \quad \text{(A31)}$$

where $a_1 = \frac{(1+s_i)}{(1-\varepsilon_i)(1-\rho)}$, $b_1 = \frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i$, and $c_1 = \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}$. In the second period there will be a new target and new bidders. Therefore, there will be new realizations of the errors and synergies. Let s^2 and ε^2 represent the synergy and firm-specific error in the second period (where the superscript represents the second period). There is, however, only one realization of ρ. We know from Corollary 7 that the first period tells the market that

$$E\left[\rho \mid \rho \geq 1 - e^{S_i+u_i-\log C}, b_1, c_1\right] > E[\rho] = 0 \quad \text{(A32)}$$

if there is a merger, and

$$E\left[\rho \mid \rho < 1 - e^{S_i + u_i - \log C}, b_1, c_1\right] \leq E[\rho] = 0 \tag{A33}$$

if there is not a merger. Therefore the market will move all prices after the first period. Let

$$1 - \rho^2 \equiv (1 - \rho)/E\left[1 - \rho \mid \rho > 1 - e^{S_i + u_i - \log C}, b_1, c_1\right], \tag{A34}$$

if there is a merger in the first period, and let

$$1 - \rho^2 \equiv (1 - \rho)/E\left[1 - \rho \mid \rho \leq 1 - e^{S_i + u_i - \log C}, b_1, c_1\right] \tag{A35}$$

if there is not. Therefore, $(1 - \rho^2)$ is the market-wide error left in prices after the market reacts to the news in the first period. Note that the truncated distribution of a log-concave variable is log-concave. Therefore, $\log(1 - \rho^2)$ has a log-concave distribution. Furthermore, this is true regardless of what information is known after the first period (S_i, u_i, C, b_1, c_1, etc.), because integrating over the unknown information S_i, u_i, (given b_1, c_1, or no information or some intermediate information) does not change the log-concavity. Therefore, we can assume that after the first-period, all of the first-period information is released, or only the fact that a merger occurs, or anything in between.

In the second period, acceptance of the merger implies that

$$E\left[\left(1 + s_i^2\right) \middle| \frac{\left(1 + s_i^2\right)}{\left(1 - \varepsilon_i^2\right)\left(1 - \rho^2\right)}, \frac{\left(1 + s_i^2\right)\left(1 - \varepsilon_j^2\right)}{\left(1 - \varepsilon_i^2\right)\left(1 + s_j^2\right)} \forall j \neq i, \frac{\left(1 + s_i^2\right)\left(1 - \varepsilon_T^2\right)}{\left(1 - \varepsilon_i^2\right)}\right] > 1. \tag{A36}$$

Since all of the variables are log-concave, Lemma 4 (with $m = \log(1 - \rho)$ and equation (A24) conditional on $\rho \leq 1 - e^{S_i + u_i - \log C}, b_1, c_1$) ensures that everything we know about period one mergers is also true in period two although the market-wide misvaluation has changed. First, note that if there is a merger in the first period, then $\rho^2 < \rho$, so the market reduces prices. Therefore, a merger in the second period is less likely to be caused by market-wide misvaluation. Second, regardless of whether there was or was not a merger in the first period, ρ^2 is an increasing function of ρ. Therefore, a high enough realization of ρ will cause the probability of a merger in the second period to be above the ex ante expected probability. Thus, a high enough ρ will cause a merger wave. The same logic will hold for the third, fourth, or n^{th} period. Q.E.D.

Proof of Theorem 3: The proof that the expectation of λ increases with every merger is similar to the proof of Corollary 7. Acceptance of a merger proposal occurs only if the inequality in equation (A27) is satisfied. Hence, Lemma 4 implies that acceptance occurs if and only if

$$\frac{(1 + s_i)}{(1 - \varepsilon_i)} > d\left(\frac{1}{1 - \rho}, \frac{1 - \varepsilon_j}{1 + s_j} \forall j \neq i, 1 - \varepsilon_T\right) = D$$
$$\iff S_i + u_i - \log D > 0, \tag{A37}$$

where $d(\cdot)$ is some nonincreasing function of its arguments, u_i, and u_T are defined above, and since $1+s_i=(1+\lambda)(1+w_i)$, $S_i=\Lambda+W_i$, where $\Lambda=\log(1+\lambda)$ and $W_i=\log(1+w_i)$. Therefore,

$$\Lambda > \log D - W_i - u_i \Longleftrightarrow \lambda = e^{\Lambda} - 1 > e^{\log D - W_i - u_i} - 1. \tag{A38}$$

It follows directly from this that

$$E\big[\lambda \mid \lambda > e^{\log D - W_i - u_i} - 1, b_1, c_1\big] > E[\lambda]. \tag{A39}$$

The proof that the $E[\rho \mid merger]$ is smaller if λ is larger follows almost directly from equation (A30) above. The conditional expectation of ρ is a monotonic increasing function of $1-e^{S_i+u_i-\log c\left(\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right)}$. The common synergy component is in s_i and s_j. Therefore, the common component is not a part of the first argument in the function $c(\cdot)$ because it cancels out. Furthermore, $c(\cdot)$ is a nonincreasing function of its arguments and the second argument of $c(\cdot)$ is increasing in λ. Thus, if the common synergy component is larger, then $1-e^{S_i+u_i-\log c\left(\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right)}$ is smaller. Furthermore, the conditional distribution of S_i+u_i is increasing in b_1 and c_1 in the sense of first-order stochastic dominance. Therefore, the conditional distribution of $1-e^{S_i+u_i-\log c\left(\frac{(1+s_i)(1-\varepsilon_j)}{(1-\varepsilon_i)(1+s_j)}\forall j\neq i, \frac{(1+s_i)(1-\varepsilon_T)}{(1-\varepsilon_i)}\right)}$ is smaller in the sense of first-order stochastic dominance, and hence the conditional expectation of ρ is lower. Q.E.D.

REFERENCES

Andrade, Gregor, Mark Mitchell, and Erik Stafford, 2001, New evidence and perspectives on mergers, *Journal of Economic Perspectives* 15, 103–120.

Berkovitch, Elazar, and M.P. Narayanan, 1990, Competition and the medium of exchange in takeovers, *Review of Financial Studies* 3, 153–174.

Brealey, Richard, and Stewart Myers, 1996, *Principles of Corporate Finance* (McGraw Hill, New York).

Caplin, Andrew, and Barry Nalebuff, 1991, Aggregation and social choice: A mean voter theorem, *Econometrica* 59, 1–23.

Eckbo, B. Espen, Ron M. Giammarino, and Robert L. Heinkel, 1990, Asymmetric information and the medium of exchange in takeovers: Theory and tests, *Review of Financial Studies* 3, 651–675.

Fishman, Michael J., 1989, Preemptive bidding and the role of the medium of exchange in acquisitions, *Journal of Finance* 44, 41–57.

Gort, Michael, 1969, An economic disturbance theory of mergers, *Quarterly Journal of Economics* 83, 624–642.

Gorton, Gary, Matthias Kahl, and Richard Rosen, 2000, Eat or be eaten: A theory of mergers and merger waves, Unpublished working paper, Wharton, University of Pennsylvania.

Hansen, Robert G., 1985, Auctions with contingent payments, *American Economic Review* 75, 862–865.

Hansen, Robert G., 1987, A theory for the choice of exchange medium in mergers and acquisitions, *Journal of Business* 60, 75–95.

Hietala, Pekka, Steven N. Kaplan, and David T. Robinson, 2003, What is the price of hubris? Using takeover battles to infer overpayments and synergies, *Financial Management* 32, 1–37.

Holmstrom, Bengt, and Steven N. Kaplan, 2001, Corporate governance and merger activity in the United States: Making sense of the 1980s and 1990s, *Journal of Economic Perspectives* 15, 121–144.

Hou, Kewei, Per Olsson, and David T. Robinson, 2002, Does takeover increase stockholder value? Unpublished working paper, Duke University.

Kaplan, Steve, 1993, Paramount Communications A and B, Unpublished case, University of Chicago.

Jovanovic, Boyan, and Peter Rousseau, 2001, Mergers and technological change: 1885–2001, Unpublished working paper, Vanderbilt University.

Jovanovic, Boyan, and Peter Rousseau, 2002, The Q-theory of mergers, *American Economic Review* 92, 198–204.

Maksimovic, Vojislav, and Gordon Phillips, 2001, The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains? *Journal of Finance* 56, 2019–2065.

Martin, Kenneth, 1996, The method of payment in corporate acquisitions, investment opportunities, and managerial ownership, *Journal of Finance* 51, 1227–1246.

McCardle, Kevin F., and S. Viswanathan, 1994, The direct entry versus takeover decision and stock price performance around takeovers, *Journal of Business* 67, 1–43.

Milgrom, Paul, and Robert Weber, 1982, A Theory of auctions and competitive bidding, *Econometrica* 50, 1089–1122.

Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.

Persons, John, and Vincent Warther, 1997, Boom and bust patterns in the adoption of financial innovations, *Review of Financial Studies* 10, 939–967.

Rhodes-Kropf, Matthew, David T. Robinson, and S. Viswanathan, 2004, Market valuations and merger activity: The empirical evidence, *Journal of Financial Economics* (forthcoming).

Rhodes-Kropf, Matthew, and S. Viswanathan, 2000, Corporate reorganizations and non-cash auctions, *Journal of Finance* 55, 1807–1849.

Rhodes-Kropf, Matthew, and S. Viswanathan, 2002, Financing auction bids, Unpublished working paper, Columbia University.

Samuelson, William, 1987, Auctions with contingent payments: Comment, *American Economic Review* 77, 740–745.

Shleifer, Andrei, and Robert Vishny, 2003, Stock market driven acquisitions, *Journal of Financial Economics* 70, 295–311.

Toxvaerd, Flavio, 2002, Strategic merger waves: A theory of musical chairs, Unpublished working paper, Hebrew University.