Do Insiders Learn from Outsiders? Evidence from Mergers and Acquisitions
Author(s): Yuanzhi Luo
Source: *The Journal of Finance*, Vol. 60, No. 4 (Aug., 2005), pp. 1951-1982
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/3694858

THE JOURNAL OF FINANCE • VOL. LX, NO. 4 • AUGUST 2005

Do Insiders Learn from Outsiders? Evidence from Mergers and Acquisitions

YUANZHI LUO*

ABSTRACT

I find that the market reaction to a merger and acquisition (M&A) announcement predicts whether the companies later consummate the deal. The relation cannot be explained by the market's anticipation of the closing decision or its perception of the deal quality at the announcement. Merging companies appear to extract information from the market reaction and later consider it in closing the deal. Furthermore, the relation varies with deal characteristics, suggesting that companies seem to have a higher incentive to learn from the market when canceling the announced deal is easier or when the market has more information that the companies do not know.

IMAGINE AFTER PRIVATE NEGOTIATIONS, the CEOs of two companies conclude that a merger will create value. They then reach a deal and announce it to the public. Outside investors, however, disagree with the proposed merger after being informed. They believe that the managers have overlooked certain weaknesses of the plan, and if carried out, it will reduce the total value of the two companies. The investors express their opposition in many ways, including trading down the companies' stock prices. What do the managers do in such a situation? Do they ignore the market and consummate the merger anyway, or do they learn from the market and cancel the deal?

In this paper, I investigate learning during mergers and acquisitions (M&As). "Learning" means, specifically, that the managers of merging companies extract information from the stock market reaction to the M&A announcement and consider the information in making the closing decision. In short, learning implies information flows from the market to the company. Typically, only a small number of top managers and their advisors make M&A decisions. Market participants, including stock analysts and institutional investors, can be better positioned than those insiders of the merging companies to analyze the international, macroeconomic, and industry issues that are relevant to the deal.

*Yuanzhi is a hedge fund portfolio manager at OTA Asset Management. The paper was first submitted and accepted when I was at Goldman, Sachs & Co. It is adapted from Chapter 1 of my Ph.D. dissertation at the University of Rochester. I am most indebted to my advisor, John Long. I thank him for instilling in me a scientific attitude to research and for offering sound advice throughout this project. During the adaptation, the questions and suggestions of an anonymous referee and Rick Green (the previous editor) have also been very helpful. In addition, the paper has benefited from comments by Andrew Alford, Elizabeth Demers, Gregg Jarrell, Chris Jones, Xiaolei Liu, Clifford Smith, Ross Watts, and participants of the 2004 AFA Annual Meeting. I thank Sandra Sizer for editorial assistance. The usual disclaimers apply.

Besides, investors may detect errors and omissions of the insiders. It is reasonable to presume that learning occurs during M&As.

Anecdotally, both learners and non-learners abound. *BusinessWeek* reports that Lucent stopped merger discussions with Alcatel because "investors clearly signaled their displeasure with it." After dropping the bid for the consulting arm of PricewaterhouseCoopers, Carly Fiorina, CEO of Hewlett-Packard, said to a group of analysts and institutional investors, "[A] number of you verbalized your concerns... and others simply voted with their positions in the stock... I realize you made some valid points." However, Quaker Oats ignored warnings from the market when buying Snapple. The day after they announced the deal, Quaker's stock price plunged nearly 10%, and Snapple also lost market value. Several analysts argued loudly against the merger. Quaker Oats nevertheless consummated the deal in 1995 by paying $1.7 billion. Under Quaker's mass-market-style management, Snapple quickly lost its original appeal to young and active consumers, and its sales deteriorated. Quaker finally sold Snapple in 1997 for $300 million, realizing a loss of $1.4 billion.[1]

Surprisingly, Jennings and Mazzeo (1991), who may be the only authors to study learning during M&As, claim that their data do not support the occurrence of learning. They investigate the relation between the bidder's stock return around the initial M&A announcement and the bidder's later deal closing and offer revision decisions but do not find any consistent association. They conjecture that either the manager's information subsumes the investor's information or the manager is under the influence of hubris that is so severe that it obliterates the incentive to learn.

In the broader literature, there has been little definitive, large-sample evidence that companies utilize information from the capital market to improve their investment decisions, although many economists—including Rock (1986), Jegadeesh, Weinstein, and Welch (1993), and Dye and Sridhar (2000)—believe that investors collectively have information about a company's investment projects that the managers of the company do not know. Dye and Sridhar (2000) comment, "[the current literature] fail(s) to recognize that information flows between capital markets and firms need not be just from firms to capital markets (as recognized in the extant literature), but also be from capital markets to firms."[2]

Learning is an important issue both in theory and in practice. The efficient market hypothesis (EMH) suggests that the capital market is superior to any

[1] Related news stories include "Lucent: Suddenly, the hole is even deeper" *BusinessWeek* 6/11/2001, "Hewlett shelves PWC deal: US10 cents below expectations: Computer maker falls short of financial targets" *National Post* 11/14/2000, "Quaker Oats Posts $1.11 Billion Quarterly Loss" *The Wall Street Journal* 4/24/1997, and "Investors sour on Quaker Oats' offer for Snapple" *The Associated Press* 11/3/1994.

[2] This idea is scattered in the literature. Rock (1986), Jegadeesh et al. (1993), and Bommel (1997) claim that the capital market can be better informed about the quality of an IPO company than the company itself. Giammarino et al. (2004) make a similar claim about SEOs. Rappaport (1987) and Dye and Sridhar (2000) argue that managers can find out the effectiveness of their strategies by closely reading the market reaction to their disclosures. Brealy and Myers (2000) believe Viacom should have dropped its bid for Paramount after the market reacted negatively.

individual or group of individuals in processing public information. If companies maximize their shareholder value, we expect that the informational efficiency of the market contributes to the economic efficiency of investment decisions inside the company. In reality, M&As are among the biggest investment decisions a company ever makes. However, nearly half of the deals in my sample experience negative market reactions around the initial announcement, and the majority of them are later consummated. The situation prompts a question: Do the companies ignore investors' opinions and, ultimately, the shareholder value in making M&A decisions?

Using new hypotheses, new test specifications, and a much larger sample than the one studied by Jennings and Mazzeo (1991), I find that the stock market reaction to the initial M&A announcement predicts the deal's final closing status. Moreover, certain deal characteristics affect the relation, as predicted by my model of learning. The evidence is consistent with the following interpretations. First, learning seems to occur in M&As. Merging companies appear to extract information from the market reaction and consider it in closing the deal. Second, the incentive to learn varies with the economic and informational conditions the companies face. They are more likely to learn from the market when canceling the announced deal is easier or when, ex ante, the market is expected to have more information that the companies do not know.

Learning is a decision, and merging companies make the decision based on cost–benefit analysis. Understanding investors' opinions incurs costs. For example, top managers must spend time and effort to communicate with investors. The prolonged deal closing caused by such activities may also help the companies' competitors on both the corporate control and the product markets. Companies commit the resources to learning only if they believe the expected gain outweighs the cost. The easier it is for the companies to renege on an announced deal, the higher the chance that additional information from the market will alter the decision, and thus the more valuable the expected benefit of learning. The merging companies' incentive to learn is also influenced by the market's relative information advantage over the companies. The market's opinion is more helpful to the merging companies if outside investors have more access to important raw deal valuation information, or if the companies have fewer resources to work out the valuation by themselves.

I develop a new method for testing learning. The correlation between the market reaction to the M&A announcement and the later deal-closing (or renegotiation) decision is not a sufficient proof of learning even if it is statistically significant. There are two possible non-learning explanations for such a relation. First, even if there is no learning, investors estimate the deal completion probability at the announcement, and this probability estimate can influence the market reaction ("probability-feedback"). Second, cross-sectional differences in deal quality may cause the link. Deals that are perceived better by both the market and the merging companies at the announcement are likely to have higher market reactions and a better chance of consummation, even if there is no learning ("common information"). My method controls these two effects in testing learning.

The paper is organized as follows. Section I introduces the intuition and the formulation of the new method for testing learning. Section II presents three testable hypotheses that differentiate learning from no-learning. The model in Section III substantiates the key claims of Sections I and II. It also motivates control variables in later tests and clarifies the interpretations of the test results.[3] Section IV describes the sample and data. Section V reports the test results. Section VI concludes.

I. The Method

My method for testing learning builds on a regression of the deal completion decision on the stock return around the M&A announcement. I argue that learning does not necessarily occur even if a relation exists in this simple test. The probability-feedback and the common-information hypotheses are the alternative explanations of the relation. In this section, I first explain these two non-learning hypotheses intuitively and then lay out the new method that corrects for them in testing learning.

The first example illustrates the probability-feedback hypothesis. Suppose two deals will each produce \$10 synergy if completed. At the announcement, investors correctly foresee that the probability of consummation is 50% for Deal 1 and 100% for Deal 2. Therefore, the total value change of the two merging companies at the announcement is $\$10 \times 50\% = \$5$ for Deal 1 and $\$10 \times 100\% = \$10$ for Deal 2. In this example, a higher market reaction is associated with a greater deal completion probability without learning because the market's expectation of the completion probability "feeds back" into the price reaction to the announcement.

To solve the probability-feedback problem, I use the market-estimated synergy at the announcement, rather than the stock returns around the announcement, as the independent variable in my tests. I define synergy as the total value change if the two companies will definitely merge. Conceptually, synergy is not a function of the anticipated completion probability. However, synergy is not directly observable and must be estimated before I perform the tests of learning.

The second example illustrates the common-information hypothesis. Common information is what both the merging companies and the market know even if the companies do not learn from the market. In this example, I preclude learning by assuming no information asymmetry between the market and the companies—all information is common information. At the announcement, Deals 1 and 2 are expected to create \$1 and \$100 of value, respectively, if completed. Other parameters of the two deals are the same. Clearly, the quality of Deal 1 is lower, and so the market reaction to its announcement is also lower.

There is always new information between the announcement and the closing, and it can be either good or bad. Suppose the ex ante distribution of the new information is the same for the two deals. Obviously, Deal 1 is more likely to be

[3] It is possible, however, to skip Section III and still understand the intuition of the paper.

cancelled at the closing because a relatively small amount of negative news can bring its expected synergy down below zero. Thus, Deal 1 has a lower market reaction and less chance of being completed. In this example, the quality difference between the two deals at the announcement is common information, and it causes the correlation between the market reaction and the deal completion probability.

I use control samples to correct the effect of common information. For example, I identify a sample in which learning is likely to occur and use a corresponding no-learning sample as the control. If the relation between the market-estimated synergy at the announcement and the deal-closing decision is stronger in the former than in the latter, learning occurs in the former. The intuition is that, if there is no learning, common information in the market synergy estimate is the only cause of the correlation, and its effect should be the same in both samples. A stronger relation in the hypothesized learning sample implies that the market's unique information in its synergy estimate, which includes everything not in the common information, affects the merging companies' later deal-closing decision in that sample. By definition, the companies in that sample learn from the market.

In tests, I use the following probit formula. Section III proves that this formula distinguishes learning from the no-learning probability-feedback and the common-information hypotheses.

$$\begin{aligned} \text{COMPLETE} \sim a_c + \Delta a\ \text{TRAIT} + b_c\ \text{SYNERGY} \\ + \Delta b\ \text{TRAIT} \times \text{SYNERGY} + cQ, \end{aligned} \tag{1}$$

where the deal-closing decision COMPLETE is 1 for completed deals and 0 for cancelled ones, and "$\sim$" represents a probit regression relation. Conceptually, SYNERGY is the expected deal synergy estimated by the market at the announcement, assuming the two companies will definitely merge. Empirically, I estimate SYNERGY before running the test. TRAIT represents the deal characteristics that are hypothesized to co-vary with the merging companies' incentive to learn from the market. In the simplest case, TRAIT is 1 for the learning sample and 0 for the no-learning control sample. Vector Q consists of control variables. The interactive TRAIT $\times$ SYNERGY is the key term in equation (1). If its coefficient Δb is statistically significant and with the predicted sign, no-learning is rejected.

The probit test (1) uses SYNERGY and TRAIT to correct for the probability-feedback and the common-information effects. First, the probability-feedback hypothesis cannot explain any relation between SYNERGY and COMPLETE because the probability of consummation does not affect SYNERGY. Second, TRAIT controls for the common-information effect. In the base case in which TRAIT is 1 for the hypothesized learning sample and 0 for the control sample, no-learning predicts that the relation between SYNERGY and COMPLETE is the same in the two samples, $\Delta b = 0$, and a statistically significant Δb rejects the null. The intuition extends to the case in which TRAIT is continuous. The learning hypothesis predicts that the company's incentive to learn changes in a particular direction with TRAIT, whereas no-learning makes no such prediction. If Δb is significant and has the sign predicted by learning, no-learning is rejected. In fact, TRAIT can be either a scalar or a vector, discrete or continuous.

Before performing the probit test (1), I estimate SYNERGY from the deal announcement return (DCAR), which is the value-weighted average of the bidder's and the target's cumulative abnormal returns (CARs) around the announcement. The estimation builds on two basic ideas. First, I assume that DCAR is the product of SYNERGY and the market-estimated probability of consummation at the announcement (Pr). Second, investors estimate Pr according to SYNERGY and other deal characteristics:

$$\text{SYNERGY}^{(i)} = \text{DCAR}/\text{Pr}^{(i-1)} \quad \left(\text{set Pr}^{(0)} = 1\right), \tag{2}$$

$$\text{COMPLETE} \sim \gamma_c^{(i)} + \Delta\gamma^{(i)}\,\text{TRAIT} + \eta^{(i)}\,\text{SYNERGY}^{(i)} + i^{(i)}\,\text{Q}, \tag{3}$$

$$\text{Pr}^{(i)} = \Phi\left(\gamma_c^{(i)} + \Delta\gamma^{(i)}\,\text{TRAIT} + \eta^{(i)}\,\text{SYNERGY}^{(i)} + i^{(i)}\,\text{Q}\right). \tag{4}$$

In these equations, $\Phi(\cdot)$ is the standard normal *c.d.f.* The interactive variable TRAIT $\times$ SYNERGY is not in (3) or (4) because I assume the null of no-learning in the estimation.[4]

Equations (2)–(4) compose an iterative system to estimate SYNERGY, with superscript i as the iteration step. Only (3) needs estimation, and equations (2) and (4) simply calculate the left-hand variables from the right-hand variables. At step i, equation (2) calculates SYNERGY$^{(i)}$ from Pr$^{(i-1)}$ (initial value Pr$^{(0)}$ = 1). The probit regression (3) estimates the coefficients on SYNERGY$^{(i)}$ and the other independent variables. Equation (4) calculates Pr$^{(i)}$ using the coefficients from equation (3). The process repeats until convergence.

Although alike, the two probit regressions (1) and (3) are conceptually different. Equation (1) investigates what affects the consummation decision, and the vantage point is at the closing. Equation (3) estimates the market's forecast of

[4] Equation (2) is a first-order approximation. It has two potential fallacies. First, SYNERGY is the expected value of deal payoff assuming the merger will definitely happen. However, DCAR reflects the expected payoff of the deal conditional on deal completion, because the market might anticipate cancellation at the announcement:

$$\text{SYNERGY} \equiv E(V_{ST}), \quad \text{and}$$

$$\text{DCAR} = \text{Pr} \times E(V_{ST} \mid \text{COMPLETE} = 1),$$

where V_{ST} is the market's valuation of the deal at the closing ($t = T$). Equation (2) assumes $E(V_{ST}) = E(V_{ST} \mid \text{COMPLETE} = 1)$, which is not true in general. However, if the ex ante deal completion probability is close to 1, then $E(V_{ST}) \approx E(V_{ST} \mid \text{COMPLETE} = 1)$. In my full sample, the completion frequency is 92%. Footnote 11 continues the discussion with empirical tests.

Second, in equation (2), I assume no information leakage before the announcement. Leakage releases information about SYNERGY to the market prior to the announcement and thus reduces the information content of DCAR. Laws and exchange regulations prohibit such leakage. Companies also regularly sign confidentiality agreements at the very beginning of their M&A negotiations. In tests, I use pre-announcement stock price movements to control for possible information leakage.

the consummation probability at the announcement, and the vantage point is at the announcement. In tests, all independent variables in equation (3) must be publicly available at the announcement, but the same is not required for equation (1). For example, the stock price movements between the announcement and the closing can be in equation (1) but not in equation (3).

The expected total value change (SYNERGY) is theoretically a better predictor of the closing decision (COMPLETE) than is the value change of any one side. Both sides must agree to complete a deal, but either side can veto it. If a deal increases the combined value, through efficient negotiation, the two companies can always find terms to split the gain and make each one better off, and the deal is likely to be completed. But if a deal reduces the combined value, at least one side has to suffer a loss, and the deal is likely to be cancelled. Thus, the bidder's and the target's interests are aligned on synergy. Henceforth, I assume frictionless negotiations in M&As, regard the two sides in a deal as one economic agent, and refer to this agent as "the company" or "the manager."

II. Testable Hypotheses

The probit formula (1) determines whether learning occurs in M&As by examining whether the incentive to learn differs among merging companies. Logically, if some companies learn more from the market than others, learning must occur in the former. I argue that the company has a higher incentive to learn (i) when the cost of canceling the announced deal is lower and (ii) when the market has more information that the company does not know. Intuitively, when cancellation is more difficult, the incentive to collect new information is lower because the deal looks more inescapable. Section III proves these claims in a model. Accordingly, I propose three proxies for either the cancellation cost or the informativeness of the market. Based on these proxies, I present three testable hypotheses that relate the incentive to learn to observable M&A characteristics.

First, Luo (2001) claims that the cancellation cost is higher for deals with formal merger agreements at the announcement (agreement deals) than for those without such agreements (pre-agreement deals). Some companies announce their deals after they have signed definitive merger contracts, and others announce before such contracts. It is a breach of contract to unilaterally renege on an agreement deal, and doing so incurs heavy penalties. For example, Getty Oil breached a merger agreement with Pennzoil and, instead, joined Texaco. The court later ordered Texaco to pay Pennzoil more than $10 billion for damages. Besides explicit costs, canceling an agreement deal can severely impair the company's reputation and credibility. In contrast, it is easier to stop a pre-agreement merger. Thus, the M&A announcement timing is a proxy for the cancellation cost.

Second, the opacity of high-tech deals makes the market's opinion less relevant. Verrecchia (1983) and Wagenhofer (1990) argue that a company has the incentive not to disclose information if such a release could hurt its competitive position. Without important raw valuation information, investors' opinions become less informative. Lang and Lundholm (1993) find that less corporate disclosure is associated with lower analyst following, higher dispersion in analyst forecasts, and higher volatility in forecast revisions. I apply these general arguments to high-tech M&As. Proprietary technical information (e.g., new drug information in a biotech merger) is important for the market to evaluate high-tech deals. However, such information is also valuable to the competitors of the merging companies. Therefore, the merging companies closely guard such information, leaving outside investors with little access to it. Additionally, high-tech deals are more likely pioneers of their types. There are fewer similar deals in the past for investors to use as comparisons for deal valuation. As a result, the market is less knowledgeable about high-tech M&As than about other deals.

Third, smaller bidders have less expertise and fewer resources to process public information on themselves. Smaller companies have less managerial talent. For example, compensation should be associated with managerial quality, and Perry and Zenner (2000) show that CEOs of smaller companies are paid much less than their counterparts in S&P 500 companies. Anecdotally, smaller bidders can afford less for in-house M&A analysis or outside investment banking services. Their managers are likely to be less specialized, with their knowledge geared more to production than to finance. Thus, smaller bidders tend to find the market more informative than do larger bidders.

To summarize, learning implies the following three testable hypotheses:

H1: Companies are more likely to learn in pre-agreement deals than in agreement deals.

H2: Companies are more likely to learn in non-high-tech deals than in high-tech deals.

H3: Smaller bidders are more likely to learn than are larger bidders.

III. The Model

In this section, I present a model to analyze learning. Section I argues intuitively that the probit formula (1) solves the "probability-feedback" and the "common-information" problems. Section II assumes that the cancellation cost and the informativeness of the market affect the company's incentive to learn from the market. I prove these claims in this section. The model also yields insights about how to pick control variables for the test (1). The basic setting of the model is broader than M&As. It can help researchers study learning in other corporate events.

The shareholder S represents the market. The manager M represents the two companies of an announced merger. Synergy is the difference between the value of the combined company, if the merger happens, and the total value of the two standalone companies, if they never have a deal. The agents S and M estimate synergy based on available information. Both agents are rational, and S is risk-neutral. The interest rate is zero, so that I ignore the time value of money from the deal announcement to the closing. At any time t, the manager M estimates synergy at $V_{\mathrm{M}t}$, and his utility from the deal is

$$U_t = z + V_{\mathrm{M}t}, \tag{5}$$

where z represents the interest difference between M and S. A merger might bring the bidder manager the prestige of controlling a larger company, or it might trigger a generous golden parachute for the target manager. The managers might also be enamored of the deal. A positive z implies either an agency cost or hubris or both.

Equation (5) is not a typical utility function because the utility U_t is for the manager M, but the wealth change $V_{\mathrm{M}t}$ is for the shareholder S. The equation simply characterizes the relation between the interests of M and S by extending the base case of no agency/hubris problems, in which M maximizes S's wealth by maximizing his own utility ($U_t = V_{\mathrm{M}t}$). This ideal situation is achieved if explicit and implicit compensation and promotion contracts completely align the interests of M and S. With agency/hubris costs, M undertakes the deals that generate enough private benefit/hubris for him but reduce S's wealth. Equation (5) captures the intuition in a simple way—by maximizing his utility U_t, M might accept deals for his private benefit/hubris even when they are bad for S (when $0 > V_{\mathrm{M}t} > -z$).

Figure 1 presents the time line of events. At $t = 1$, M estimates synergy at V_{M1} and decides whether to announce the deal. The announcement happens between $t = 1$ and 2. At $t = 2$, S values synergy at V_{S2}, which is conceptually identical to SYNERGY. At $t = 3$, after observing the market reaction, M values synergy at V_{M3}. All three moments, $t = 1$, 2, and 3, happen around the announcement. I assume that the interval between $t = 1$ and 3 is very short, so that M and S may exchange information in the interim, but no new information is created. However, new information can emerge between $t = 3$ and T. At $t = T$, M estimates synergy at $V_{\mathrm{M}T}$ and makes the closing decision. If $V_{\mathrm{M3}} \neq V_{\mathrm{M1}}$ and the change is influenced by V_{S2} (SYNERGY), M learns from S.

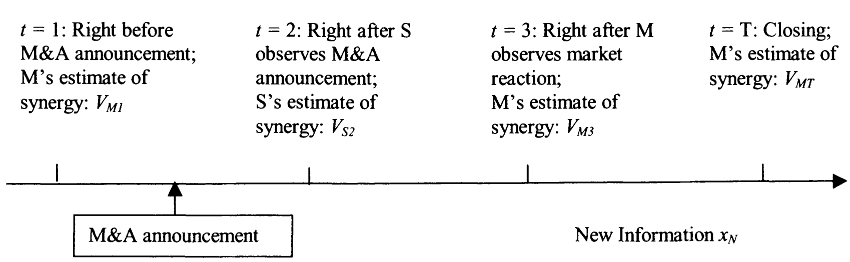

Figure 1. The timeline of events. This figure shows the four time points of interest between the M&A announcement and the closing. The first three moments, $t = 1$–3, are all close to the announcement. I assume that the interval between $t = 1$ and 3 is very short, so that the manager M and the shareholder S may exchange information during the interval, but no new information is created. On the other hand, the time span between $t = 3$ and $t = T$ can be substantial, and new information (x_{N}) may emerge.

The manager M makes both the announcement and the closing decisions. At $t = 1$, M's reservation utility from no deal is normalized at zero. Thus, an announcement happens if and only if $z + V_{\mathrm{M1}} \geq 0$. At $t = T$, M either completes (COMPLETE = 1) or cancels (COMPLETE = 0) the announced deal. The cancellation cost is $C > 0$. Therefore, M expects to get $z + V_{\mathrm{M}T}$ from a consummation and $-C$ from a cancellation. His utility maximization implies:

$$\begin{aligned} &\text{COMPLETE} = 1 \quad \text{if and only if } z + V_{\mathrm{M}T} + C \geq 0 \quad \text{and} \\ &\text{COMPLETE} = 0 \quad \text{otherwise.} \end{aligned} \tag{6}$$

A. Information Distribution

There are four mutually exclusive information sets for the deal, three around the announcement and the fourth between the announcement and the closing. Figure 2 illustrates the information distribution between the manager M and the shareholder S at $t = 2$. M and S share common information, which includes the information M releases at the announcement, for example, the fact that the announcement happens, contract terms, and disclosed business plans. S might have exclusive information that M does not have because the market has strong public information processing power. M might also hold inside information that is not disclosed. In general, learning means that M receives all or part of S's unique information at $t = 3$. For tractability, my model contrasts no-learning

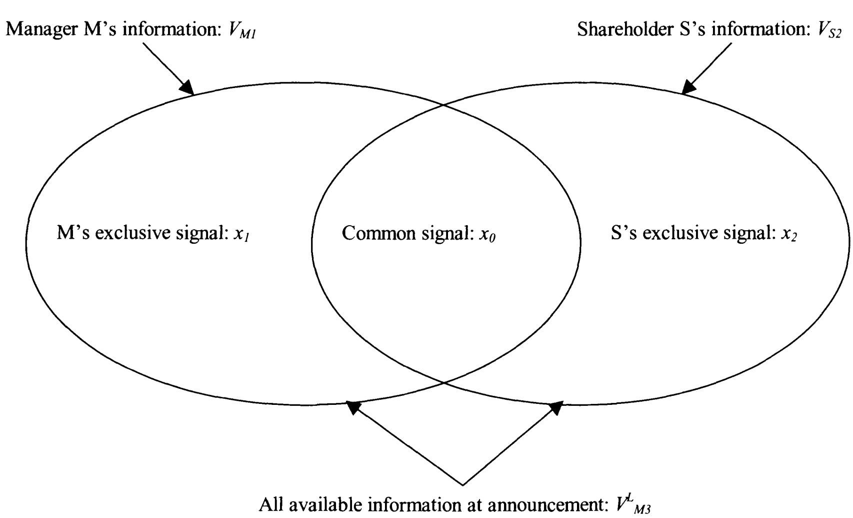

Figure 2. Information distribution right after the M&A announcement. At $t = 2$, the information asymmetry between the manager M and the shareholder S may be bi-directional: each agent may have exclusive information that the other party does not have, and both may share common information. This setting is the most general form of information distribution between two agents. The common signal (x_0) includes the information S receives from M during the announcement. At $t = 3$, if M precisely observes S's exclusive signal (x_2), M's deal valuation is $V^{\mathrm{L}}_{\mathrm{M3}}$.

only with perfect learning, which means M receives all S's unique information at $t = 3$.

The setting of Figure 2 subsumes the traditional belief that the manager is the insider with private information and the shareholder is the outsider whose information is a subset of that of the insider. The old view corresponds to an empty shareholder unique information set. A key insight from Figure 2 is that the manager M has the incentive to learn as long as the shareholder S's unique information set is not empty, no matter how small or large M's inside information set.

I quantify all valuations of the deal by M and S based on the four deal-specific information sets. If an agent knows none of the information sets, I define his estimate of synergy as X, which is the universal cross-sectional mean. If he only has the common information, his valuation is set as $X + x_0$. The variable x_0 is the change caused by the common information, and I refer to x_0 as the "common signal." At $t = 1$, M has both the common information and M's unique information, thus his valuation is

$$V_{\mathrm{M1}} = X + x_0 + x_1. \tag{7}$$

The variable x_1 represents the additional valuation update and is thus called "M's unique signal." Similarly, at $t = 2$, S has both the common information and S's unique information. Her valuation is

$$V_{\mathrm{S2}} = X + x_0 + x_2, \tag{8}$$

and x_2 is "S's unique signal."

Signals x_0, x_1, and x_2 are mutually independent because they correspond to non-overlapping information sets. Rationality of M and S implies that the signals have zero means. This setup is analogous to the "random walk" model of the stock price under EMH, in which the change of the stock price is independent of all past information and has a zero mean because investors are rational and risk-neutral. The estimates of synergy by S and M in my model are like the "price" of the stock in the random walk model, and x_0, x_1, and x_2 are like changes of the "price" responding to new information.

I prove that, at $t = 3$, if the manager M learns from the shareholder S (M obtains all S's unique information), M's estimate of synergy is

$$V_{\mathrm{M3}}^{\mathrm{L}} = V_{\mathrm{M1}} + x_2 = X + x_0 + x_1 + x_2. \tag{9}$$

$V_{\mathrm{M3}}^{\mathrm{L}}$ is the "true" value of the deal at $t = 3$ in the sense that $V_{\mathrm{M3}}^{\mathrm{L}}$ incorporates all available information at that time. The other valuations by M or S based on partial information, V_{M1} and V_{S2}, should be unbiased estimates of $V_{\mathrm{M3}}^{\mathrm{L}}$ with random errors. Furthermore, the errors of V_{M1} and V_{S2} should be independent of each other because they are caused by non-overlapping information sets: V_{M1} does not incorporate S's private information, and V_{S2} does not have M's private information. If $V_{\mathrm{M3}}^{\mathrm{L}} = X + x_0 + x_1 + x_2 + e$, then the error of V_{M1} would be $V_{\mathrm{M3}}^{\mathrm{L}} - V_{\mathrm{M2}} = x_2 + e$, and the error of V_{S2} would be $V_{\mathrm{M3}}^{\mathrm{L}} - V_{\mathrm{S2}} = x_1 + e$. For the two errors to have zero means and be mutually independent, e must equal zero.

If the manager M does not learn, his valuation at $t = 3$ is

$$V_{M3}^{N} = V_{M1} = X + x_0 + x_1. \tag{10}$$

In this model, I use the superscript "L" for learners, the superscript "N" for non-learners, and no-superscript for the general case.

The manager M might also receive new information after the announcement ($t = 3$ to $t = T$), and his deal valuation changes by x_N ("new signal"), which is independent of x_0, x_1, and x_2 and has a zero mean:

$$V_{MT} = V_{M3} + x_N. \tag{11}$$

It is economically reasonable and technically convenient to assume that x_0, x_1, x_2, and x_N follow normal distributions: $x_0 \sim N(0, \Sigma_0)$, $x_1 \sim N(0, \Sigma_1)$, $x_2 \sim N(0, \Sigma_2)$, and $x_N \sim N(0, \Sigma_N)$. According to equation (8), $V_{M1} \sim N(X, \Sigma_M)$ and $V_{S2} \sim N(X, \Sigma_S)$, where

$$\Sigma_M = \Sigma_0 + \Sigma_1 \quad \text{and} \quad \Sigma_S = \Sigma_0 + \Sigma_2. \tag{12}$$

The variances Σ_0, Σ_1, Σ_2, and Σ_N indicate the importance of the corresponding signals. In this model, a signal is deterministic, not random, to the agent who knows it. The variance is for the distribution from which the signal is drawn, not for the signal itself. The larger the variance, the more valuable the signal is ex ante, because knowing the signal resolves a bigger uncertainty. As a special case, a zero variance corresponds to an empty information set.

B. How to Test Learning

I use the model to analyze my test formula (1). The test (1) builds on the intuition that, in the regression of the closing decision (COMPLETE) on the market's estimate of synergy (SYNERGY $= V_{S2}$), the slope is higher for learners than for non-learners. This belief needs a proof. In addition, to prevent no-learning hypotheses from having similar empirical implications, the test (1) might need control variables. Therefore, it is important to understand what determines the slope under no-learning. Proposition 1 addresses both issues.

PROPOSITION 1: *In the probit regression,*

$$\text{COMPLETE} \sim a + b \times \text{SYNERGY}, \tag{13}$$

the slope b is bigger for learners than for non-learners ($b^L > b^N$). Specifically,

$$b^L = \frac{1}{\sigma^L}, \tag{14}$$

$$b^N = \frac{\Sigma_0/\Sigma_S}{\sigma^N}, \tag{15}$$

$$\sigma^L = \sqrt{\Sigma_1 + \Sigma_N}, \quad \text{and} \tag{16}$$

$$\sigma^{\mathrm{N}} = \sqrt{\frac{\Sigma_0}{\Sigma_{\mathrm{S}}}\Sigma_2 + \Sigma_1 + \Sigma_{\mathrm{N}}}. \tag{17}$$

In these equations, Σ_0, Σ_1, Σ_2, Σ_{N}, and Σ_{S} are the variances of the common signal (x_0), the manager inside signal (x_1), the market unique signal (x_2), the new signal (x_{N}), and SYNERGY, respectively; $\Sigma_{\mathrm{S}} = \Sigma_0 + \Sigma_2$; $\sigma^{\mathrm{L}}(\sigma^{\mathrm{N}})$ is the standard deviation of the residual term in the latent OLS model of (13) for learners (non-learners). I assume that the learners precisely observe x_2, but the non-learners do not observe x_2 at all. See Appendix for the proof of this proposition.

In essence, the test formula (1) compares the slope b in equation (13) between a learning sample and a no-learning control sample, and the three testable hypotheses H1–H3 each correspond to a pair of such samples. The coefficient of the interactive term in (1), Δb, is the difference between the learner's slope b^{L} and the non-learner's slope b^{N}. Equation (15) shows that b^{N} is positive if Σ_0 is positive, which is the common-information effect.

Proposition 1 explains why learners have a higher slope than non-learners ($b^{\mathrm{L}} > b^{\mathrm{N}}$). When the manager does not learn, the market unique signal (x_2) in SYNERGY does not influence the closing decision (COMPLETE). In the regression (13), x_2 becomes an "error-in-variable" in SYNERGY. First, the "error" attenuates b^{N} relative to b^{L} (compare the numerators of equations (14) and (15): $1 > \Sigma_0/\Sigma_{\mathrm{S}}$). Second, the "error" reduces the explanatory power of SYNERGY and adds to the residual term of the model (compare the denominators of equations (14) and (15): $\sigma^{\mathrm{L}} \le \sigma^{\mathrm{N}}$). Both effects make the slope b smaller for non-learners than for learners.

Equation (15) shows what affects the slope b in (13) under no-learning. If the residual size, σ^{N}, is constant (e.g., the test includes control variables that are uncorrelated with SYNERGY and capture all the cross-sectional variation of σ^{N}), the slope b^{N} is determined by, and increases in, $\Sigma_0/\Sigma_{\mathrm{S}}$. The ratio $\Sigma_0/\Sigma_{\mathrm{S}}$ measures the importance of the common signal x_0 in SYNERGY. A lower $\Sigma_0/\Sigma_{\mathrm{S}}$ implies that the market has more unique information because the manager knows a smaller part of the market's signal. Thus, the more unique information the market has, the smaller the ratio $\Sigma_0/\Sigma_{\mathrm{S}}$ and the slope b^{N}.

Even if the regression (13) does not include control variables, more market unique information still leads to a lower b^{N}. According to equations (15) and (17), the residual size σ^{N} varies when equation (13) does not have control variables, and b^{N} decreases in σ^{N}. The more market unique information at the announcement, the bigger is σ^{N} and the lower is b^{N} because the market unique signal reduces the explanatory power of SYNERGY and adds to noise in the model under no-learning. See Appendix for a proof.

Therefore, the test (1) does not need control variables for the effect of the market unique information on Δb because learning and no-learning have empirically distinguishable inferences. Under learning, the more unique information the market has, the higher the company's incentive to learn from the market, and the higher the slope b. Under no-learning, more market unique information implies a lower b. Take hypothesis H3 as an example: if smaller bidders learn more from the market than do larger bidders, the slope b is higher for smaller bidders. However, if no company learns, the market's unique information is absent in the closing decision for both groups. The resulting attenuation effect is stronger, and the slope b is lower, for smaller bidders because the market has more unique information for them. Thus, learning and no-learning predict different signs of Δb.

I do not find any compelling reason for deal characteristics to mechanically cause the same effect on Δb as H1, H2, or H3 predicts. Several factors might affect the slope b^{N} in (13) under no-learning. Equation (17) shows that the manager's inside signal (x_1) and the new signal (x_{N}) both affect the residual size σ^{N} and so b^{N}. Beyond the model, if the average deal quality (X), the agency/hubris cost (z), and the cancellation cost (C) vary within the sample of interest or the control sample, their variations might also affect σ^{N} and b^{N}. However, Δb is a difference between the two samples. The variable Δb does not change if the effect of these deal characteristics on b^{N} is the same between the two samples. The focus of test (1) on Δb, rather than on the slope b of any sample, diminishes the need to control for deal characteristics.

Although test (1) does not require any control variable according to my analysis, proper control variables help increase the power of the test by reducing the residual size σ. Equations (16) and (17) show that the manager's private signal (x_1) and the new signal (x_{N}) affect the residual size (both σ^{L} and σ^{N}). Conceptually, x_1 and x_{N} affect the closing decision (COMPLETE) but are not part of the market's estimate of synergy (SYNERGY). I control x_1 and x_{N} with post-announcement stock price movements. The manager inside information at the announcement can later get to the public and affect the stock price. Disclosure laws and stock exchange regulations require merging companies to disclose new deal information in a thorough and timely fashion after the announcement.[5]

C. Who Learns and Who Does Not

I prove that, if it is costly to learn from the market, companies have a higher incentive to learn (i) if the cost to cancel the announced deal is lower or (ii) if the market is considered to have more exclusive information ex ante. These claims underpin my hypotheses H1–H3. The proof contains an innocuous simplification of the model.

Learning from the market incurs a cost on the company, such as top managers' time and effort and the delay in the execution of the merger caused by learning activities. The stock price reaction to the M&A announcement is noisy: it contains both the effect of the announcement and the effects of other, concurrent events. Naively following the price reaction can be as dangerous as ignoring it. To understand the opinions of investors, managers not only pay

[5] Herz and Baller (1971, vol. I. p. 58) comment, "A failure to follow up adequately, once public disclosure has been made of a proposed business combination, can create the same potential liabilities as the initial failure to disclose."

attention to stock price movements but also actively communicate with market participants, for example, by meeting with stock analysts and portfolio managers and analyzing their feedback.

Given the cost of learning, the higher the ex ante value of the market signal, the more likely the manager will learn. The value of the market signal is the maximum price the manager is willing to pay to acquire it. With the assumption that the manager M does not receive new information after the announcement, Proposition 2 presents a closed-form relation among the value of the market signal, the cancellation cost, and the informativeness of the market. Intuitively, the information after the announcement does not alter the effect of the cancellation cost or the effect of the market informativeness on the company's incentive to learn from the market. Ignoring the post-announcement signal makes the algebra tractable.

PROPOSITION 2: *Assuming the new signal, x_N, is not informative ($x_N = 0$), the ex ante value of the market unique signal, x_2, to the manager is*

$$K = \sigma_2 \varphi(-W/\sigma_2) - W\Phi(-W/\sigma_2). \tag{18}$$

The variable K decreases in the cancellation cost and increases in informativeness of x_2:

$$\partial K/\partial C = -\Phi(-W/\sigma_2) < 0, \tag{19}$$

$$\partial K/\partial \sigma_2 = \varphi(W/\sigma_2) > 0, \tag{20}$$

where C is the cancellation cost, σ_2 is the SD of x_2, $\sigma_2^2 = \Sigma_2$, $W = C + z + V_{M1}$, and $\Phi(\cdot)$ and $\varphi(\cdot)$ are the standard normal c.d.f. and p.d.f., respectively.

See the Appendix for a proof.

IV. Sample and Data

I obtain my initial M&A sample from the SDC Domestic M&A Database (SDC). M&A characteristics, including the announcement date, the company identity, and whether the deal is consummated, are mainly from SDC, with minor corrections drawn from Dow Jones Interactive (DJI). The announcement timing—whether a definitive agreement has been signed at the announcement—is hand-collected from news reports in DJI and the "history events" and "history dates" fields in SDC. Stock price data are from the CRSP daily files. Table I records the detailed sample-selection process.

My final sample excludes hostile M&As. I assume that the two sides in the merger plan are cooperative from announcement to closing. They always agree on one estimate of synergy and consummate the deal as long as the estimate does not deteriorate severely. During the period between announcement and closing, I assume that the negotiation is frictionless, and that other deal characteristics remain constant. For example, no new bidders emerge, and target shareholders do not intervene by accepting hostile tender offers. These

Table I

Sample Selection

This table reports my sample-selection process. My initial sample of mergers and acquisitions (M&As) is from the SDC Domestic M&A database (SDC). I also supplement data from SDC with data from the Dow Jones Interactive (DJI). In the table, the first column contains the number of observations after each query step, and the second column describes the query step. To obtain the cumulative abnormal return (CAR) of a stock, I estimate the following market model:

$$R_{it} = a_i + \beta_i R_{Mt} + \varepsilon_{it},$$

where R_{it} is the daily stock return, R_{mt} is the daily return of the value-weighted CRSP market portfolio, and ε_{it} is the daily abnormal return. All stock return data are from CRSP. I estimate the model between trading days −250 to −50 before the M&A announcement. I set the bidder's α at zero because the bidder often outperforms prior to the announcement. The target's α is unconstrained.

No. of Obs. After Query	Query Description
	Machine search in SDC
	SDC Domestic M&As announced between 1/1/1990 and 12/31/1999
36,001	Deal type included: disclosed value M&As, leveraged buyouts, tender offers, and exchange offers
17,557	Percent of shares acquirer is seeking to own after transaction: 50% or higher
17,556	Target nation: United States
5,479	Target is public
4,982	Acquirer nation: United States
3,602	Acquirer is public
3,235	Deal value is $10 million or higher
	Hand collecting data in DJI and SDC
2,569	Exclude:
	(1) Hostile target initial reaction, unilateral proposal, unsolicited offer, deal started with tender or block purchase from a third party
	(2) Competing bid
	(3) Offer is challenged while open
	(4) No news record or unclear record for announcement timing
	Two SDC announcement dates (SDC deal numbers: 514511020 and 390133020) are changed according to DJI
	(5) Announcement is deal rejection or completion or deal is still pending
	(6) Self tender, asset sale, division sale, bankruptcy process, exchange offer, bidder already a majority owner or exercise of an early option
	(7) Multilateral deal, or deal in which parties have multiple equity classes
2,133	Both bidder and target are identified in CRSP, matched by CUSIP, ticker and/or company name, and have [−1, +1] trading-day deal-announcement CARs
2,114	Deals have [−1, +7] trading-day deal-announcement CARs; the full sample

simplifications suit friendly deals reasonably well, but not hostile ones. Therefore, the sample excludes deals with hostile initial target reactions, unsolicited bids, unilateral proposals, and deals that start with tender offers.

The final sample also excludes multiple-bidder deals. The market reaction to each such bid can be contaminated by the other bids, so it is difficult to infer the market's estimate of synergy from the market reaction. Also, the outcome of one bid depends on the other bids. Companies might cancel a bid when a new bid emerges, even if their valuation of the original bid does not change. My model

Table II

Descriptive Statistics of M&A Announcement Returns

This table describes the market reaction to the M&A announcement. I use the cumulative abnormal returns (CARs) of the merging companies around the announcement to represent the market reaction. I define the event window as the trading days [−1, +7] around the announcement. The deal return (DCAR) is the value-weighted average of the bidder's and the target's announcement CARs.

					Correlations with	
Variable	Mean	*SD*	*t*-Statistics	Frequency ≥ 0	Bidder Ann. CAR	Target Ann. CAR
DCAR	0.8%	8.7%	4.3	54.4%	0.86	0.55
Bidder's announcement CAR	−1.6%	9.0%	−8.0	42.4%		0.17
Target's announcement CAR	17.4%	23.3%	34.4	81.5%		

does not suit such multiple-bidder M&As. Therefore, I exclude challenged bids, competing bids, and multilateral deals.[6]

I require that all M&As are by U.S. public companies. To limit data-collection by hand, I set the sample period at 10 years, January 1, 1990 to December 31, 1999 by the announcement date. The deal value must be at least $10 million to reduce measurement errors. I exclude deals in which the parties have multiple equity classes because their price data are often incomplete. I also exclude block purchases from third parties, exercises of earlier acquired options, self tenders, asset sales, division sales, bankruptcies, and exchange offers, because they are not M&As in the traditional sense. The final full sample consists of 2,114 M&As.

Table II describes the market reaction to the M&A announcement. I measure the market reaction with the CARs of the announcing companies over trading days [−1, +7] around the announcement. The −1 trading day controls for possible delays of published news. The +7 trading days allow investors to receive and fully analyze the announced information. According to Table II, almost half of the M&As reduce shareholder value at the announcement. The value-weighted DCAR is positive or flat only in 54% of the observations. The two sides have very different experiences. Only 42% of the bidders have nonnegative announcement returns, whereas 82% of the targets gain. Their return correlation is only 0.17. The bidder return dominates DCAR because the bidder is often much larger than the target. The correlation between DCAR and the bidder return is 0.86.

V. Tests

A. Preliminary: Market Reaction and Deal Completion

Tables III and IV show the relation between the market reaction to the M&A announcement and the later closing decision (COMPLETE). I use either the

[6] If a new bidder emerges while the original bid for a target is still open, the original bid is "challenged", and the new bid is "competing." SDC identifies 214 challenged bids in the sample of 3,235 after my machine search. I include challenged deals in unreported tests, and the conclusions are the same.

Table III

M&A Completion Frequency and the Sign of Market Reaction to the M&A Announcement

This table reports the relation between the sign of the combined deal return (DCAR) and the deal completion frequency. The table shows the relation in the full sample and in seven subsamples: pre-agreement deals (EarlyAnn = 1), agreement deals (EarlyAnn = 0), non-high-tech deals (LowTech = 1), high-tech deals (LowTech = 0), deals with small bidders (\$2.3 million < BMV < \$515.4 million), deals with medium bidders (\$515.8 million < BMV < \$2,664 million), and deals with large bidders (\$2,698 million < BMV < \$505,796 million). If two companies announce their M&A after they have signed a definitive merger contract, the deal is an agreement deal. Otherwise, it is a pre-agreement deal. If both the bidder and the target are high-tech companies, the deal is a high-tech deal. Otherwise, it is a non-high-tech deal. I use the bidder's and the target's high-tech status codes in SDC to judge whether they are high-tech companies. The variable BMV is the bidder's market capitalization at the close of the trading day −41 before the announcement. The three bidder size categories each have the same number of observations. I use the sign rather than the magnitude of DCAR because, intuitively, companies should complete the deals that increase shareholder value, no matter how much the increase is. In the table, the completion frequency is in percentage, and the number of observations is in parentheses. The last row of the table reports the difference in completion frequency between the deals with DCAR ≥ 0 and those with DCAR < 0. The *p*-values of the differences are in italics.

	Full Sample	Announcement Timing		High-Tech Deal Classification		Bidder Stock Market Capitalization		
	(a)	(b) Pre-agreement	(c) Agreement	(d) Non-high-tech	(e) High-Tech	(f) Small	(g) Medium	(h) Large
Unconditional	91.9%	73.7%	95.3%	92.1%	91.2%	87.2%	93.0%	96.2%
(Obs.)	(2114)	(334)	(1780)	(1660)	(454)	(674)	(675)	(675)
DCAR ≥ 0	93.0%	80.7%	95.5%	93.4%	91.7%	89.6%	94.5%	96.0%
(Obs.)	(1149)	(192)	(957)	(933)	(216)	(385)	(364)	(351)
DCAR < 0	90.6%	64.1%	95.1%	90.5%	90.8%	84.1%	91.3%	96.3%
(Obs.)	(965)	(142)	(823)	(727)	(238)	(289)	(311)	(324)
Difference = (DCAR ≥ 0) − (DCAR < 0)	**2.5%****	**16.6%*****	**0.4%**	**2.9%****	**0.9%**	**5.5%****	**3.2%**	**−0.3%**
p-value	*0.038*	*0.001*	*0.714*	*0.033*	*0.733*	*0.033*	*0.105*	*0.848*

***, **, *: Statistically significant at the 1%, 5%, and 10% levels, respectively.

Table IV

Probit Regressions of the M&A Closing Decision on the Market Reaction to the M&A Announcement

This table reports nine probit regressions of the M&A closing decision on the market reaction to the M&A announcement controlling for deal characteristics. The dummy COMPLETE equals 1 for completed deals and 0 for cancelled deals. Each regression corresponds to a specific sample and model combination. I measure the bidder's and the target's post-announcement CARs from the trading day +8 after the announcement to the trading day −2 before the closing. I measure the two companies' pre-announcement CARs over the trading days [−40, −2] before the announcement. I obtain from SDC the data for Deal value, whether the bidder and the target are headquartered in the same state, and whether the exchange medium is pure cash. For each test, the number of observations used in the regression is in parentheses, and the p-values of the coefficients are in italics.

	Full Sample—Two Tests		Announcement Timing		High-Tech Deal Classification		Bidder Equity-Market Capitalization		
	(a)	(b)	(c) Pre-Agreement	(d) Agreement	(e) Non-High-Tech	(f) High-Tech	(g) Small	(h) Medium	(i) Large
Constant	1.53***	1.52***	0.770***	1.80***	1.53***	1.57***	1.08***	1.76***	2.11***
	0.000	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
Deal announcement CAR (DCAR)	**1.35*** **		**4.20*** **	**0.618**	**2.04*** **	**0.204**	**1.51** **	**2.37** **	**−0.544**
	0.005		***0.000***	***0.304***	***0.001***	***0.806***	***0.014***	***0.012***	***0.696***
Bidder's announcement CAR		**1.02** **							
		0.027							
Target's announcement CAR		**0.168**							
		0.390							
Bidder's post-announcement CAR	0.611***	0.616***	1.07**	0.663**	0.765***	−0.121	1.12***	0.291	0.283
	0.006	*0.006*	*0.033*	*0.016*	*0.002*	*0.816*	*0.004*	*0.522*	*0.512*
Target's post-announcement CAR	1.22***	1.22***	1.00**	1.21***	1.39***	1.18**	0.500	2.59***	1.24*
	0.000	*0.000*	*0.027*	*0.003*	*0.000*	*0.027*	*0.212*	*0.000*	*0.090*
Bidder's pre-announcement CAR	0.545*	0.514*	0.773	0.733*	0.521	0.652	0.802**	0.778	0.284
	0.071	*0.087*	*0.131*	*0.092*	*0.169*	*0.203*	*0.046*	*0.213*	*0.751*
Target's pre-announcement CAR	−0.150	−0.151	−0.357	−0.119	−0.0990	−0.327	−0.0828	−0.233	−0.688
	0.406	*0.409*	*0.252*	*0.632*	*0.653*	*0.318*	*0.736*	*0.531*	*0.131*
Deal value (billion $)	0.00276	0.00394	0.136	−0.0111	0.0286	−0.0115	1.01**	0.252	−0.00151
	0.816	*0.747*	*0.464*	*0.263*	*0.322*	*0.447*	*0.037*	*0.153*	*0.908*
Deal value/BMV	−0.123***	−0.114***	−0.133**	−0.111***	−0.128***	−0.162	−0.115***	−0.618**	−0.831***
	0.000	*0.000*	*0.011*	*0.001*	*0.000*	*0.113*	*0.003*	*0.011*	*0.000*
Bidder and target headquartered in same state (1 = yes)	0.250**	0.254**	0.0278	0.510***	0.264**	0.146	0.277**	0.514**	0.702*
	0.014	*0.012*	*0.878*	*0.008*	*0.022*	*0.513*	*0.044*	*0.017*	*0.086*
Pure cash deal (1 = yes)	−0.228**	−0.228**	−0.0460	−0.332**	−0.114	−0.552***	0.259	−0.231	−0.227*
	0.029	*0.029*	*0.824*	*0.011*	*0.356*	*0.007*	*0.109*	*0.244*	*0.332*
(Obs.)	(2004)	(2004)	(302)	(1702)	(1574)	(430)	(666)	(669)	(669)

***, **, *: Statistically significant at the 1%, 5%, and 10% level, respectively.

combined DCAR or individual bidder and target announcement returns to represent the market reaction. Under learning, these returns predict COMPLETE. The hypotheses H1–H3 further propose that the predictive power varies with certain deal characteristics. Accordingly, I examine the relation in the full sample as well as in seven subsamples related to H1–H3. Table III reports univariate tests, and Table IV contains multiple regressions. The results are consistent with learning. However, such evidence is inconclusive because the tests do not fully control for the probability-feedback and the common-information effects.

Table III shows that the deal return (DCAR) predicts the closing decision (COMPLETE). For the full sample, the completion frequency is 93% when DCAR ≥ 0, but only 90.6% when DCAR < 0. The difference is significant at the 5% level. Companies consummate M&As more often when DCAR is nonnegative than when it is negative. Intuitively, the difference in the completion rate between DCAR ≥ 0 and DCAR < 0 measures the predictive power of DCAR on COMPLETE. Columns (b)–(h) show that the difference is larger in pre-agreement deals than it is in agreement deals, in non-high-tech deals than in high-tech deals, and in deals with small bidders than in those with large bidders. The results are consistent with H1–H3.

Luo (2001) claims that most announced M&As in the 1990s are with definitive contracts, and retracting such a deal is often prohibitively expensive. The high penalty makes cancellations rare, regardless of the market reaction. In my full sample, 91.9% of M&As are completed.[7] Nevertheless, how the market reacts to the announcement contains substantial information about the likelihood of cancellation. In Table III, the sign of DCAR makes a 2.5% difference in the 8.1% (=1 − 91.9%) overall cancellation rate for the full sample. In subsamples in which I hypothesize learning, the spread is 16.6% in the 26.3% (=1 − 73.7%) cancellation rate for pre-agreement deals, 2.9% in 7.9% (=1 − 92.1%) for non-high-tech deals, 5.5% in 12.8% (=1 − 87.2%) for deals with small bidders, and 3.2% in 7.0% (=1 − 93%) for deals with medium bidders.

Table IV reports probit regressions of the deal-closing decision (COMPLETE) on the market reaction to the announcement controlling for deal characteristics. Both tests (a) and (b) are for the full sample, and tests (c)–(i) are for the seven subsamples related to hypotheses H1–H3. Consistent with learning, the market reaction significantly predicts COMPLETE in the full sample. Across the subsamples, the effect of the market reaction on COMPLETE varies in patterns that are consistent with H1–H3.[8]

According to Table IV test (a), the deal return (DCAR) predicts COMPLETE at the 1% level in the full sample. According to test (b), the bidder's announcement return predicts COMPLETE at the 5% level in the full sample, and the target's announcement return is not significant. The pattern is consistent with

[7] The high completion frequency seems specific to my sample period and sample selection. Dodd (1980) reports 80 completed deals among 151 M&As in the 1970s. Jennings and Mazzeo (1991) report 351 completed ones in 472 deals in the 1980s.

[8] Jennings and Mazzeo (1991) do not find such a relation. Their total sample contains 472 M&As, and the biggest sample size of their regressions is 188. Their event window spans 11 or 21 days and always ends on the announcement day.

my argument that the combined deal return is a better determinant of the closing decision than the returns of the individual companies. Conceptually, the gains or losses of the bidder and the target are contaminated by their bargaining positions, which are not central to the closing decision. Thus, I often use DCAR, rather than the bidder's or the target's return, to represent the market reaction.

I regress COMPLETE on DCAR within the seven subsamples in the tests (c)–(i) of Table IV. The coefficient on DCAR is higher, and the corresponding p-value is lower, for pre-agreement deals than for agreement deals, for non-high-tech deals than for high-tech deals, and for deals with smaller bidders than for those with larger bidders. Although the coefficient is not monotonic in bidder size across the tests (g)–(i), it is still higher for small and medium bidders than for large bidders. Overall, the evidence is consistent with the learning hypotheses H1–H3.

B. Tests Controlling for Probability-Feedback and Common-Information

Table V tests the hypotheses H1–H3 controlling for the probability-feedback and the common-information effects with the probit formula (1). Among the independent variables, SYNERGY is the market's valuation of the deal at the announcement and is estimated prior to the test according to equations (2)–(4). SYNERGY is often similar to the DCAR because the estimated completion probability (Pr) in equation (2) is often close to 1. The rank correlation between SYNERGY and DCAR is 0.98. EarlyAnn is 1 for pre-agreement deals and 0 for agreement deals. LowTech is 1 for non-high-tech deals and 0 for high-tech deals. BMV is the bidder's market capitalization on trading day -42 before the announcement.

Table V includes control variables for the sake of caution as well as for higher test power. The stock returns before the M&A announcement capture information leakage, and the returns afterward are the proxy for post-announcement information. The dollar size of the deal and its ratio to the bidder's market capitalization control for any size effect. The two other control variables—if the bidder and the target are headquartered in the same state and if the exchange medium is pure cash—significantly predict the closing decision, and I include them to increase the test power.[9,10]

Tests (a)–(c) of Table V examine H1–H3 individually. Consistent with these learning hypotheses, the interactive terms EarlyAnn $\times$ SYNERGY, LowTech $\times$ SYNERGY, and BMV $\times$ SYNERGY are all significant at the 5% level. SYNERGY affects the closing decision (COMPLETE) more in pre-agreement deals than in agreement deals, in non-high-tech deals than in high-tech deals,

[9] Luo (2002) reports regression results with no controls and with only post-announcement stock returns as controls, and the conclusions are the same. Luo (2002) also proposes and tests an extra hypothesis.

[10] Jennings and Mazzeo (1991) also include the bidder's pre-offer toehold position in the target as an independent variable. Toehold has no implication in my theory. I include toehold in unreported tests; it is not significant, and including it does not change my conclusions.

Table V

Tests of Learning Controlling for the Probability-Feedback and the Common-Information Problems

This table reports five probit regressions. The dependent variable is always COMPLETE, which equals 1 for completed deals and 0 for cancelled deals. The variable SYNERGY is the proxy for the market's estimate of synergy at the announcement assuming the deal will definitely be completed. In the tests (a)–(d), I estimate SYNERGY from the deal announcement return (DCAR) with the following iterative process:

$$\text{SYNERGY}^{(i)} = \text{DCAR}/\text{Pr}^{(i-1)} \quad (\text{set Pr}^{(0)} = 1), \tag{2}$$

$$\text{COMPLETE} \sim \gamma_c^{(i)} + \Delta\gamma^{(i)}\,\text{TRAIT} + \eta^{(i)}\,\text{SYNERGY}^{(i)} + \iota^{(i)}\,\text{Q}, \tag{3}$$

$$\text{Pr}^{(i)} = \Phi\left(\gamma_c^{(i)} + \Delta\gamma^{(i)}\,\text{TRAIT} + \eta^{(i)}\,\text{SYNERGY}^{(i)} + \iota^{(i)}Q\right). \tag{4}$$

In these equations, $\Phi(\cdot)$ is the *c.d.f.* of the standard normal distribution. The sign "$\sim$" represents a probit regression relation. The variable TRAIT = {EarlyAnn, LowTech, BMV}; Q = {Bidder's pre-announcement CAR, Deal value, Deal value/BMV, Bidder, and target headquartered in same state (1 = yes), Pure cash offer (1 = yes)}; $\Delta\gamma$ and ι are also vectors. The superscript i is for the iteration step. Equation (2) calculates SYNERGY$^{(i)}$ from DCAR and Pr$^{(i-1)}$ (initial Pr$^{(0)}$ = 1). Equation (3) gives the probit regression coefficients. Equation (4) calculates Pr$^{(i)}$ using the coefficients from equation (3). Only equation (3) needs estimation. Equations (2) and (4) simply calculate the left-hand variables. During the iteration, I censor Pr$^{(i)}$ at the lower end at 0.00001 to avoid the division of zero. I also censor SYNERGY$^{(i)}$ at −100% and +100%. The process continues until convergence.
For tests (a)–(d), at convergence ($i = 20$):

$$
\begin{aligned}
&\text{SYNERGY} = \text{DCAR}/\,\text{Pr}, \quad \text{and}\\
&\text{Pr} = \Phi\{1.49230 - 1.07112\ \text{EarlyAnn} + 0.23564\ \text{LowTech} + 0.01485\ \text{BMV(billion \$)}\\
&\quad + 1.46144\ \text{SYNERGY} - 0.02849\,(\text{Deal value billion \$}) - 0.10608\,(\text{Deal value/BMV})\\
&\quad + 0.74705\,(\text{Bidder's pre-announcement CAR})\\
&\quad + 0.29232\,(1 \text{ for bidder \& target headquartered in same state, 0 otherwise})\\
&\quad - 0.22244\,(1 \text{ for pure cash deals, 0 otherwise})\}.
\end{aligned}
$$

For test (e), I substitute DCAR for SYNERGY:

SYNERGY = DCAR.

The number of observations for the tests are in parentheses, and the p-values of coefficients are in italics.

	Test (a)[a]	Test (b)[a]	Test (c)[a]	Test (d)[a]	Test (e)[b]
Constant	1.55***	1.54***	1.54***	1.54***	1.54***
	0.000	*0.000*	*0.000*	*0.000*	*0.000*
EarlyAnn (1 = pre-agreement deals)	−1.07***	−1.08***	−1.07**	−1.07**	−1.09**
	0.000	*0.000*	*0.000*	*0.000*	*0.000*
LowTech (1 = non-high-tech deals)	0.270**	0.291**	0.272**	0.276**	0.274**
	0.018	*0.010*	*0.017*	*0.015*	*0.016*
BMV (billion $)	0.0149**	0.0148**	0.0184***	0.0182***	0.0191***
	0.024	*0.025*	*0.008*	*0.009*	*0.008*
SYNERGY	0.623	0.0821	1.71***	−1.08	−0.162
	0.241	*0.908*	*0.000*	*0.890*	*0.854*
EarlyAnn × SYNERGY	**2.36****			**1.88****	**3.16****
	0.010			***0.046***	***0.014***

(continued)

Table V—*Continued*

	Test (a)[a]	Test (b)[a]	Test (c)[a]	Test (d)[a]	Test (e)[b]
LowTech × SYNERGY		**2.08****		**1.72***	**1.74****
		0.018		***0.058***	***0.010***
BMV × SYNERGY			**−0.170****	**−0.140***	**−0.153***
			0.034	***0.095***	***0.089***
Bidder's post-announcement CAR	0.708***	0.734***	0.717***	0.716***	0.705***
	0.003	*0.002*	*0.003*	*0.003*	*0.003*
Target's post-announcement CAR	1.17***	1.11***	1.11***	1.14***	1.16***
	0.000	*0.000*	*0.000*	*0.000*	*0.000*
Bidder's pre-announcement CAR	0.758**	0.743**	0.778**	0.777**	0.777**
	0.018	*0.020*	*0.014*	*0.016*	*0.015*
Target's pre-announcement CAR	−0.198	−0.178	−0.173	−0.192	−0.208
	0.307	*0.356*	*0.367*	*0.322*	*0.282*
Deal value ($billion)	−0.0312**	−0.0312**	−0.0288**	−0.0299**	−0.0313**
	0.017	*0.018*	*0.022*	*0.021*	*0.015*
Deal value/BMV	−0.108***	−0.121***	−0.119***	−0.117***	−0.109***
	0.002	*0.000*	*0.000*	*0.000*	*0.000*
Bidder and target headquartered	0.325***	0.315***	0.328***	0.328***	0.335***
in same state (1 = yes)	*0.004*	*0.005*	*0.003*	*0.004*	*0.003*
Pure cash offer (1 = yes)	−0.242**	−0.222**	−0.232**	−0.237**	−0.234**
	0.031	*0.049*	*0.038*	*0.035*	*0.037*
(Obs.)	(2004)	(2004)	(2004)	(2004)	(2004)

***, **, *: Statistically significant at the 1%, 5%, and 10% level, respectively.
[a]SYNERGY = DCAR/Pr, estimated iteratively according to equations (2)–(4).
[b]SYNERGY = DCAR.

and in deals with smaller bidders than those with larger ones. BMV in test (c) is a continuous variable. In essence, I test learning in deals with smaller bidders and use those with larger bidders as the control. The learning hypothesis H3 predicts that BMV × SYNERGY has a negative sign.[11,12]

[11] The completion frequency of pre-agreement deals is by far the lowest (74%) among the seven subsamples. This potential distortion caused by the approximation of $E(V_{ST}) = E(V_{ST} \mid \text{COMPLETE} = 1)$ in equation (2) should be the most severe (see footnote 4 for a theoretical discussion). I here re-test H1 adjusting the potential distortion of the approximation. The conclusion is unchanged.

This footnote assumes (i) between the announcement and the closing, the changes in the market's and the company's estimates of synergy are perfectly correlated, and (ii) the companies' post-announcement CARs are all driven by the change in the investor's deal valuation. Both assumptions tend to exaggerate the possible distortion effect of the approximation. For example, if the manager's and the investor's deal valuations are mutually independent, the approximation is correct. Under these assumptions,

$$\text{SYNERGY} = \text{DCAR/Pr} - \sigma_N \frac{\varphi[\Phi^{-1}(\text{Pr})]}{\text{Pr}},$$

where σ_N is the standard deviation of the combined bidder and target post-announcement CAR between the announcement and the closing. In a revised test (a) of Table V with this new SYNERGY estimate, the term EarlyAnn × SYNERGY is still significant at the 5% level. See Luo (2002) for details.

[12] Interestingly, H1 coincides with casual observations. Quaker Oats and Snapple announced their deal after they had signed a definitive merger agreement. Both the deals between Alcatel and Lucent and between Hewlett-Packard and PricewaterhouseCoopers were announced before such a contract.

The announcement timing (EarlyAnn) is a decision of the company and can be influenced by the anticipation of learning. However, the interpretation of Table V test (a) is not consequently changed because learning and the no-learning predict opposite signs of EarlyAnn × SYNERGY. According to Luo (2001), companies choose early announcement with less commitment (EarlyAnn = 1) when the market possesses more unique information. If learning does not occur, according to Proposition 1, the probit slope will be lower for EarlyAnn = 1 than for EarlyAnn = 0 because more market unique information implies a stronger attenuation effect on the probit slope. Therefore, no-learning implies that EarlyAnn × SYNERGY will have a negative sign, which is against the evidence.

Tests (d) and (e) of Table V investigate the marginal effects of the hypotheses H1–H3 over each other, although testing the occurrence of learning does not require the three hypotheses to be mutually exclusive.[13] Test (d) includes EarlyAnn, LowTech, BMV, and their interactions with SYNERGY as regressors. All three interactive variables have the predicted signs and are statistically significant at the 10% level, consistent with H1–H3 jointly. To ensure that the results are not an artifact of the estimation process of SYNERGY, I substitute the DCAR for SYNERGY in test (e). In equation (2), SYNERGY would equal DCAR if the forecasted deal completion probability (Pr) were one. Again, all three interactive variables with DCAR have the predicted signs and are significant at the 10% level, consistent with learning.

In summary, the learning hypothesis predicts that the coefficients on the interactive variables EarlyAnn × SYNERGY, LowTech × SYNERGY, and BMV × SYNERGY are different from zero and with particular signs, whereas the no-learning hypothesis does not have such implications. Table V is consistent with learning. The results indicate that the announcement timing, the high-tech status, and the bidder size seem to influence the company's incentive to learn from the market during M&As.[14]

C. Possible Learning in Renegotiations

In addition to the closing decision, merging companies can also apply what they learn from the market to the decision of renegotiation. For example, the

[13] Luo (2001) shows that the effects of H1–H3 are correlated. For example, larger bidders are more likely to have agreement deals.

[14] An alternative story is that companies in agreement deals, high-tech deals, and deals with large bidders learn less from the market because they suffer more agency/hubris problems (higher z). If hubris is a trait of personality, as Bertrand and Schoar (2002) claim, we have no reason to believe that the managers in the aforementioned subsamples are more or less influenced by hubris than the others. If hubris is a function of the compensation environment, as hinted by Jensen and Meckling (1976), we would expect that the hubris problem is less severe when the manager's interest is more tied to the investor's interest. With data from COMPUSTAT ExecuComp, which is available for only a small part of my full sample, I compare the equity-based-compensation and the stock ownership of the bidder CEO as well as of the bidder top five executives between the corresponding subsamples. I find no evidence for the alternative story. The results are available on request.

bidder and the target might agree to reduce the offer price after they find out from investors that the original bid benefits the target too much at the expense of the bidder, and vice versa. Ultimately, learning implies that investors have unique information, and that companies acquire and utilize the information in all M&A decisions. The learning hypothesis would be less credible if companies do not learn in renegotiations.

I find evidence consistent with learning in M&A renegotiations. The bidder's return around the announcement predicts the later offer revision decision. My tests show that, the lower the bidder's announcement return, the more likely the offer price is revised down, and vice versa. Jennings and Mazzeo (1991) do not find such a relation. Their results might be specific to their sample and test design.

As a caveat, my results on renegotiations do not completely rule out other explanations, such as the probability-feedback and the common-information stories. A thorough investigation of learning in renegotiations is out of the reach of my theory and therefore out of the scope of this paper. Renegotiations are inherently a bargaining issue, but both my theoretical model and my test formula (1) focus on synergy rather than bargaining. In the following tests, I use the bidder's and the target's returns around the announcement, rather than SYNERGY or the combined deal return (DCAR), as independent variables because of the bargaining nature of renegotiations.

I categorize an ex post consummated deal as "renegotiated" if SDC records both a change in the bid price and a change in contract terms between the announcement and the consummation. There are 94 such renegotiated deals, with the offer price revised up in 42 deals and down in the other 52. All the renegotiated deals are announced after January 1995. To avoid potential sample-selection biases by SDC, I exclude deals before January 1, 1995 from all the tests of this subsection. In my full sample, 1,561 deals are announced on or after January 1, 1995, and 1,458 of them are completed. I exclude cancelled deals from the tests because I cannot determine which side initiates the cancellation.[15]

Table VI shows the relations between the renegotiation decision and deal characteristics. Column (a) reports pair-wise correlations, and (b) reports a probit regression with the dependent variable as 1 for renegotiated deals and 0 for the others. Among the deal characteristics, only the announcement timing predicts the renegotiation decision, and the relation is strongly significant in both columns. Understandably, offers without a definitive contract are more likely to be revised.

I investigate if the market reaction predicts the renegotiation decision in Table VII. Test (a) is an ordered probit regression that includes all completed deals. The dependent variable is 1 for increased offers, 0 for not-renegotiated deals, and -1 for reduced offers. The three levels create two intercepts in the test. I do not include deal characteristics in test (a) because they are not

[15] SDC often records a change in contract terms as "terms amended" in its History Event column. In collecting the renegotiated sample, I manually excluded those changes in the SDC offer price that are caused by rounding errors.

Table VI

Renegotiation and M&A Deal Characteristics

This table reports the relations between the renegotiation decision and M&A deal characteristics. The sample consists of all completed deals announced between January 1, 1995 and December 31, 1999. The sample period has a later start point to avoid possible sample-selection problems in the renegotiation data from SDC. An M&A is "renegotiated" if (i) it is eventually completed, (ii) SDC reports that the contract terms are amended, and (iii) its final offer price is different from its initial offer price. The first column reports the correlations between the renegotiation dummy (one for renegotiated deals and zero for others) and deal characteristic variables. The second column reports a probit regression of the renegotiation decision on the deal characteristics. The number of observations for the tests are in parentheses, and the p-values of coefficients are in italics.

	(a) Correlations Variable = 1: Renegotiated Deals, 0: Others	(b) Probit Regression Dependent Variable = 1: Renegotiated Deals, 0: Others
Constant		−1.64***
		0.000
EarlyAnn (1 = pre-agreement deals)	0.12***	0.668***
	0.000	*0.000*
LowTech (1 = non-high-tech deals)	0.04	0.095
	0.156	*0.483*
BMV (billion $)	−0.02	−0.00024
	0.520	*0.923*
Deal value (billion $)	−0.03	−0.034
	0.258	*0.309*
Deal value/BMV	−0.02	−0.065
	0.413	*0.489*
Bidder and target headquartered in same state (1 = yes)	−0.01	−0.059
	0.667	*0.621*
Pure cash deal (1 = yes)	0.02	0.140
	0.361	*0.307*
Price collar (1 = yes)	0.03	0.172
	0.292	*0.228*
(Obs.)	(1458)	(1390)

***, **, *: Statistically significant at the 1%, 5%, and 10% level, respectively.

monotonically related to the dependent variable. For example, a pre-agreement announcement implies that the dependent variable is more likely to be either 1 or −1 than to be 0. In test (a), the bidder announcement return is positive and significant at the 10% level. The target announcement return is not significant. The evidence is consistent with learning. If investors believe that the bidder overpays (underpays) at the original offer, the bidder's announcement return is lower (higher), and the offer is likely to be revised down (up).

Table VII test (b) examines the relation between the market reaction and the renegotiation decision only within the deals that are renegotiated ex post. The dependent variable of the probit regression is 1 for increased offers and 0 for reduced offers. At the announcement, the market likely anticipates the difference in the revision probability between the ex post renegotiated deals and

Table VII

Renegotiation and Market Reaction

This table reports two probit regressions of the renegotiation decision on the market reaction to the M&A announcement. Test (a) is an ordered probit test with a three-level dependent variable, which is 1 for renegotiated deals with increased offers, 0 for not-renegotiated deals, and −1 for renegotiated deals with decreased offers. Test (a) has two intercepts because of the three levels of the dependent variable. The sample of test (a) consists of all completed deals announced between January 1, 1995 and December 31, 1999. Test (b) is an ordinary probit test with a two-level dependent variable, which is 1 for renegotiated deals with increased offers and 0 for renegotiated deals with decreased offers. The sample of test (b) only consists of completed and renegotiated deals announced between January 1, 1995 and December 31, 1999. Both tests include 94 renegotiated deals, among which the offer price increases in 42 deals and decreases in the other 52 deals. In both tests, I use the bidder's and the target's announcement CARs to represent the market reaction to the M&A announcement. The number of observations for the tests are in parentheses, and the p-values of coefficients are in italics.

Dependent Variable	Test (a)[a]	Test (b)[b]
Intercept	−1.94***	−0.0762
	0.000	*0.690*
Intercept2	3.81***	
	0.000	
Bidder's Announcement CAR	**0.895***	**9.01*****
	0.089	***0.001***
Target's Announcement CAR	**0.0414**	**1.45**
	0.846	***0.104***
Bidder's post-announcement CAR	0.588**	3.23**
	0.014	*0.016*
Target's post-announcement CAR	0.793***	5.03***
	0.004	*0.001*
Bidder's pre-announcement CAR	0.0214	0.0363
	0.945	*0.975*
Target's pre-announcement CAR	−0.0587	0.0846
	0.775	*0.916*
Sample	All completed deals	Completed and renegotiated deals
(No. of obs.)	(1457)	(94)

***, **, *: Statistically significant at the 1%, 5%, and 10% level, respectively.

[a]1: Offer higher, 0: No change, −1: Offer lower.

[b]1: Offer higher, 0: Offer lower.

the ex post not-renegotiated ones, and the price reaction can be more subdued for the former. Therefore, mixing the two samples together, as in test (a), reduces the power of the test. Consistent with the analysis, the bidder announcement return is significant at the 1% level in test (b), stronger than that in test (a). The target's announcement CAR is positive with a p-value of 10%. Again, the evidence is consistent with learning in renegotiations.

VI. Conclusions

The strength of the market economy lies in its efficient use of resources, which include information. The stock market price is often seen as the most informationally efficient indicator of company value, and the corporation as the most productive element in the economy. Thus, researchers have come to ask: Does the efficiency of the capital market contribute to the efficiency of the company? However, the finance literature has failed to offer convincing evidence on whether or how the company utilizes the market information to improve investment decisions.

I show that companies seem to learn from the market during M&As. The merging companies appear to extract information from the market reaction to the announcement and apply that information to the closing decision. In a sample of more than 2,000 domestic M&As announced in the 1990s, the combined bidder and target abnormal return around the announcement predicts whether the companies will later consummate the deal. The relation is not fully explained by non-learning hypotheses, such as the probability-feedback and the common-information effects. In robustness tests, I also find that the bidder's abnormal return around the announcement predicts the deal renegotiation decision. The evidence, taken in its entirety, favors the learning hypothesis.

Furthermore, merging companies respond to the market reaction differently and in predictable patterns. Their deal completion decision is more sensitive to the market's opinion at the announcement when (i) the companies announce the deal before signing a definitive contract, (ii) they are not in high-tech industries, and (iii) the bidder is small- or mid-cap. The evidence is most consistent with the interpretation that the companies are more likely to learn when the cost of reneging is lower or when the market has more information that the companies do not know. The announcement timing is correlated with cancellation cost. It is easier to cancel a deal without a contract than one with a contract. Deal high-tech status and bidder size appear to be proxies for the informativeness of the market. The market should be less informative about high-tech deals because those deals are less transparent to investors. Smaller bidders should find the market more informative because these bidders do less deal valuation by themselves.

Looking forward, I expect more definitive evidence on learning to emerge, not only in M&As, but also in other areas of corporate finance. After all, the market's pursuit of better information is universal, and the company's pursuit of better investment policies is also universal. The model and the test process of this paper can also be applied to other studies of learning. In particular, the method to control for the probability-feedback and the common-information problems only require two sequential public actions by the company on the same issue (e.g., announcement and implementation) and the market's reaction to the first action. Such examples include stock repurchases, stock issuances, corporate strategy changes, and the appointment of top managers.

Appendix

LEMMA 1: *The probit regression model*

$$\text{COMPLETE} \sim a + b \times \text{SYNERGY}, \tag{13}$$

has the following latent OLS regression model:

$$V_{MT} = a_{\mathrm{OLS}} + b_{\mathrm{OLS}} \times \mathrm{SYNERGY} + \varepsilon. \tag{A1}$$

Their coefficients have the following relations:

$$a_{\mathrm{OLS}} + z + C = \sigma a, \tag{A2}$$

and

$$b_{\mathrm{OLS}} = \sigma b, \tag{A3}$$

where SYNERGY is V_{S2} *in the model, and* σ *is the SD of* ε.

Proof of Lemma 1: According to equation (6):

$$\mathrm{Prob}(\mathrm{COMPLETE} = 1) = \mathrm{Prob}(V_{MT} + z + C \geq 0). \tag{A4}$$

According to equation (A1):

$$\begin{aligned}\mathrm{Prob}(V_{MT} + z + C \geq 0) &= \mathrm{Pr}(\varepsilon \leq a_{\mathrm{OLS}} + b_{\mathrm{OLS}} V_{\mathrm{S2}} + z + C) \\ &= \Phi[(a_{\mathrm{OLS}} + z + C + b_{\mathrm{OLS}} V_{\mathrm{S2}})/\sigma],\end{aligned} \tag{A5}$$

where $\Phi(\cdot)$ is the standard normal *c.d.f.* According to equation (13):

$$\mathrm{Prob}(\mathrm{COMPLETE} = 1) = \Phi(a + bV_{\mathrm{S2}}). \tag{A6}$$

Combining (A4)–(A6),

$$\Phi[(a_{\mathrm{OLS}} + z + C + b_{\mathrm{OLS}} V_{\mathrm{S2}})/\sigma] = \Phi(a + bV_{\mathrm{S2}}). \tag{A7}$$

Q.E.D.

Lemma 1 shows that the latent OLS model of (13) has V_{MT} as its dependent variable. The variable V_{MT} is M's valuation of the deal at the closing.

Proof of Proposition 1: According to equations (10) and (11), if the manager M does not learn from the shareholder S,

$$V_{MT}^{\mathrm{N}} = X + x_0 + x_1 + x_{\mathrm{N}}. \tag{A8}$$

Under no-learning, the coefficients in the latent OLS model (A1) are

$$b_{\mathrm{OLS}}^{\mathrm{N}} = \frac{\mathrm{Cov}\left(V_{MT}^{\mathrm{N}}, V_{\mathrm{S2}}\right)}{\mathrm{Var}\left(V_{\mathrm{S2}}\right)} = \frac{\Sigma_0}{\Sigma_{\mathrm{S}}}, \tag{A9}$$

$$a_{\mathrm{OLS}}^{\mathrm{N}} = E\left(V_{MT}^{\mathrm{N}}\right) - b_{\mathrm{O}}^{\mathrm{N}} E(V_{\mathrm{S2}}) = X\frac{\Sigma_2}{\Sigma_{\mathrm{S}}}, \quad \text{and} \tag{A10}$$

$$\varepsilon^{\mathrm{N}} = (\Sigma_2/\Sigma_{\mathrm{S}})x_0 - (\Sigma_0/\Sigma_{\mathrm{S}})x_2 + x_1 + x_{\mathrm{N}}. \tag{A11}$$

The SD of ε^{N} is

$$\sigma^{\mathrm{N}} = \sqrt{\frac{\Sigma_0}{\Sigma_{\mathrm{S}}}\Sigma_2 + \Sigma_1 + \Sigma_{\mathrm{N}}}. \tag{17}$$

According to equations (9) and (11), if M learns from S,

$$V_{\mathrm{M}T}^{\mathrm{L}} = X + x_0 + x_1 + x_2 + x_{\mathrm{N}}. \tag{A12}$$

Under learning, the coefficients in the latent OLS model (A1) are

$$b_{\mathrm{OLS}}^{\mathrm{L}} = 1, \tag{A13}$$

$$a_{\mathrm{OLS}}^{\mathrm{L}} = 0, \tag{A14}$$

$$\varepsilon^{\mathrm{L}} = x_1 + x_{\mathrm{N}}. \tag{A15}$$

The SD of ε^{L} is

$$\sigma^{\mathrm{L}} = \sqrt{\Sigma_1 + \Sigma_{\mathrm{N}}}. \tag{16}$$

From here, the results of Proposition 1 can be easily obtained by using Lemma 1. Q.E.D.

Proof: Without control for σ^{N}, b^{N} decreases in the amount of market unique information.

The market's estimate of synergy (SYNERGY) comprises the common signal (x_0) and the market unique signal (x_2). The better the manager understands the market, the less unique information should the market have. In the model, the variance represents the informativeness of a signal. Thus, $-1 \le \mathrm{Corr}(\Sigma_0, \Sigma_2) \le 0$, where Σ_0 and Σ_2 are the variances of x_0 and x_2, respectively. For tractability, the proof assumes either $\mathrm{Corr}(\Sigma_0, \Sigma_2) = -1$ or $\mathrm{Corr}(\Sigma_0, \Sigma_2) = 0$.

If the regression (13) does not include any control variable, substitute equation (17) into equation (15):

$$b^{\mathrm{N}} = \frac{\Sigma_0/\Sigma_{\mathrm{S}}}{\sqrt{(\Sigma_0/\Sigma_{\mathrm{S}})\,\Sigma_2 + \Sigma_1 + \Sigma_{\mathrm{N}}}}. \tag{A16}$$

Case 1. Σ_{S} is fixed: $\mathrm{Corr}(\Sigma_0, \Sigma_2) = -1$.

The information content of SYNERGY is constant cross-sectionally. Reformulate equation (A16) into equation (A17):

$$b^{\mathrm{N}} = \frac{1/\sqrt{\Sigma_{\mathrm{S}}}}{\sqrt{\dfrac{1}{\Sigma_0/\Sigma_{\mathrm{S}}} + \dfrac{(\Sigma_1 + \Sigma_{\mathrm{N}})/\Sigma_{\mathrm{S}}}{(\Sigma_0/\Sigma_{\mathrm{S}})^2} - 1}}. \tag{A17}$$

Thus, b^{N} increases in $\Sigma_0/\Sigma_{\mathrm{S}}$ or, equivalently, decreases in $\Sigma_2/\Sigma_{\mathrm{S}}$. The ratio $\Sigma_2/\Sigma_{\mathrm{S}}$ represents the relative amount of market unique information in SYNERGY.

Case 2. Σ_0 is fixed: $\mathrm{Corr}(\Sigma_0, \Sigma_2) = 0$.

Cross-sectionally, the amount of common information is uncorrelated with the amount of market unique information. Reformulate equation (A16) into equation (A18)

$$b^{\mathrm{N}} = \frac{\Sigma_0}{(\Sigma_0 + \Sigma_2)\sqrt{\dfrac{\Sigma_0}{\Sigma_0/\Sigma_2 + 1} + \Sigma_1 + \Sigma_{\mathrm{N}}}}. \tag{A18}$$

Thus, b^{N} decreases in Σ_2, which is the absolute amount of market unique information. Q.E.D.

Proof of Proposition 2: In this proof, I assume that the new signal x_{N} is not informative:

$$x_{\mathrm{N}} = 0. \tag{A19}$$

Substitute equation (A19) into equation (11):

$$V_{\mathrm{M}T} = V_{\mathrm{M}3}. \tag{A20}$$

According to equations (5), (9), and (A20), if M completes the deal, his utility at $t = T$ is

$$U_{\mathrm{M}T} = z + V_{\mathrm{M}1} + x_2. \tag{A21}$$

If M cancels the deal, his utility at $t = T$ is

$$U_{\mathrm{M}T} = -C. \tag{A22}$$

Under the assumption of $x_{\mathrm{N}} = 0$, M always completes the announced deal if he does not learn. If he learns, according to equations (6), (9), and (A20), he cancels the deal if and only if

$$C + z + V_{\mathrm{M}1} + x_2 < 0. \tag{A23}$$

Thus, according to equations (A21)–(A23), the expected value of learning x_2 is

$$K = \int_{x_2=-\infty}^{x_2=-(C+z+V_{\mathrm{M}1})} [(-C) - (z + V_{\mathrm{M}1} + x_2)]\, f(x_2)\, dx_2, \tag{A24}$$

where $f(\cdot)$ is the *p.d.f.* of x_2. I derive the equations (18)–(20) from (A24) using the two general formulas: $\int_{t=-\infty}^{t=x} t\varphi(t)\,dt = -\varphi(x)$ and $\frac{d}{dx}\varphi(x) = -x\varphi(x)$. Q.E.D.

Equation (A24) has additional important properties. The signal x_2 is always valuable ($K > 0$). When investors do not have unique information, the benefit of learning is 0 ($\sigma_2 \rightarrow 0$, $K \rightarrow 0$). The higher the agency/hubris cost z, the less valuable is x_2 ($\partial K/\partial z = \partial K/\partial W = -\Phi(-W/\partial_2) < 0$). When C or z is very high, the value of x_2 is close to 0 ($C \rightarrow +\infty$ or $z \rightarrow +\infty$, $K \rightarrow 0$).

REFERENCES

Bertrand, M., and A. Schoar, 2002, Managing with style: The effect of managers on firm policies, Working paper, MIT.

Bommel, J., 1997, Messages from market to management: The case of IPO's, Working paper, INSEAD.

Brealy, R., and S. Myers, 2000, *Principles of Corporate Finance*, 6th ed. (Irwin/McGraw-Hill, Boston).

Dodd, P., 1980, Merger proposals, management discretion and shareholder wealth, *Journal of Financial Economics* 8, 105–137.

Dye, R., and S. Sridhar, 2000, Strategy-directing disclosures, Working paper, Northwestern University.

Giammarino, R., R. Heinkel, B. Hollifield, and K. Li, 2004, Corporate decisions, information and prices: Do managers move prices or do prices move managers? *Economic Notes* 33, 83–110.

Herz, J. W., and C. H. Baller, eds., 1971, *Business Acquisitions: Planning and Practice*, vols. I and II (Practising Law Institute, New York City).

Jegadeesh, N., M. Weinstein, and I. Welch, 1993, An empirical investigation of IPO returns and subsequent equity offerings, *Journal of Financial Economics* 34, 153–175.

Jennings, R. H., and M. A. Mazzeo, 1991, Stock price movements around acquisition announcements and management's response, *Journal of Business* 64, 139–163.

Jensen, M. C., and W. H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, 305–360.

Lang, M., and R. Lundholm, 1993, Cross-sectional determinants of analysts ratings of corporate disclosures, *Journal of Accounting Research* 31, 246–271.

Luo, Y., 2001, Learning from the market as a motive for disclosure: The case of takeover announcement timing, Working paper, SSRN.

Luo, Y., 2002, Information from capital markets to firms during corporate takeovers, Ph.D. Dissertation, University of Rochester.

Perry, T., and M. Zenner, 2000, CEO compensation in the 1990s: Shareholder alignment or shareholder expropriation? *Wake Forest Law Review* 35, 123–152.

Rappaport, A., 1987, Stock market signals to managers, *Harvard Business Review* Nov–Dec 1987, 57–62.

Rock, K., 1986, Why new issues are underpriced? *Journal of Financial Economics* 15, 187–212.

Verrecchia, R., 1983, Discretionary disclosure, *Journal of Accounting and Economics* 5, 179–194.

Wagenhofer, A., 1990, Voluntary disclosure with a strategic opponent, *Journal of Accounting and Economics* 12, 341–364.