No News Is News:

Do Markets Underreact to Nothing?[*]

Stefano Giglio

University of Chicago, Booth School of Business and NBER

Kelly Shue

University of Chicago, Booth School of Business and NBER

July 30, 2014

Abstract

As illustrated in the tale of "the dog that did not bark," the absence of news and the passage of time often contain information. We test whether markets fully incorporate this information using the empirical context of mergers. During the year after merger announcement, the passage of time is informative about the probability that the merger will ultimately complete. We show that the variation in hazard rates of completion after announcement strongly predicts returns. This pattern is consistent with a behavioral model of underreaction to the passage of time and cannot be explained by changes in risk or frictions. (*JEL* G02, G14, G34)

The dog did nothing in the night time ... that was the curious incident.

- Sir Arthur Conan Doyle

The absence of news reports and the passage of time often contain important information. For example, a citizen who lives through a sustained period without terrorist attacks should update positively on the effectiveness of the government's antiterrorism programs. A manager who observes that an employee has executed a difficult task without incident should update positively on the employee's quality.

"No news" is also news in many financial contexts. For example, if a firm does not lay off workers or declare bankruptcy after a macroeconomic shock, investors should update positively on the firm's underlying strength. On the other hand, if a firm repeatedly fails to announce new investment projects, investors may be justified in updating negatively on the firm's growth prospects. Finally, investment returns that seldom display newsworthy variation can reveal information about the underlying investment decisions. Overly consistent returns may be suggestive of fraud, as in the case of Bernie Madoff's investment fund.

Rational agents should perform Bayesian updating on the passage of time. In efficient financial markets with rational investors, the passage of time can lead to price movements even in the absence of explicit news. Alternatively, agents may be boundedly rational and imperfectly update on the passage of time. A large body of literature in behavioral economics shows that agents tend to underreact to less vivid and salient sources of information.[1] This suggests that agents may underreact to the information tied to the passage of time, which is likely to be less salient than the events typically covered by explicit news stories.

In this paper, we study the extent to which markets incorporate the information content of "no news." By "no news," we do not literally refer to situations in which zero explicit news stories are

---

[1]See, for example, Agostinelli, Sherman, Fazio, and Hearst (1986), Gabaix, Laibson, Moloche, and Weinberg (2006), Gifford (2005), Healy (1981), Hearst (1991), Kahneman (1973, chapter 2), Maki (1989), Neisser (1963), Pashler (1998), Pedzek, Maki, Valencia-Laver, Whetstone, Stoeckert, and Dougherty (1988), Radner and Rothschild (1975), Sargent (1993), Simon (1955), Treisman and Gormican (1988), and Treisman and Souther (1985) for theories and evidence from psychology, and Merton (1987), Peng (2005), and Peng and Xiong (2006) for applications to financial markets.

released. Rather, we define "no news" as the information content tied to the passage of time, that is, what market participants should know by observing the passage of time even if they are unable to observe explicit news.

We focus on a financial context in which we can easily quantify the information contained in the passage of time: mergers. Mergers offer a convenient empirical setting for several reasons. First, each merger has a clear starting point: the announcement of the intention to merge. Second, the returns of merger investment strategies heavily depend on a well-defined and stochastic ending point: the merger either completes, or the parties withdraw for reasons, such as loss of financing, antitrust rulings, or target shareholder resistance. To best capture this uncertainty, we focus on mergers without known expiration dates, so that both the timing and outcome of merger resolution are stochastic. Between merger announcement and resolution, there exists an interim period that usually lasts several months to a year. We show empirically that the passage of time during this interim period contains information about whether the deal will ultimately complete. We then compute how prices should move during this interim period if markets fully incorporated the information tied to the passage of time.

Using a sample of over 5,000 mergers, we estimate the hazard rate of merger completion, defined as the probability that a merger will complete in event week t conditional on it not completing or withdrawing prior to week t. If the hazard rate of completion is nonconstant over the event life of a merger, then the passage of time contains information about merger completion. We find that hazard rates of completion do indeed vary strongly over event time and are hump shaped. Hazard rates rise from zero in the first weeks after announcement, peak around event week 25, and then decline to zero one year after announcement. In contrast, hazard rates of withdrawal are essentially flat. These patterns hold throughout the calendar time period of our sample, 1970 to 2010. They also hold after accounting for potential heterogeneity, such as the form of merger financing or the size of the target.

Rational markets should incorporate all available information, including predictable variation in hazard rates tied to the passage of time. If risk remains constant, high hazard rates should not predict

4

high returns. For example, if the market believes the merger is likely to complete tomorrow, the price of the target should be high today, so that the mean return between today and tomorrow should be the risk-free rate plus compensation for risk. When we look empirically at returns, we instead find a strong positive correlation between hazard rates and returns in the event year following merger announcement. This relationship is robust and holds even when we estimate hazard rates using an earlier sample and mean returns using a later sample.[2] In other words, returns are predictable and they move with the hump-shaped hazard rates: the average return across all deals is just above 20 bp per week in the first weeks after announcement, and it rises above 40 bp per week between event weeks 20 and 30. Average returns then decline sharply as more time passes after announcement.

What explains the strong predictability of returns by hazard rates? We explore two possible explanations: underreaction to the passage of time (the behavioral explanation) and changes in risk or trading frictions over the event lives of mergers (the rational explanation).

First, we develop a behavioral model in which underreaction to information tied to the passage of time generates the observed returns predictability. The model links movements in the target's price to market beliefs about event-time variation in hazard rates. If agents correctly update using the passage of time and systematic risk does not change over the event lives of mergers, then mean weekly returns should be constant in event time. Returns should not vary systematically with the passage of time, and they should not be predicted by the hazard rate.

However, if agents underreact to the information contained in the passage of time, they will behave as though they believe that the hazard rate of completion does not vary over event time as much as the true hazard rate. This implies that agents will tend to underestimate the hazard rate when hazard rates are high and overestimate it when hazard rates are low.

Underreaction to the passage of time further implies that mean returns should be high when hazard rates are high (because markets underestimate merger completion probabilities and receive positive

---

[2]For cash-financed mergers, the relevant return is the return from holding the target. For equity-financed mergers, the relevant return is that from a strategy in which one takes a long position in the target and a short position in the acquirer.

Electronic copy available at: https://ssrn.com/abstract=2139564

surprises on average) and low when hazard rates are low (because markets overestimate merger completion probabilities and are disappointed on average). In other words, underreaction implies that hazard rates and mean returns should be positively correlated. This matches our empirical findings, both for the full sample of mergers, and in various subsamples that account for various sources of deal heterogeneity.

Importantly, these predictions hold even if investors observe explicit news in the interim period between merger announcement and resolution. For example, investors may be exposed to news reports of target shareholder voting results or insider information leaks about merger completion probability. As mentioned above, we define "no news" as the information that market participants should know by observing the passage of time even if they are unable to observe explicit news. If explicit news is observed by market participants, our methodology remains valid because rational investors should update on both explicit news and the passage of time; the passage of time should still not predict returns. We do not rule out the possibility that markets also underreact to explicit news. However, we show that, at a minimum, markets underreact to the passage of time. In fact, any release of explicit news about merger completion probability should be a bias against our findings that aggregate hazard rates tied to the passage of time predict returns. If agents receive explicit news, they should estimate merger completion probability with less error, and therefore aggregate historical hazard rates should be less predictive of returns.

Using our simple model, we estimate the market's beliefs about completion hazard rates that would generate the observed average returns in each event week. The implied beliefs track the empirically measured hazard rates but display approximately 40% less variation over time. This is consistent with an underreaction hypothesis in which agents only partially incorporate the information content of the passage of time. Using parametric and nonparametric tests, we confirm that the underreaction is strongly statistically significant and holds after accounting for potential deal heterogeneity.

While our results are consistent with the behavioral model of underreaction, the positive relationship between returns and hazard rates could also reflect compensation for risk or frictions (the

6

rational explanation). We begin by noting that the correlation between hazard rates and returns is a phenomenon measured over the event life of the merger, and therefore cannot be explained by changes in risk or risk premia over calendar time.[3] Next, we test whether our results can be explained by event-time variation in three types of risk: (1) systematic risk as captured by the Fama-French factors, (2) downside risk, in which returns covary more with the market during market downturns, and (3) idiosyncratic risk. To measure risk, we examine the returns of trading strategies that modify the common merger arbitrage strategy described by Mitchell and Pulvino (2001): for each calendar month, we invest in all mergers active between certain event windows. We test whether a trading strategy that invests in deals active in event weeks when hazard rates are high (estimated from the aggregate sample of mergers in a preceding period) delivers a higher alpha than does a strategy that invests in deals in event weeks when hazard rates are low.

Our High Hazard strategy delivers a significant monthly alpha (relative to the three Fama-French factors) of 105 bp. This is significantly higher than the 33 bp average of strategies that buy deals in the Low Hazard weeks. This is also significantly higher than the 71 bp of the traditional Buy-and-Hold strategy, which invests in deals for their entire event lives. These alphas represent the economic magnitude of potential mispricing: event-time variation in hazard rates predicts a substantial difference in alpha of approximately 72 bp per month between the High and Low Hazard strategies.

We also find that all risk exposures (Fama-French betas, momentum beta, downside beta, beta with respect to the option factors of Agarwal and Naik [2004], and idiosyncratic risk) do not vary significantly in event time. Therefore, although risk is a potential contributor to the positive returns in the traditional Buy-and-Hold merger arbitrage strategy, it cannot explain why returns covary with hazard rates over the event lives of mergers.

Finally, we consider alternative rational explanations based on event-time-varying frictions and asymmetric information. For example, to lock in capital gains, large institutional investors tend to

---

[3]Because mergers occur in waves, we may be concerned that event time is correlated with calendar time, and therefore changes in risk or risk premia over calendar time may matter. However, in a regression of returns on hazard rates, the results remain similar after we control for calendar time (year x month) fixed effects, which removes calendar year-month variation in risk and risk premia.

sell the target immediately after announcement. If not enough arbitrage capital takes the other side of the deal, the downward price pressure could result in low returns immediately after announcement, followed by rising returns as arbitrage capital enters. It is also possible that the degree of asymmetric information changes in event time, such that the buyer's required compensation for the asymmetric information also changes. These explanations predict that mean returns should be correlated in event time with proxies for market liquidity and asymmetric information. Instead, we show that almost all the event-time variation in these market conditions is concentrated in the first two weeks after announcement, whereas the event-time variation in returns that we document occurs on a different time scale, in the months following announcement. Further, we show that potential "last day effects" associated with delisting returns after the last day of trading before merger completion cannot generate our returns patterns.

We conclude that changes in risk, frictions, and asymmetric information in event time are unlikely to drive the relationship between hazard rates and returns. Rather, the empirical evidence supports the behavioral hypothesis that markets fail to incorporate all information contained in the passage of time while waiting for merger resolution.

Given that sophisticated investors are likely to exist in these markets, we explore why these returns patterns are not arbitraged away. Transaction costs may prevent arbitrageurs who are aware of the phenomenon from trading against it. To test this hypothesis, we study how the High Hazard strategy performs when executed on subsamples of mergers for which arbitrage is likely to be more difficult because of higher transaction costs. We find that the alphas of our High Hazard strategy are significantly larger for smaller deals, for deals with lower volume and turnover, for deals with higher bid-ask spreads, and for mergers that took place in the first part of our sample, when arbitrage activity was less intense. We also simulate realistic trading strategies that limit the total exposure to each deal and account for direct and indirect transaction costs following the procedure developed by Mitchell and Pulvino (2001). Accounting for trading costs and limitations in the size of the positions pushes our alphas toward zero, although we still find significant differences in the alphas between

8

our High and Low Hazard strategies. This is consistent with a limits to arbitrage view in which boundedly rational retail investors generate the mispricing and sophisticated investors are unable to fully arbitrage away mispricing due to transaction costs.

To the best of our knowledge, this is the first paper to empirically investigate underreaction to the passage of time. However, our findings build upon and complement related findings in behavioral finance. For example, Da, Gurun, and Warachka (2012) show that markets underreact to the slow release of news. Corwin and Coughenour (2008) and Barber and Odean (2008) show that investors focus on familiar or attention-grabbing stocks, whereas Hirshleifer and Teoh (2003), Hirshleifer, Hou, Teoh, and Zhang (2004), and DellaVigna and Pollet (2009) study limited attention with respect to firm disclosure. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) find evidence of limited attention with regard to firms' economic linkages. Chan (2003), Gilbert, Kogan, Lochstoer, and Ozyildirim (2012), Hirshleifer, Lim, and Teoh (2009), Huberman and Regev (2001), and Tetlock (2011) study under- and overreaction to explicit news in different financial settings. Overall, the existing literature argues that investors underreact to less salient news when sifting through a set of explicit news stories. This paper shows that investors also underreact to the absence of news, which alone can contain valuable information.

Our research also relates to several recent papers that explore price movements in the absence of trading activity. Marin and Olivier (2008) and Gao and Ma (2012) find that markets partially incorporate the information tied to the absence of insider trading. Bagnoli, Kross, and Watts (2002) show that prices drop following delays in earnings reports because delays convey negative information.[4] Compared with these papers, which highlight why the absence or delay of activity should be informative, we develop a behavioral model and quantify the extent to which markets underreact to the information content of the passage of time.

Understanding how agents process the absence of news is economically important because the

---

[4]A related theory literature includes Campbell and Hentschel (1992), who show that the absence of major price movements predicts low future volatility, Galai, Raviv, and Wiener (2007), who model the relationship between time spent in distress and liquidation, and Jung and Kwon (1988), who model the information content of the absence of disclosures.

9

passage of time often contains valuable information that can reduce asymmetric information problems (for example, between voters and politicians, employers and employees, and investors and insiders). Underreaction to no news can therefore lead to misallocation of resources in a variety of contexts. In addition, the distortionary effects of underreaction to no news can be amplified by the fact that no news tends to be slow moving and persistent.

1   Data

We combine data on merger activity from two sources. The first data source, generously shared by Mark Mitchell and Todd Pulvino (MP), covers merger activity from 1970 to 2005. It is an updated version of the data described by Mitchell and Pulvino (2001). The second data source is Thomson One (TO), formerly known as the SDC, which covers merger activity from 1985 to 2010. Because MP covers a longer time series, whereas TO offers more comprehensive coverage of recent years, we combine the two datasets as follows: we use the MP dataset for years up to and including 1995 and the TO dataset afterward. The exact year of the split is determined by a comparison of the relative coverage of the two datasets in each year. Our results are robust to using only MP or TO data.

We define the takeover premium for cash-financed deals (in which the shareholders of the target obtain a fixed amount of dollars per share upon deal completion) as the ratio of the initial offer price at deal announcement to the price of the target two days before deal announcement. For equity-financed deals, in which the shareholders of the target obtain Δ shares of the acquirer per share of the target owned upon deal completion, the takeover premium is defined as $\Delta * P^A_{t=-2}/P^T_{t=-2}$. We refer to Δ as the exchange ratio, whereas P^A and P^T are the acquirer's and target's share prices, respectively.

We apply the following filters to our initial sample of mergers (a detailed description of the sample size remaining after each filter is reported in the Appendix).

1. The merger is all cash-financed or all equity-financed. We exclude hybrid forms of financing or

deals with contingency terms (e.g., collar agreements) because they are more difficult to price using the available data on equity prices. For equity-financed deals, we require that there exists data on the exchange ratio for the deal.

2. The merger takes the form of a simple one-step merger without a known expiration date for investors to tender shares. We exclude tender offers, which have known expiration dates, because the information content of the passage of time near and beyond the expiration date is likely to be obvious to market participants.[5]

3. For cash-financed mergers, equity price data are available for the target from the Center for Research in Security Prices (CRSP). For equity-financed mergers, equity price data for both the target and acquirer are available from CRSP. We restrict our analysis to U.S. targets in the case of cash mergers and U.S. targets and acquirers in the case of equity mergers.

4. We exclude deals for which the typical hazard rates of completion or withdrawal are less applicable. First, we exclude deals that compete with a previous bid for the same target that was announced within the past three years because competing bids are relatively more likely to withdraw and follow more deal-specific heterogeneity in timing. Second, we exclude deals in which the target price exceeds the offer price immediately after announcement. In these cases, it is likely that the market expects either a competing offer or a favorable revision of deal terms and deal completion is less likely to be the primary form of uncertainty.

Note that these filters only exclude deals from the sample or investment strategy based upon information that was publicly available at the time of the deal. After applying these filters, we are left with 5,377 deals, which are summarized in Table 1. If a deal does not complete, it can either be formally withdrawn on a particular date or remain pending. Seventy-three percent of deals complete

---

[5]In some cases, one-step mergers have projected completion dates or drop-dead dates (as opposed to expiration dates) that are disclosed at merger announcement. Discussions with M&A lawyers suggest that substantial uncertainty remains even with these date projections because of the many uncertain steps involved in the merger process that can greatly affect the timing and probability of completion. Therefore, we keep all one-step mergers in the sample. In the Appendix, we discuss how hazard rates depend on the drop-dead date.

within the first year after announcement, and 21% withdraw within the year. The median time to completion is 88 days.

2 Hazard Rates

In this section, we document how the hazard rate of completion varies over the event lives of mergers. Variation in hazard rates represents one important reason why the passage of time after merger announcement should contain information about whether the deal will ultimately complete. Other reasons why the passage of time may contain information are discussed in Section 4.4.

2.1 Empirical hazard rates

Let t refer to the number of weeks after the merger announcement. Note that t measures event time rather than calendar time. Let $S(t)$ be the probability that the merger survives until time t; that is, it does not complete or withdraw prior to t. Let $h(t)$ be the hazard rate of completion at time t, that is, the probability that the merger completes during period t conditional on surviving up to t. We also estimate a separate hazard rate of withdrawal $w(t)$, although we will show that this hazard rate remains roughly constant over event time.

We use the standard Nelson-Aalen estimator of competing hazard rates. We estimate the hazard rates of completion (withdrawal) as the fraction of deals that complete (withdraw) during each period t among those that have survived until time t, taking into account that once a deal has completed it cannot withdraw, and vice versa. The Nelson-Aalen estimator assumes that all merger completion and withdrawal events are drawn from the same underlying distribution and provides an estimate of this distribution at each point in event time. In reality, it is possible that deal completions and withdrawals follow different hazard processes depending on the observable or unobservable characteristics of each deal. We explicitly account for heterogeneity among deals in the next sections, and for now estimate a single hazard curve for all deals.

Figure 1 shows the estimated hazard rates of completion and withdrawal for the sample of all

mergers. We report estimates using the full sample of mergers (1970–2010), and separately over the early and late parts of the sample (1970–1990 and 1991–2010).[6]

Three main results emerge from these figures. First, the hazard rates of completion are strongly nonconstant. They start at around zero during the first weeks, then rise to about 6% per week around week 20, and gradually decline to zero by the end of the first year after announcement. Second, hazard rates of withdrawal are instead essentially constant. Third, hazard rate patterns estimated using the early and late calendar time samples are similar, suggesting that hazard rate patterns have not changed significantly over the past several decades. These results indicate that the passage of time from announcement contains useful information about the completion probability of a merger: for example, knowing that a year has passed since the announcement reveals that the completion probability is likely to have fallen close to zero.

2.2 Heterogeneity in hazard rates

The previous analysis showed significant variation in hazard rates over event time under the assumption that all mergers have the same hazard rate curve. However, the hazard rate for any specific merger may differ from the hazard rate we estimate using aggregate data because of observed or unobserved heterogeneity.

We tackle the possibility of observed heterogeneity by considering a set of observable merger characteristics and studying whether hazard rates indeed differ across deals based on these characteristics. The twelve dimensions we consider are: type of financing (cash or equity), size of the target, calendar time period of the deal, merger arbitrage spread (the relative difference between the effective offer price and the target price two days after merger announcement), the premium at announcement (the relative difference between the effective offer price and the target price two days before merger announcement), the drop-dead date of the deal, whether the deal occurs during a merger wave, whether

---

[6]In Figure A1 we extend the 50-week event window after deal announcement to 100 weeks. Note that all hazard rates are estimated using the STATA "sts graph" command and estimate a model of competing hazard rates, in which all deals that do not complete or withdraw within the plotted event window are counted as censored as of the end of the event window.

the merger is diversifying for the acquirer, target share price (low-priced stocks are more likely to be small and illiquid), whether the deal (if equity-financed) has a floating or fixed exchange ratio, whether the deal is friendly or hostile, and target termination fee.

In the Appendix, we divide all deals into two approximately equal-sized groups based on the value of each characteristic. We then nonparametrically estimate the hazard rate in each group. The Appendix shows that in all cases, the hazard rate displays a strong hump shape: completion hazard rates are low at the beginning and the end of the year after announcement and rise during the intermediate period. In many cases, accounting for heterogeneity leads only to differences in the estimated scale of the hazard curve, whereas the hump shape and the timing of the hump remain similar. These scaling differences do not affect our results, as our later analysis uses the location of the hump (e.g., early or late in event time), not the level of hazard rate. Only with respect to the type of financing, size, diversification, target share price, and drop-dead dates does the location of the peak of the hump vary across groups. Our main analysis will take into account heterogeneity in these observable characteristics.[7]

Our results are also robust to unobserved heterogeneity for two reasons. First, we will test a behavioral hypothesis that predicts a positive relationship between each individual merger's latent hazard rate and returns. To the extent that our measured hazard rate approximates each merger's individual hazard rate with noise, this is a bias against our empirical findings in support of the behavioral hypothesis. Second, in Proposition 1 in the Appendix, we prove a theoretical result under commonly used proportionality assumptions about the nature of the unobserved heterogeneity: if unobserved heterogeneity in hazard rates is present, the mean of the true hazard rates will display more event-time variation than does our measured hazard rate. In our case, we estimate a hazard rate that is strongly hump-shaped: the true hazard rate will necessarily display even more time variation. This implies that even if unobserved heterogeneity is present, there is information content in the

---

[7]We do not perform the main analysis by drop-dead date even though it generates hazard rate variation because the sample for which we observe drop-dead dates is too small to perform the analysis after subdividing into two groups. However, we present results using all available data on drop-dead dates in Figure A2 and Table A1.

passage of time. Armed with the result that hazard rates of completion vary significantly over the event lives of mergers, we now study the implications for returns.

3   Returns and Hazard Rates

In this section, we document a surprising positive correlation between hazard rates and average weekly returns over event time. Further, this relationship continues to hold when we estimate hazard rates using an aggregate sample in a previous period and returns in a later period. For cash mergers, the relevant return is the weekly return from investing in the target. For equity mergers, the relevant return is the weekly return from going long the target and shorting Δ shares of the acquirer. Each event week's return includes the gains from any delisting, that is, the upside from attaining the acquirer's offer price if the merger completes in that week. Note that we use actual returns for each event week, and do not scale any daily return to a weekly horizon for deals that complete in the middle of a week.[8]

We start by plotting average returns across deals in event time. Because very few deals survive until one full year after announcement and returns are noisy, in all subsequent analysis, we focus on event weeks 1 through 45. [9]

Figure 2 plots completion and withdrawal hazard rates in the top panel and mean weekly returns in the bottom panel. Because of noise in the returns data, we plot returns over event time by fitting a smoothed local mean[10] to the panel series of returns for each deal in each event week. The figure shows smoothed returns using the optimal bandwidth. In unreported results, we also plot the curve using 0.5 and 1.5 times the optimal bandwidth, as well as fitting a local linear regression, and find qualitatively similar results. The figure shows that the hazard rate of completion and mean weekly

---

[8]We test our underreaction hypothesis using stock prices as opposed to option prices. In theory, option prices may offer insights into the market's perceptions of completion probabilities, as well as beliefs about downside and jump risk. Unfortunately, less than 10% of our sample consists of deals in which options for the target are traded. In addition, the options subsample tends to exclude small stocks, for which we find the greatest degree of underreaction.

[9]Only 6.2% of deals survive beyond 45 weeks after announcement. All results in this paper are substantively unchanged if we include returns after week 45, although the confidence intervals for average returns (as plotted in Figure 2) are very wide for all weeks after week 45.

[10]In particular, we fit a local polynomial of degree 0 using the Epanechnikov kernel.

returns tend to move together. In the first weeks after announcement and towards the end of the first year after announcement, completion hazard rates are below the average and returns are below the average as well. In the intermediate weeks, hazard rates are high and returns are high as well.

In Figure 2, we also plot 90% pointwise confidence bands for each point in the returns curve. These confidence bands grow wider as we approach one year after merger announcement because fewer deals survive as time passes after announcement. We also conduct a more formal test of whether returns are constant over event time. We estimate a regression of returns on indicators for each event week following deal announcement, with controls for calendar year-month fixed effects and with standard errors double-clustered by merger and calendar year-month. The results are reported in Table 2 (first column, first row), and show that we can statistically reject that returns are constant across event weeks, when pooling together all deals.

Next, we test the strength of the relationship between returns and completion hazard rates. The second column of the table reports the results of a regression of weekly returns on hazard rates, with controls for calendar year-month fixed effects. Observations are at the merger by event week level. The fixed effects control for possible calendar time variation in unobservables that might affect returns, for example, calendar-time variation in risk or risk premia. We allow standard errors to be double clustered at both the the calendar year-month level and at the merger level. We find that hazard rates (estimated from the aggregate sample) significantly predict returns over event time. The relationship continues to hold when we adopt a split-sample approach that is free of "lookahead" bias. In row (4) of the table, we show that hazard rates estimated using the first half of our sample (pre-1991) predict returns in the second half of the sample.[11] The other rows of the table report regressions of returns on hazards after conditioning on observable sources of heterogeneity; we postpone a discussion of these results for Section 4, after presenting a model of underreaction, because the model will be useful in

---

[11]We also estimate the regression computing returns using only the long side of equity deals (i.e., just buying the target without shorting the acquirer). This means that, although the arbitrage hedge is incomplete, we are not affected by potential deal revisions that may change the exchange ratio Δ. In addition, we can include many more (especially small) deals for which we do not have information about Δ in our data. The coefficient of the regression remains significant and increases in magnitude: 0.069 (the standard error is 0.013).

interpreting the results.

Overall, we find that returns following merger announcement are nonconstant and strongly predictable using aggregate hazard rates. The strong predictability is surprising because we expect rational markets to incorporate all available information, including predictable variation in hazard rates. For example, if the market understands that the merger is likely to complete tomorrow, the price of the target should be high today, such that the mean return only reflects compensation for risk.

What explains this return predictability? In the remainder of this paper, we explore two possible explanations: underreaction to the passage of time (the behavioral explanation) and changes in risk or frictions over the event lives of mergers (the rational explanation).

4   A Simple Behavioral Model of Underreaction

To understand what time variation in hazard rates implies for returns when markets imperfectly update on the passage of time, consider the following parsimonious pricing model for the returns of the target of a cash merger after the announcement of the intention to merge.

4.1   The model

Let t represent event time after merger announcement. Let r be the constant risk-free rate. Let $\hat{P}(t)$ be the price of the target's shares after merger announcement, but before the deal has completed or withdrawn. If at any point the deal completes, the value of the target jumps to P_C, the amount of cash per share promised to the target's equity holders. If at any point the deal is withdrawn, the price jumps to $P_0(t)$, where $P_0(t)$ is some latent process.[12] We model $P_0(t)$ as follows:

$$dP_0(t) = \mu P_0(t)dt + \sigma P_0(t)dZ(t), \tag{1}$$

---

[12]Using insights from Malmendier, Opp, and Saidi (2011), who show that merger announcement can change the underlying value of the target even if the merger never completes, we do not constrain $P_0(t)$ to represent the value of the target if the merger had never been announced. Rather, $P_0(t)$ represents the value that the target share price would revert to if the acquirer were to withdraw at time t.

where $Z(t)$ is a standard Brownian motion. We assume that there is an end time, T, such that any deal that does not complete by time T is assumed to never complete (in accordance with the empirical evidence that shows that hazard rates of completion fall to zero approximately one year after merger announcement).

If the merger has not completed or withdrawn prior to time t, the price of the target is determined as follows:

$$
\begin{aligned}
\hat{P}(t) = \ & E_t\{\int_t^T e^{-r(z-t)}e^{-\int_t^z[\hat{h}(k)+\hat{w}(k)]dk}\hat{h}(z)P_C dz \\
& + \ \int_t^T e^{-r(z-t)}e^{-\int_t^z[\hat{h}(k)+\hat{w}(k)]dk}\hat{w}(z)P_0(z)dz \\
& + \ e^{-r(T-t)}e^{-\int_t^T[\hat{h}(k)+\hat{w}(k)]dk}P_0(T)\},
\end{aligned}
\tag{2}
$$

where $\hat{h}(t)$ and $\hat{w}(t)$ are risk-neutral hazard rates.

Because we wish to focus on a possible behavioral explanation, assume for now that all risk is idiosyncratic and the market believes that all risk is idiosyncratic. This means that we can interpret $\hat{h}(t)$ and $\hat{w}(t)$ as market beliefs about the true hazard rates, as opposed to the risk-neutral hazard rates that also reflect the risk attitude of the market (we postpone a thorough discussion of risk to a later section).

Let $h(t)$ and $w(t)$ be the true hazard rates, as opposed to the market beliefs represented by $\hat{h}(t)$ and $\hat{w}(t)$. $\hat{P}(t)$ in Equation (2) is a function of market beliefs about hazard rates, whereas the average realized one period return from holding the target is a function of both $\hat{P}(t)$ and the true hazard rates:

$$
E\left[ret_t\right] = \left(\frac{P_C}{\hat{P}(t)} - 1\right)h(t)dt + \left(\frac{P_0(t)}{\hat{P}(t)} - 1\right)w(t)dt + \left(\frac{d\hat{P}(t)}{\hat{P}(t)}\right)\left[1 - h(t) - w(t)\right]dt.
\tag{3}
$$

Combining Equations (2) and (3), we can decompose the expected one-period return as follows:

$$
\begin{aligned}
E\left[ret_t\right] &= rdt \\
&+ \left(\frac{P_C}{\hat{P}(t)} - 1\right)\left[h(t) - \hat{h}(t)\right]dt \\
&+ \left(\frac{P_0(t)}{\hat{P}(t)} - 1\right)\left[w(t) - \hat{w}(t)\right]dt.
\end{aligned}
\tag{4}
$$

Note that

$$
\left(\frac{P_C}{\hat{P}(t)} - 1\right) \geq 0 \text{ and } \left(\frac{P_0(t)}{\hat{P}(t)} - 1\right) \leq 0.
$$

The model generates simple testable predictions concerning the relationship between hazard rates and mean returns at each event time t. First, if markets have correct beliefs about hazard rates $(h(t) = \hat{h}(t), w(t) = \hat{w}(t))$, the mean target return will always equal the risk-free rate r (because all risk is assumed to be idiosyncratic). Second, if the market underestimates completion hazard rates $(\hat{h}(t) < h(t))$, mean returns will be higher than the risk-free rate r. This occurs because the market, underestimating the probability of completion, will receive positive surprises on average, generating abnormally high returns. Finally, if the market overestimates the completion probability $(\hat{h}(t) > h(t))$, the target's stock will be overvalued at time t and experience a return that is lower than the risk-free rate. Incorrect beliefs about the withdrawal hazard rate would similarly lead to deviations from the risk-free rate, following Equation (4). Note that returns in each period will deviate from the risk-free rate only if beliefs differ from the true hazard rate in the current period (future differences between beliefs and true hazard rates do not matter, except in affecting the second and third terms through $\hat{P}(t)$).

These predictions directly map to the behavioral hypothesis of market underreaction to no news. Suppose that markets fail to use the passage of time to update on changes to the hazard rate, but have correct beliefs on average over the event life of a merger. In other words, the market believes that $\hat{h}(t) = \overline{h}$ and $\hat{w}(t) = \overline{w}$, where $\overline{h}$ and $\overline{w}$ represent the average of the true hazard rates. This implies

that the market will have approximately correct beliefs about the hazard rate of withdrawal because $w(t)$ is approximately constant over time. However, the market will underestimate the completion hazard rate during event weeks in which the true hazard rate is high. During these times, the model predicts that we should observe particularly high returns for the target's stock. In contrast, in event periods in which the true completion hazard rate is low, markets will overestimate the hazard rate, and the model tells us that we should expect to see particularly low average returns for the target. In other words, underreaction to the passage of time implies that mean returns should be positively correlated with true hazard rates, exactly as we observe in the data.

Figure 3 shows an example of how the relationship between hazard rates and returns varies depending on whether beliefs are correct. The top panel shows an example of true completion and withdrawals hazard rates. It also plots a sample set of beliefs, for illustrative purposes only, in which the market holds correct beliefs about hazard rates for the first several weeks after deal announcement (the dotted line and the solid lines coincide).[13] After a certain number of weeks, and up to a year after announcement, agents fail to use the passage of time to update on changes in the hazard rate. The beliefs about the completion hazard rate are constant but correct on average. As a consequence, in this example, markets underestimate the true completion hazard rate between weeks 10 and 37 and overestimate the hazard rate from week 37 onward.

The lower panel of Figure 3 shows the model predictions for average excess returns over the risk-free rate in each event week among the set of of deals that have not yet completed or withdrawn. During event periods in which beliefs are correct, mean excess returns are zero (the return is equal to the risk-free rate). When markets underreact to no news but have correct beliefs on average about hazard rates, the returns curve follows the shape of the hazard rate of completion: returns are positively correlated with hazard rates.

These predictions extend to a model in which merger returns contain risk that is systematic and in which risk and risk premia are allowed to be nonconstant in calendar time. As long as risk and

---

[13]Figure 3 does not correspond to the actual beliefs we estimate; it is an example of model-implied returns given a certain set of beliefs.

risk premia do not vary systematically over event time, rational updating on the passage of time implies that merger returns should be constant over the event life of the merger (although mean returns may exceed the risk-free rate). Underreaction to no news still implies a positive relationship between hazard rates and returns. These predictions also extend to a model of equity mergers; returns for these deals are those from a portfolio in which investors long the target and short the acquirer. Finally, these predictions hold even if agents also have incorrect beliefs about the average completion rate over the merger's event life. As long as hazard rate beliefs exhibit flatter event-time variation than true hazard rates, the model predicts a positive relationship between hazard rates and mean returns.

Figure 2 and Table 2 document that the main prediction of the model is borne out in the data: we find a significant positive correlation between hazards and returns over event time. After a brief discussion of underreaction to explicit news versus underreaction to the passage of time, we will present additional direct tests of the model of underreaction, as well as tests that take into account several dimensions of deal heterogeneity.

4.2 Underreaction to no news versus underreaction to explicit news

During the interim period between merger announcement and resolution, investors may observe explicit news. For example, investors may see news coverage of shareholder voting results or insider information leaks about merger completion probability.[14] By "no news," we do not refer to the situation in which no explicit news is released. Rather, we define "no news" as the information content tied to the passage of time, that is, what the market should know by observing the passage of time even if agents are unable to observe explicit news.

In this section we show that the release of explicit news is not a problem for our methodology; rather, it is a bias against our findings. While we cannot rule out the possibility that markets

---

[14]For example, Ahern and Sosyura (2014) study explicit news released during merger negotiations. Although they focus on news released prior to merger announcement, their analysis suggests that news released after announcement may also strongly affect market expectations.

underreact to explicit news, we can show that, at a minimum, markets underreact to the passage of time.

The key insight is that it is possible to underreact to multiple sources of information at the same time. If multiple signals observed by the market convey the same information and this information is not reflected in prices, then the market must be underreacting to all signals. "No news," defined as the passage of time, is one public signal that the market should observe. If the passage of time predicts returns, then the market must be underreacting to that signal and possibly to other signals that contain the same information.

For example, suppose that, historically, the probability of merger completion drops drastically just after event month six. The market observes a new merger approaching month six. The information content of the passage of time tells us that this merger is now unlikely to complete. Consider first the case in which there is no explicit news released after merger announcement. If the price does not drop as the merger approaches month six, then we know that markets underreacted to the information content of the passage of time.[15]

Now suppose that, right before month six, the market also observes a news report warning that this deal is unlikely to complete. At this point, the market observes two signals conveying the same information: the passage of time beyond month six and the negative news report. If we again find that the price does not drop, then we know that markets have underreacted to both the information content of the news report and the passage of time.

Given that prices do not drop as the merger approaches month six, could it be that agents underreacted to explicit news but correctly incorporated the information content of the passage of time? No: if markets had incorporated the information content of the passage of time, they would have realized that the merger was unlikely to complete because it was past month six, and the price would have dropped.

---

[15]To see why the price should decline as we reach month six, recall that, up to month six, the probability of completion is positive, so the value of the target incorporates the probability that the deal will complete and the target shareholders will gain from the completion. As the completion probability drops, the value of the target drops because now the merger is unlikely to complete.

In our paper, we show that prices do not incorporate information that is available by looking at the passage of time. Therefore, markets (at a minimum) underreact to the passage of time. We cannot exclude that prices also fail to incorporate information contained in explicit news. In fact, the release of explicit news should be a bias against our measures of underreaction. The motivation is as follows: if a deal is heavily covered by explicit news, it is easier for agents to understand deal completion probabilities as some of the information content of the passage of time also may be covered by the explicit news. This should lead to reduced correlation between hazard rates and returns.

To explore the effect of explicit news on measured underreaction, we merge our main dataset with data on the number of times per day the firm is mentioned in major news sources (e.g., DJ Newswire and Reuters) scraped from Factiva. We are able to match 53% of the firms in our dataset. We find that news arrival is concentrated in the first few weeks after deal announcement and drops to a steady level afterward. In particular, it does not display any hump shape (see Figure A4). We then classify each deal as "high news coverage" or "low news coverage" based on the average number of newspaper citations per day in the period between the day after announcement and on deal completion, deal withdrawal, or one year after announcement (whichever happens first). We consider the availability of news as one additional source of heterogeneity in our tests of the model, to which we turn next.

4.3 Testing the model

We present three tests of the model using data on hazard rates and returns. We perform our tests for the aggregate sample that includes all deals, as well as within subsamples obtained by dividing deals along dimensions of heterogeneity. We choose the dimensions of heterogeneity for which we observe significant hazard rate variation across types: cash versus equity financing, small versus large target size, low versus high target stock price, and diversifying versus nondiversifying mergers.[16] In addition to these characteristics, we also report results grouped by time period (early sample vs. late sample),

---

[16]Even though it generates hazard rate variation, we do not perform the analysis by drop-dead date: the sample for which we observe drop-dead dates is too small to perform the analysis after subdividing into two groups. However, in Figure A2 and Table A1, we present results using all available data on drop-dead dates.

an analysis that only uses returns from the later sample and hazards computed in the earlier sample (which is robust to lookahead bias), and a classification of deals by high or low news coverage.

4.3.1 Model test: Regression of returns on hazard rates. The first, least restricted, test of the model is the regression of returns on completion hazard rates. The model predicts that, as long as systematic risk does not change over event time, hazard rates should be positively correlated with returns if agents underreact to the passage of time. We showed above that this positive correlation holds for the sample that pools all deals (Table 2, row 1). The remaining rows of Table 2, Columns 1 and 2, repeat the test of constant returns in event time and the estimation of the relationship between hazards and returns for groups of deals sorted by characteristics.

The first column shows that we can reject the null hypothesis of constant weekly returns over event time for all but the subsample of large targets and targets with high share price (which is correlated with being a large target in terms of market capitalization). This result makes sense because it is easier for arbitrageurs to correct mispricing among large stocks. The lack of returns variation over event time for these large stocks is consistent with our results in Section 7, in which we discuss limits to arbitrage and show that mispricing is concentrated in deals with higher transaction costs.

Column 2 tests whether returns are correlated with hazard rates. We find a strong positive relation for all samples, except for large targets, targets with high share price, and deals with high news coverage. The lack of significance in the large targets and high share price samples is again consistent with our findings in Section 7 that mispricing is concentrated in deals with higher transaction costs. The lack of significance in the sample with high explicit news coverage is also not surprising because, as we discussed in Section 4.2, the presence of explicit news makes it more likely that the market will learn some of the information contained in the passage of time from the explicit news sources. More information results in less underreaction and therefore reduced correlation between returns and hazard rates.

Overall, this first test supports the behavioral model's predictions that agents underreact to the information contained in the passage of time, except in subsamples in which arbitrageurs are likely to

be very active or in which the information content of the passage of time is also likely to be available from explicit news sources.

4.3.2 Parametric tests of the model. As a second, more direct, test of the model, we perform a parametric test. Suppose beliefs regarding the hazard rate take the form:

$$\hat{h}(t) = (1 - \theta)\overline{h} + \theta h(t).$$

In this case, an agent with $\theta = 1$ has completely correct beliefs, whereas an agent with $\theta = 0$ has beliefs that are flat and equal to the average hazard rate $\overline{h}$. θ captures the degree of underreaction, with lower values corresponding to more underreaction. Using data on hazards and returns, we can estimate θ and test whether it is significantly different from one. We report the details of the estimation in the Appendix. Column 3 of Table 2 reports θ for all deals and separately for each group of deals sorted by their characteristics. The table shows that the parameter θ is estimated to be as low as 0.31 for low-priced target stocks, and as high as 0.92 for high-priced target stocks. We can reject the full rationality hypothesis ($\theta = 1$) in all cases, except for the case of large targets, high-price targets, and deals with high news coverage. The inability to reject the hypothesis that $\theta = 1$ in the large targets and high share price samples is again consistent with our findings in Section 6 that the degree of mispricing is low among deals that are easier to arbitrage (e.g., larger deals with greater liquidity and lower transaction costs). The inability to reject the hypothesis that $\theta = 1$ among the high news coverage sample is again consistent with the idea the explicit news makes it more likely that the market will learn some of the information contained in the passage of time from the explicit news sources, leading to reduced mispricing. The overall value of θ estimated using all mergers is 0.6, which is significantly different from one and suggests a noticeable amount of underreaction.

4.3.3 Nonparametric test of the model. We can also use the model and observed returns to flexibly estimate the market's beliefs with regard to completion hazard rates. This test is nonparametric in the sense that we relax the assumption that $\hat{h}(t) = (1 - \theta)\overline{h} + \theta h(t)$ and allow for beliefs

to take on a more flexible form. In particular, we parameterize the model using the main sample moments of the data: $P_c = 1.3P_0(0)$, corresponding to an approximate 30% takeover premium as shown in Table 1, and $r = 2\%$ per year.[17] Using Equation (4), we estimate the values for beliefs $\hat{h}(t)$ such that the model-implied returns match the observed average return in each event week. To focus on implied beliefs concerning the completion hazard rate, we also impose that beliefs about the withdrawal hazard rate are correct, $\hat{w}(t) = w(t)$. Given that $w(t)$ is approximately constant, the results are robust to assuming that beliefs about withdrawal hazard rates are flat and equal to the mean of $w(t)$.[18] Because the model assumes that all risk is idiosyncratic, we also adjust the average return across all event weeks to be equal to the risk-free rate. In practice, as shown by Mitchell and Pulvino (2001), the average return across all event weeks exceeds the risk-free rate mainly because of transaction costs in operating the arbitrage strategy. We show later that these transaction costs are approximately constant in the event windows that are relevant for our findings.

Figure 4 compares the estimates of true hazard rates with the beliefs implied by fitting the model to the observed returns for the full sample of deals. Consistent with an underreaction hypothesis, we find that implied beliefs of completion hazard rates are flatter than the estimates of true hazard rates. Hazard rates are overestimated at the beginning and the end of the event period, and underestimated in the intermediate period.[19]

---

[17]Of course, $P_0(0)$ need not equal the price of the target prior to merger announcement, as noted by Malmendier, Opp, and Saidi (2011). Our model calibration yields similar results if we instead assume that $P_c = 1.2P_0(0)$ or that $P_c = 1.4P_0(0)$.

[18]In theory, the positive correlation between hazard rates of completion and returns also may be driven by incorrect beliefs about the hazard rate of withdrawal. However, the true hazard rate of withdrawal is flat over event time, so agents will have approximately correct beliefs about the hazard rate of withdrawal even if they ignore the passage of time. To generate the observed returns pattern only through incorrect beliefs about withdrawal, the market must believe that withdrawal rates are hump shaped even though they are flat over event time.

[19]Figure 4 implies that agents overestimate the hazard rate of completion during the first two months following merger announcement. This may seem surprising if we consider that most deals cannot legally complete so soon after announcement due to regulatory barriers, a fact that should be obvious to many market participants. However, the market may anticipate that positive explicit news about merger completion probability will be released in the period immediately following the announcement, which would lead the target's price to converge upward toward the offer price. The release of definitive news about future completion will have the same effect on target prices as if the actual completion event occurs. Therefore, the overestimation of hazard rates observed in the first two months could be explained by the agents being overly optimistic about the probability of obtaining good explicit news about future completion. If agents overestimate the probability of receiving good news in the next period, they will set prices too high in the current period and receive negative surprises (lack of good news) in the next period, leading to low returns in the first weeks after announcement.

Next, we perform a nonparametric test of the difference between the implied and estimated hazard rate curves. We do this for the full sample of mergers, as well as within subsamples that condition on heterogeneity for each deal group. We use the Cramer-Von Mises test, which is common in the medical literature, to compare the two hazard rates, using the critical values for the test as reported byKlein and Moeschberger (2003). Column 4 of Table 2 shows the p-value of the test. For all groups, except for large targets, we can reject the null hypothesis that the two hazard rates are equal at the 10% level.

Taken together, all three tests show strong support for the behavioral model, except for those cases in which we expect the behavioral bias to be weakest: deals for which arbitrage is easier (large deals and deals with high target price, which are strongly correlated) or for which the presence of explicit news makes it more likely that the market will learn some of the information contained in the passage of time from the explicit news sources, leading to reduced mispricing.

4.4 Other information content in the passage of time

In this paper, we focus on the hazard rates of merger completion because they are easily measured and clearly nonconstant over event time. However, variation in hazard rates need not be the only reason why the passage of time after merger announcement contains information. The value of the target, acquirer, or combined entity may change systematically with the passage of time for other reasons. For example, the arrival rate of receiving competing bids from other potential acquirers may be nonconstant over event time. In Figure A5, we show that the hazard rate of receiving competing bids is slightly higher in the weeks immediately following merger announcement than in later weeks (although it is relatively flat compared with hazard rates of completion). In addition, the expected value of the target if the deal does not complete may vary over the event lives of mergers.

It is possible that these other real changes to merger value tend to vary systematically with hazard rates. Therefore, we cannot distinguish between the following:

1. markets underreact to the event-time variation in hazard rates of completion, and

27

2. markets have correct beliefs about the event-time variation in the hazard rates of completion but underreact to other changes to merger or target value that move in event time with hazard rates.

Importantly, both interpretations are consistent with the behavioral hypothesis in implying that hazard rates (and the real events correlated with hazard rates) predict returns because markets underreact to the information content of the passage of time.

5 Risk and other rational explanations

In this section we study the possibility that risk varies in event time with hazard rates. If so, the pattern in returns documented in Section 3 could reflect compensation for risk within a rational framework. We focus on event-time variation in risk because the positive correlation between hazard rates and returns is a phenomenon measured in event time rather than in calendar time. Moreover, we observe over 5,000 mergers staggered across calendar time and control for all calendar-time variation in risk or in risk premia through the use of calendar year-month fixed effects in all regressions. Because we do not observe event-time variation in risk premia, we will focus on event-time variation in risk.[20]

We explore three types of risk that may vary in event time. First, we study systematic risk as captured by the Fama-French factors. Second, we consider downside risk, that is, the possibility of severely negative returns concentrated in bad times. Third, we investigate idiosyncratic risk, which may be important for arbitrageurs because of underdiversification or holding costs.

We measure risk exposures by constructing trading strategies that invest in deals only in specific event-time windows. This allows us to capture potential event-time variation in risk and to estimate the economic magnitude of the variation in returns not explained by risk. Overall, we find that the event-time variation in risk cannot explain the strong correlation between hazard rates and returns.

Finally, we explore whether event-time variation in market frictions or asymmetric information

---

[20]Although we can never observe changes in risk premia over event time, it is not easy to justify why risk premia should change over event time if risk does not also change.

can produce the observed returns pattern. We show that although there is time variation in the volume, turnover, and bid-ask spread of the stock of the target, this variation is concentrated in the short period immediately following merger announcement, and it cannot explain the year-long event-time variation in mean returns. Further, we show that any potential frictions associated with the delisting return and the last day of trading before merger completion cannot generate our returns patterns.

5.1 Constructing portfolio strategies

To understand how the risk exposures of deals vary over event time, we construct calendar-time returns for a set of portfolio strategies, each of which is exposed only to deals active during specific event windows. Our strategies modify the traditional Buy-and-Hold merger arbitrage strategy, described by Mitchell and Pulvino (2001), which buys deals after announcement and holds them until either completion or withdrawal.

The first step in the construction of these portfolio strategies is to identify three event windows based only on the behavior of the completion hazard rate as estimated from an aggregate sample: a first period in which the hazard rate is below its mean (Low Hazard 1 period), a second period in which the hazard rate is above its mean (High Hazard period), and a third period later in a merger's event life when the hazard rate is again below its mean (Low Hazard 2 period). The cutoff points are event weeks 11 and 39. We also adopt a split-sample approach in which we choose event windows using hazard rates from the first half of our sample (pre-1991) and execute the trading strategy in the second half of our sample. Because the shape of the hazard curves remains stable over time, this split-sample method yields very similar cutoff weeks.

Given these cutoff weeks, we construct a series of monthly returns for each of the three strategies. In each calendar month, we invest in all deals that, at the beginning of the month, are active in the relevant event windows for each of the three strategies. To distinguish the three strategies more sharply, we leave 2 weeks around each cutoff point and do not invest in deals that are active in those

29

event weeks. All the results that follow are very robust to the exact choice of the cutoffs, as discussed in the Appendix.

The Low Hazard 1 strategy only invests in deals that, at the beginning of each calendar month, are active in event weeks 1 through 10. The High Hazard strategy only invests in deals that are active in event weeks 12 to 38. Finally, deals active in event weeks 40 to 45 are selected by our Low Hazard 2 strategy. In the middle of a calendar month, if a deal falls out of the relevant event window (e.g., the deal approaches event week 11 and the relevant event window is weeks 1 through 10), we exit out of the deal and invest the proceeds in the risk-free rate for the rest of the month. Similarly, if the deal completes in the middle of a calendar month, we capture the gains from completion and invest the proceeds in the risk-free rate.

For each calendar month, we construct an equal-weighted return using all selected deals. If no deals are active in the relevant event window in a given calendar month, the strategy invests in the risk-free rate for that month. Following standard merger arbitrage strategy, we go long the target for cash deals. For equity deals, we buy the target and short Δ shares of the acquirer for each share of the target bought. This ensures that the return following deal completion does not depend on the price of the acquirer at the time of completion. Note that equity deals involve a short position that exposes the trade to potentially large losses. Therefore, when we construct our portfolios at the beginning of each calendar month, we exclude deals that involve extreme unbalanced positions for the long and short sides relative to the position implied by the initial terms of the deal. In particular, we exit from deals if the premium falls below one (i.e., deals in which the arbitrageur loses money for sure if the deals complete). We also exit from deals if the premium moves above 200% of the initial premium. In these cases the market expects either a competing offer or a major revision of deal terms, and deal completion is less likely to be the primary form of uncertainty. Importantly, we filter deals using information available prior to the trade, and we always take into account the gains and losses of exiting a deal.[21]

---

[21] All results are robust to applying these filters to cash deals as well.

To ensure that our strategy returns do not mistakenly capture price movements due to the initial announcement of the intention to merge, all our strategies start investing on the second trading day after announcement or later depending on the event windows. Finally, although all the deals are equally weighted, we will separately explore the returns of strategies that only invest in large or small deals, as measured by the market capitalization of the target.

5.2 Event-time variation in systematic risk

We begin by testing whether our High Hazard strategy still experiences higher returns than the Low Hazard 1 and 2 strategies, after controlling for systematic risk. Table 3, panel A, shows the Fama-French alphas of the three strategies, using the full sample. The alphas of the two Low Hazard strategies are 27 and 39 bp per month, whereas that of the High Hazard strategy is 105 bp per month (13% per year). On average, the difference in alphas between the High and Low Hazard strategies is 72 bp per month.

Using these estimates, we test the underreaction hypothesis: after controlling for risk, returns and hazard rates are correlated in event time. The corresponding test in terms of the three strategies' alphas can be expressed as

$$(\alpha_{high} - \alpha_{low_1}) + (\alpha_{high} - \alpha_{low_2}) \leq 0$$

relative to the alternative $(\alpha_{high} - \alpha_{low_1}) + (\alpha_{high} - \alpha_{low_2}) > 0$.

As shown in the third column of Table 3, panel A, the alpha of the High Hazard strategy is significantly higher than the alphas of the Low Hazard strategies, and the p-value of the corresponding test is less than 0.01.

The first column of Table 3 also reports the alpha of the traditional Buy-and-Hold strategy, which invests in deals from announcement until completion or withdrawal. Because the Buy-and-Hold strategy invests in deals in both the High and Low Hazard event windows, it is unsurprising that the Buy-and-Hold alpha exceeds the alphas of our Low Hazard strategies but falls short of the alphas of our High Hazard strategies. Because the aim of this paper is to test for behavioral underreaction

31

rather than to maximize portfolio returns, we focus on comparing the High and Low Hazard strategy alphas. Nevertheless, we can also test whether our High Hazard strategy significantly outperforms the traditional Buy-and-Hold strategy. We find that the High Hazard alpha exceeds the Buy-and-Hold alpha by 34 bp per month, with a p-value of 0.07.

Note that our underreaction hypothesis centers on event-time variation in average returns, but does not have direct predictions regarding the average level of returns over the life of the merger. Consistent with the merger arbitrage literature, we find that the average return of a Buy-and-Hold strategy exceeds the risk-free rate even after controlling for standard risk factors. Within our underreaction model, this is equivalent to saying that the average return around which we expect event-time variation is not the risk-free rate r but is some higher value $\mu > r$. In this paper, we do not take a strong stand on why the Buy-and-Hold strategy yields a positive alpha. The literature studying the Buy-and-Hold return argues that the positive alpha reflects compensation for transaction costs and, to a lesser extent, downside risk (see Mitchell and Pulvino, 2001; Baker and Savasoglu, 2002). We will show in later sections that although downside risk and transaction costs may contribute to the average level of returns, they cannot explain the event-time variation in returns.

In panel A of Table 3, we use the same sample to estimate the hazard rates (which generate the event window cutoffs for the trading strategies) and to simulate the trading strategies. In panel B of the table, we report the alphas of portfolio strategies that only use information about hazard rates already available at the time of the investment. We choose the event window cutoffs for the High and Low Hazard strategies based upon hazard rates estimated using pre-1991 data, but we only invest in deals active during the later 1991-2010 sample. Given that Figure 1 shows that the shape of the hazard rate curves remained stable for the past four decades, it is not surprising that all results remain similar when performing the test using the early hazard rates and later returns.

Finally, note that we adopted a conservative approach to compute the alphas presented in panel A. Because fewer deals survive into the event window covered by the Low Hazard 2 strategy (many deals have withdrawn or completed before then), it is more likely to find months with no active deals

32

for the Low Hazard 2 strategy than it is for the High Hazard strategy. Because the return of a month with no active deals is set equal to the risk-free rate, this may artificially bias the alpha and betas of the Low Hazard 2 strategy towards zero, thus hiding the true risk and return properties of deals during the last event weeks. To avoid this problem, in panel A we only include returns from calendar months in which active deals were available for investment. This, while more conservative for tests of event-time variation in strategy alphas, does not represent the returns of a more realistic trading strategy that must invest in the risk-free rate during calendar months in which no deals are active. For completeness, in panel C we repeat the exercise including all months for all strategies, thus forming a tradable portfolio strategy. All the previous results hold.

5.2.1 Betas over event time. We now directly explore event-time variation in risk by looking at the betas of the various strategies. Table 4 reports the strategy betas with respect to the three Fama-French factors. Panel A shows that the betas are all quite small, between 0.04 and 0.28. Panel B shows that the betas for the High Hazard strategy are not significantly different from the betas for the two Low Hazard strategies.

To capture event-time variation in risk exposures in greater detail, we also look at the variation in the betas across each event week (as opposed to dividing the one-year event window into three regions). For each of the 45 event weeks following merger announcement, we construct a calendar time series of returns of a portfolio that only invests in deals that are active in that event week. We then construct a panel of calendar time returns for each of the 45 event-week-specific portfolios.

We plot the estimates of the betas in Figure 5. The figure points to two important features. First, the betas with respect to all the Fama-French factors are again generally small (for example, the market beta is always less than 0.25). The relative magnitudes of the betas cannot account for the difference in the returns that we observe in our High Hazard period (between the two vertical bars) and the Low Hazard periods (the far left and right regions). Second, there does not seem to be significant time variation in any of the betas over event time, although estimation of the betas becomes more noisy as we move toward the right, because fewer deals survive as event time passes.

In panel B of Table 4 we formally test whether betas vary positively with hazard rates over event time. We find that the relationship between betas and hazard rates is a well-estimated zero for all three Fama-French factors. The point estimates are actually negative for SMB and HML, and the economic magnitude of the relation is also extremely small: for example, a one-standard-deviation increase in the hazard rate corresponds to a increase in the market beta of around of 0.02. Overall, the results indicate that that there is no significant event-time variation in systematic risk as captured by the Fama-French factors.

5.3   Event-time variation in downside risk

An alternative explanation of the relation between hazard rates and returns is event-time variation in downside risk, that is, the risk of experiencing particularly bad returns during times when the market return is also very low (see Price, Price, and Nantell, 1982). Exposure to downside risk, similar to exposure to a short position in a put on the market portfolio, has been studied in previous research focusing on the Buy-and-Hold strategy. For example, Mitchell and Pulvino (2001) find that their Buy-and-Hold strategy is indeed exposed to downside risk, although the magnitude of the exposure is small and insufficient to explain the Buy-and-Hold alpha.

In this section we test whether the returns pattern can be explained by differential exposures to downside risk in event time. We start by calculating the raw performance of each strategy in periods in which the market return is low, defined as all months in which the market portfolio experiences a return below -3% (alternative cutoffs of -2%, -4%, and -5% yield similar results). Panel A of Table 5 shows that, during market downturns, betas increase slightly for all strategies. However, the downside betas of the High Hazard strategy do not significantly differ from the downside betas of the two Low Hazard strategies. In addition, the average returns of the three strategies are not particularly low. In months when the market loses more than 3%, only the Low Hazard 2 strategy has a negative average return. The reason is that part of our strategy invests in equity deals, which are protected from large downturns by the short position in the acquirer. In the Appendix, we further

explore downside risk by plotting the relation between strategy returns and market returns, allowing for a piecewise linear functional form (so that the beta in down markets can be different from the beta in normal markets). Figure A7 confirms that the exposure to downside risk is essentially the same across the three event-window strategies.

In panel B of Table 5, we measure exposure to downside risk by adding the out-of-the-money Put and Call factors constructed by Agarwal and Naik (2004) to the standard set of Fama-French factors. We find that the exposure of all strategies to these factors are extremely small and do not vary significantly across high and low hazard periods. In unreported results, we also find that exposures to the in-the-money Put and Call factors are similarly low and do not vary in event time. For completeness, we also estimate exposures to the momentum factor and again find that they are low and do not vary over event time.

We conclude our analysis of downside risk by presenting additional graphical evidence. In Figure 6 we plot the yearly return for the market, the High Hazard strategy, and the Buy-and-Hold strategy. The figure shows that the returns of both merger strategies are less volatile than the market return, and are not particularly exposed to aggregate downturns. Note also that the High Hazard strategy and the Buy-and-Hold strategy have similar volatility and are correlated, but the High Hazard strategy almost always has higher returns: this is the graphical counterpart to the higher alpha reported in Table 3.

5.4 Event-time variation in idiosyncratic risk

A final potential risk-based explanation for the correlation between hazard rates and returns is event-time variation in idiosyncratic risk. In cases in which merger completion comes as a surprise, prices should jump. Therefore, we may expect higher return volatility during event windows when hazard rates of completion are high. Why should idiosyncratic risk be priced at all? It's possible that the arbitrageurs that operate in merger markets are constrained to hold portfolios consisting only of mergers, and require compensation for holding a particularly volatile portfolio. As shown by Pontiff

(2006), idiosyncratic risk may matter even for diversified arbitrageurs due to holding costs of arbitrage positions in individual securities.

We find that variation in idiosyncratic risk cannot explain the correlation between hazard rates and returns. Instead of being hump shaped, the volatility of returns from a strategy that invests in all available deals increases in event time (the standard deviation of returns is 0.023, 0.039, and 0.068 for the Low Hazard 1, High Hazard, and Low Hazard 2 strategies, respectively).[22]

Finally, it is worth noting that we present conservative estimates of the return volatility that merger arbitrageurs are likely to experience. Our High Hazard strategy invests in approximately 27 deals on average in each calendar month, which is sufficient to obtain large diversification benefits that greatly reduce returns volatility. The lower fifth percentile of months by number of deals is diversified across nine deals. However, our sample of deals is artificially limited because we restrict our analysis to pure cash and equity-financed deals, which do not have contingency terms. A merger arbitrageur would likely be able to invest in a larger set of deals, such as mergers, that are financed using a mix of cash and equity.

5.5   Event-time variation in frictions

In this section, we consider three other potential explanations for the returns patterns that are consistent with rational markets. First, it is possible that event-time variation in buying and selling pressures leads to predictable returns patterns. In particular, conversations with merger arbitrageurs suggest that large institutional investors, such as mutual funds, tend to sell the target immediately after merger announcement to lock in gains and because the target's risk profile no longer fits with the fund's core strategy (e.g., a small-value investment fund). In fully efficient markets, these sell orders should not affect prices. However, if arbitrage capital moves slowly to take the other side of the trade, the selling pressure can take a while to disappear (Shleifer, 1986). This may generate low

---

[22]We also find that the High Hazard strategy displays large positive skewness of 3.75 relative to the two Low Hazard strategies (which display skewness of -0.46 and 2.63). Thus, the volatility of the High Hazard strategy partly reflects its disproportionate upside potential. To the extent that positive skewness is valued by investors, a rational model would actually predict lower expected returns for the High Hazard strategy, contrary to our findings.

returns immediately after announcement, followed by slowly increasing returns in the following days. If selling pressures are responsible for the observed pattern in returns, we expect that measures of liquidity should covary negatively with returns in event time.

Second, it is possible that variation in the degree of asymmetric information generates the returns patterns. Insiders may have greater access to rumors about merger completion than do others, and this informational advantage may be greatest during event periods when hazard rates are high. If asymmetric information is responsible for the observed patterns in returns, we expect that the bid-ask spread should covary positively with returns in event time.

Turning to the data, we find that the event-time variation in buying and selling pressure and asymmetric information, as proxied by volume, turnover, and bid-ask spread occur on a very different timescale than that of the return predictability we document. Figure 7 compares the evolution over event time of average returns with the evolution of the median volume, turnover, and bid-ask spread of target equity (each scaled by the value in the first week following the announcement). The figure shows that volume and turnover are very high in the one or two weeks following merger announcement, and then drop to a steady level starting in event week four. A very similar pattern occurs for the bid-ask spread, except that the drop towards the steady low level occurs even more immediately following merger announcement. All the variation in volume, turnover, and bid-ask spread is concentrated around announcement, and there seems to be no significant event-time variation in the months that follow, where the relevant variation in returns is concentrated.

Finally, we consider whether the hump-shaped returns pattern can be explained by frictions surrounding delisting returns and the "last day" effect. In the last day of trading before the merger formally completes and the target delists, deal completion is usually considered certain by all market participants because all parties have publicly agreed to the merger. However, the last recorded target stock price may still trade at a small discount to the "deal consideration," the final price paid by the acquirer for each share of the target. For example, the target may trade at $9.99 on the last day given an deal consideration of $10.00, a difference that corresponds to an additional 10 bp return

37

in a single day. Conversations with merger arbitrageurs suggest that this return may not always be realizable by investors because of illiquidity or fees. We show that our returns pattern cannot be explained by the last day effect by simulating a trading strategy that only earns returns based upon traded prices and does not earn any returns based upon the difference between the last traded price and the deal consideration. This simulation purges the strategy returns of any last day effects and is conservative in that it removes some of the positive returns that investors may have actually earned. We find our results qualitatively unchanged, although the alphas of all strategies are reduced. For brevity, we report our results below rather than in table form. The High Hazard alpha is 72 bp per month compared to 24 bp and 31 bp for the Low Hazard 1 and Low Hazard 2, respectively. The High Hazard strategy deliver significantly higher alphas than the Low Hazard strategies with a p-value of 0.026. In other words, the last day effect cannot explain the event-time variation in returns.

Overall, we show in this section that the returns pattern cannot be explained by event-time variation in systematic risk, downside risk, idiosyncratic risk, buying and selling pressure, asymmetric information, or last day frictions. All of these elements can contribute to the positive Buy-and-Hold returns. In addition, the presence of frictions can limit arbitrage as shown in the next section. However, they cannot by themselves generate the event-time variation in returns.

6   Limits to Arbitrage

To further support the behavioral model of underreaction, we explore why sophisticated arbitrageurs (who are likely to recognize the information content of the passage of time) allow the mispricing to persist. We find evidence suggestive of the existence of both behavioral and sophisticated investors. However, limits to arbitrage prevent sophisticated investors from fully arbitraging away the mispricing because of transaction costs.[23]

We first look at the Fama-French alphas when the strategies are executed on subsamples of mergers

---

[23]It is also possible that the hazard rate calculations we perform here seem simple once people understand them, but relatively few players have made these calculations in the past. We thank David Hirshleifer for this suggestion.

according to target characteristics that correlate with the transaction costs faced by arbitrageurs: total dollar volume, average daily turnover, bid-ask spread, and size (market cap) of the target. For each characteristic, we split the set of mergers occurring in each calendar year by the median value of the characteristic, as measured during the second week after announcement. We look at the characteristics after announcement to capture the features of the market during the time period when arbitrageurs are likely to operate. We also report results separately for early (pre-1991) and late (post-1991) calendar periods. This tests the idea that sophisticated arbitrage capital has increased over time, so later calendar periods may display less returns predictability.

Table 6 reports the Fama-French alphas of each strategy for the four stock characteristics plus the early and late calendar period division. For each characteristic, the left column corresponds to more difficult arbitrage conditions: small target market cap, low volume and turnover, high bid-ask spread, and the early sample period. For all characteristics, we find significantly higher alphas for the High Hazard strategy corresponding to the sample with more difficult arbitrage conditions.

In the Appendix, we also study the returns to trading strategies that take into account the transaction costs associated with the activity of a realistic merger arbitrage fund. We propose two methods to account for realistic direct and indirect transaction costs, both based on the "RAIM" strategy presented by Mitchell and Pulvino (2001). Table A2 shows that the alphas of all strategies are noticeably reduced and approach zero once we take into account transaction costs. However, our main finding that the alpha of the High Hazard strategy is significantly higher than the alphas of the two Low Hazard strategies still holds.

Taken together, the results in this section offer an explanation of why the behavioral underreaction is allowed to persist. Not all market participants underreact to the passage of time. A subset of investors, possibly small retail investors, are boundedly rational and generate the mispricing. More sophisticated investors are unable to fully arbitrage away the mispricing due to transaction costs.

7   Conclusion

The absence of news and the passage of time often contain important information. However, no news is likely to be less salient and vivid than traditional news stories. This may lead boundedly rational investors to underreact to the passage of time.

We test how markets react to the passage of time using the empirical context of mergers. Following the initial merger announcement, uncertainty relating to merger completion can take several months to a year to be resolved. We find that hazard rates of merger completion vary strongly over time after the merger announcement, implying that the passage of time can predict merger completion. If markets are rational, prices should correctly incorporate this information and average returns should be constant over event time absent any compensation for risk or frictions. When we examine target return patterns, we find that the aggregate merger completion hazard rates are positively correlated with target returns in event time.

We then investigate two possible explanations for this returns predictability. We first show that the positive correlation between returns and hazard rates can be explained by a behavioral model in which the agents underreact to the passage of time. If agents do not fully appreciate the variation in hazard rates associated with the passage of time, they will behave as if hazard rates are less time-varying (flatter) than in reality. This leads to periods in which agents over- or underestimate the true hazard rates of completion. When true hazard rates are high, they will be underestimated by the agents and returns will be high because of positive surprises from actual merger completions. When true hazard rates are low, they will be overestimated by the agents, and returns will be low because of negative surprises. This underreaction can explain the observed correlation between hazard rates and returns. We propose three empirical tests of the behavioral model that use returns and hazards data, and all the tests strongly support the implications of the model.

Although the positive relationship between returns and hazard rates is consistent with underreaction to no news, it could also be explained by changes in risk or other frictions over the event lives

of mergers. Using portfolios, we find that merger returns have low betas in general, and systematic risk does not vary with hazard rates over the event lives of mergers. We also show that downside risk and idiosyncratic risk do not vary in event time and cannot explain the observed pattern in returns. Finally, we show that event-time variation in selling pressures and asymmetric information is unlikely to explain the observed returns patterns. We conclude that aggregate hazard rates of merger completion predict merger returns because markets underreact to the information content of the passage of time.

Using the empirical context of mergers, we demonstrate that underreaction to the passage of time can be costly, resulting in returns variation of up to 72 bp per month as time passes after merger announcement. Of course, some investors are likely to be highly sophisticated and rational. We find evidence consistent with the existence of limits to arbitrage. We show that underreaction is concentrated in the subset of deals with lower liquidity and higher transaction costs, suggesting that trading frictions prevent sophisticated investors from arbitraging away the mispricing.

Evidence of underreaction in mergers markets is also suggestive of a more general phenomenon, in which agents underreact to the passage of time because it is often less salient than explicit news stories. Underreaction to no news can be persistent, and can potentially exacerbate asymmetric information problems in other contexts, such as the interactions between voters and politicians, managers and employees, or investors and insiders. We leave an exploration of the extent to which underreaction to no news pervades other contexts to future research.

References

Agarwal, V., and N. Y. Naik. 2004. Risks and portfolio decisions involving hedge funds. *Review of Financial Studies* 17:63–98.

Agostinelli, G., S. Sherman, R. Fazio, and E. Hearst. 1986. Detecting and identifying change: Additions versus deletions. *Journal of Experimental Psychology: Human Perception and Performance* 12:445–54.

Ahern, K., and D. Sosyura. 2014. Who writes the news? Corporate press releases during merger negotiations. *Journal of Finance* 69:241–91.

Andrade, G., M. Mitchell, and E. Stafford. 2001. New Evidence and Perspectives on Mergers. *Journal of Economic Perspectives* .

Bagnoli, M., W. Kross, and S. G. Watts. 2002. The information in management's expected earnings report date: A day late, a penny short. *Journal of Accounting Research* 40:1275–1296.

Baker, M., and S. Savasoglu. 2002. Limited arbitrage in mergers and acquisitions. *Journal of Financial Economics* 64:91–115.

Barber, B. M., and T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21:785–818.

Breen, W. J., L. S. Hodrik, and R. A. Korajczyk. 2002. Predicting equity liquidity. *Management Science* 19:119–157.

Campbell, J. Y., and L. Hentschel. 1992. No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* 31:281–318.

Chan, W. 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics* 70:223–260.

Cohen, L., and A. Frazzini. 2008. Economic Links and Predictable Returns. *Journal of Finance* 63:1977–2011.

Corwin, S., and P. Schultz. 2012. A simple way to estimate bid-ask spreads from saily high and low prices. *Journal of Finance* 67:719–59.

Corwin, S. A., and J. F. Coughenour. 2008. Limited attention and the allocation of effort in securities trading. *Journal of Finance* 63:3031–67.

Da, Z., U. G. Gurun, and M. Warachka. 2012. Frog in the pan: Continuous information and momentum. *Manuscript.*

DellaVigna, S., and J. M. Pollet. 2009. Investor inattention and friday earnings announcements. *Journal of Finance* 64:709–49.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg. 2006. Costly information acquisition: Experimental analysis of a boundedly Rational Model. *American Economic Review* 96:1043–68.

Galai, D., A. Raviv, and Z. Wiener. 2007. Liquidation triggers and the valuation of equity and debt. *Journal of Banking and Finance* 31:3604–20.

Gao, G., and Q. Ma. 2012. The sound of silence: What do we know when insiders do not trade? *Manuscript*.

Gifford, S. 2005. Limited attention as the bound on rationality. *B.E. Journal of Theoretical Economics* 5:1–42.

Gilbert, T., S. Kogan, L. Lochstoer, and A. Ozyildirim. 2012. Investor inattention and the market impact of summary statistics. *Management Science* 58:336–50.

Healy, A. 1981. The effects of visual similarity on proofreading for misspellings. *Memory and Cognition* 9:453–60.

Hearst, E. 1991. Psychology and nothing. *American Scientist* 79:432–43.

Hirshleifer, D., K. Hou, S. H. Teoh, and Y. Zhang. 2004. Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38:297–331.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289–325.

Hirshleifer, D., and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36:337–86.

Huberman, G., and T. Regev. 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *Journal of Finance* 56:387–96.

Jung, W.-O., and Y. K. Kwon. 1988. Disclosure when the market Is unsure of information endowment of managers. *Journal of Accounting Research* 26:146–53.

Kahneman, D. 1973. *Attention and effort.* Upper Saddle River, NJ: Prentice Hall.

Klein, J. P., and M. Moeschberger. 2003. *Survival analysis: Techniques for censored and truncated data, second edition.* New York: Springer.

Maki, R. 1989. Recognition of added and deleted details in scripts. *Memory and Cognition* 17:274–82.

Malmendier, U., M. Opp, and F. Saidi. 2011. Cash is King - Acquirer and target revaluation after merger bids. *Manuscript.*

Marin, J. M., and J. P. Olivier. 2008. The dog That did not bark: Insider trading and crashes. *Journal of Finance* 63:2429–76.

Menzly, L., and O. Ozbas. 2010. Market segmentation and cross-predictability of returns. *Journal of Finance* 65.

Merton, R. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42:483–510.

Mitchell, M., and T. Pulvino. 2001. Characteristics of risk and return in risk arbitrage. *Journal of Finance* 56:2135–75.

Neisser, U. 1963. Decision-time without Reaction-time: Experiments in Visual Scanning. *American Journal of Psychology* 76:376–85.

Pashler, H. 1998. *Attention.* East Sussex, UK: Psychology Press.

Pedzek, K., R. Maki, D. Valencia-Laver, T. Whetstone, J. Stoeckert, and T. Dougherty. 1988. Picture memory: Recognizing added and deleted details. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14:468–76.

Peng, L. 2005. Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis* 40:307–30.

Peng, L., and W. Xiong. 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80:563–602.

Pontiff, J. 2006. Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics* 42:35–52.

Price, K., B. Price, and T. Nantell. 1982. Variance and lower partial moment measures of systematic risk: Some analytical and empirical results. *Journal of Finance* 37:843–55.

Radner, R., and M. Rothschild. 1975. On the allocation of effort. *Journal of Economic Theory* 10:358–76.

Sargent, T. 1993. *Bounded rationality in macroeconomics.* Oxford: Oxford Unviersity Press.

Shleifer, A. 1986. Do demand curves for stocks slope down? *Journal of Finance* 41:579–90.

Simon, H. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69:99–118.

Tetlock, P. 2011. All the news that's fit to print: Do investors react to stale information? *Review of Financial Studies* 24:1481–512.

Treisman, A., and S. Gormican. 1988. Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review* 95:15–48.

Treisman, A., and J. Souther. 1985. Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General* 114:285–310.
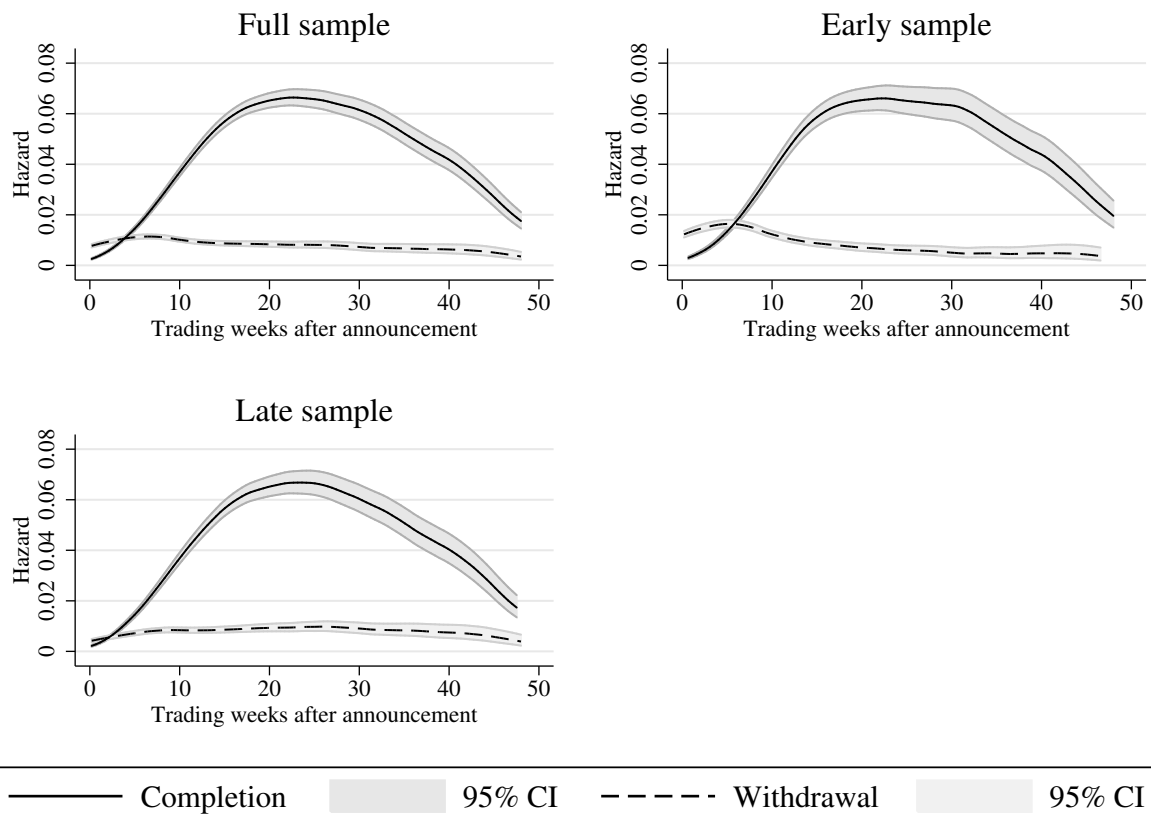
Figure 1. Hazard rates

The figure reports hazard rates of completion and withdrawal for all mergers over event time estimated using the Nelson-Aalen estimator of competing hazard rates. The top left panel uses the full sample, 1970–2010, to estimate the hazard rates. The top right panel uses data for the period 1970–1990. The bottom panel uses data for the period 1991–2010.
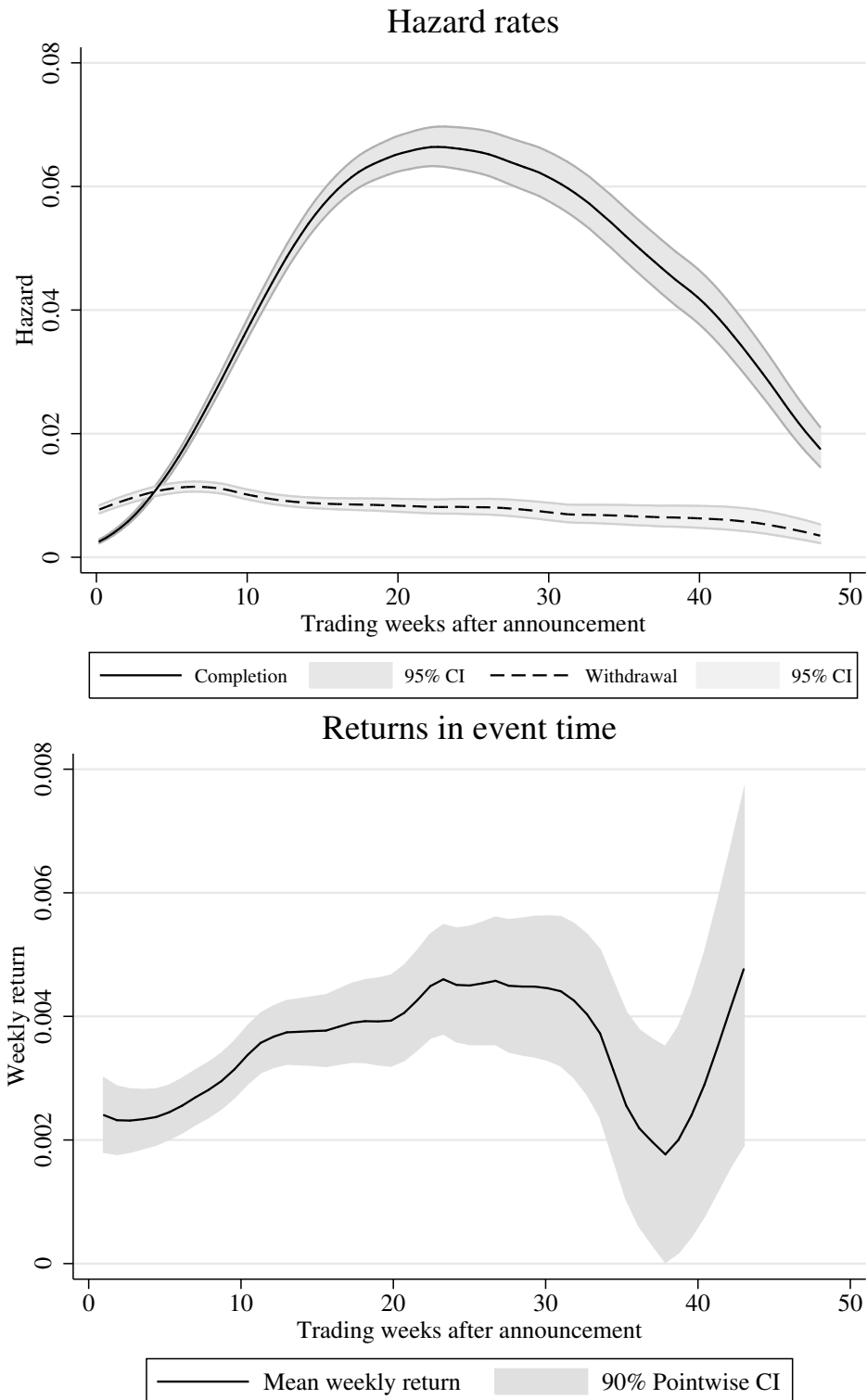
Figure 2. Hazard rates and mean weekly returns

The top panel reports estimated hazard rates for all mergers over event time, as in Figure 1. The bottom panel reports the average weekly return across deals over event time, as estimated using a local mean smoother. If a deal completes or withdraws before the end of an event week, the weekly return is calculated as the return from the beginning of the event week to the completion or withdrawal event (in these cases, we do not scale returns to represent full weekly returns).

46
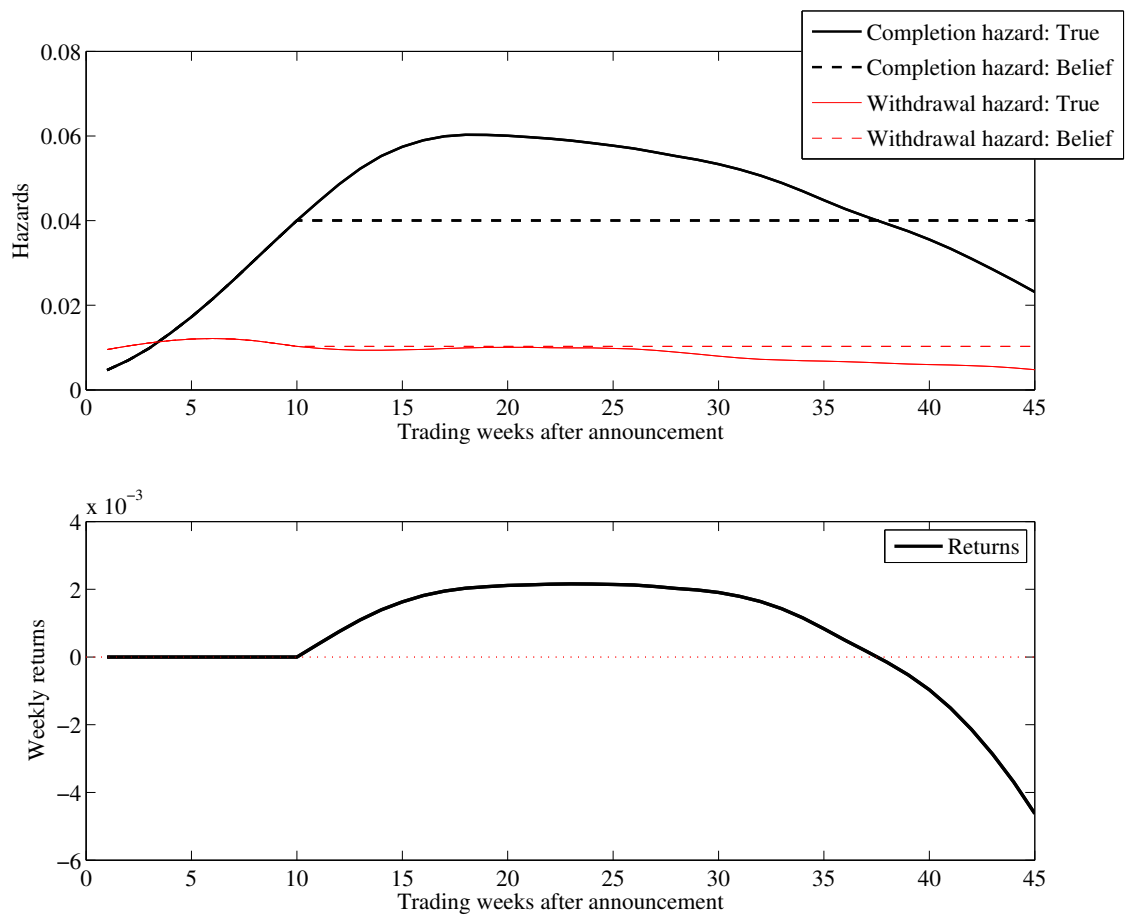
Figure 3. Model predictions of returns given beliefs

The figure reports an illustrative example of the returns implied by the model, obtained under a certain set of beliefs. The top panel reports the true hazard rates (solid lines) and an example of beliefs (dotted lines). The bottom panel plots the model-implied weekly excess returns over event time.

47
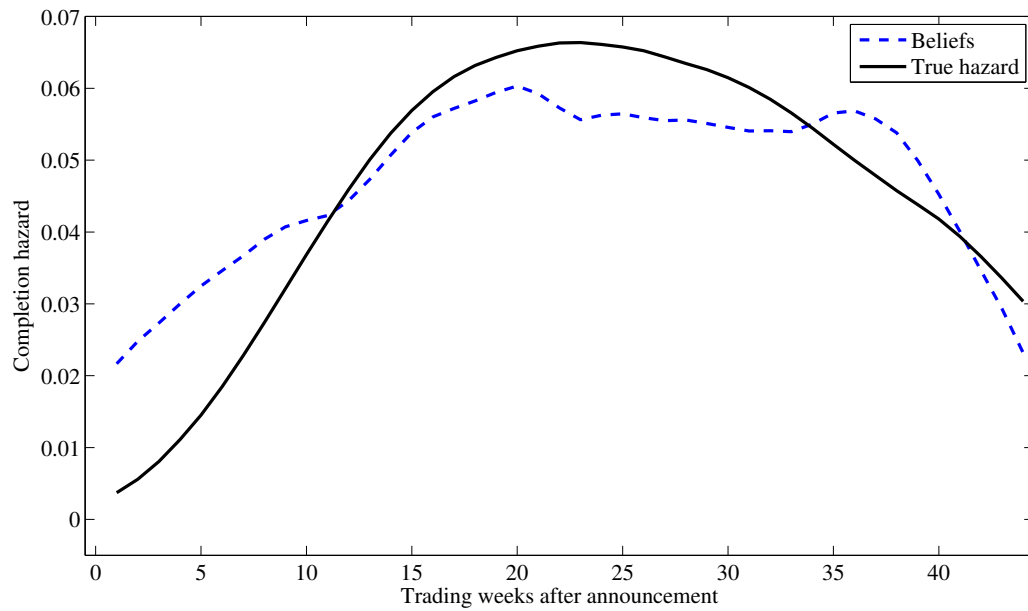
Figure 4. Implied hazard rates beliefs

The figure plots beliefs about hazard rates of completion over event time estimated from the model described in Section 4. Inputs into the model include estimates of the true hazard rates constructed in Section 2 and observed returns.
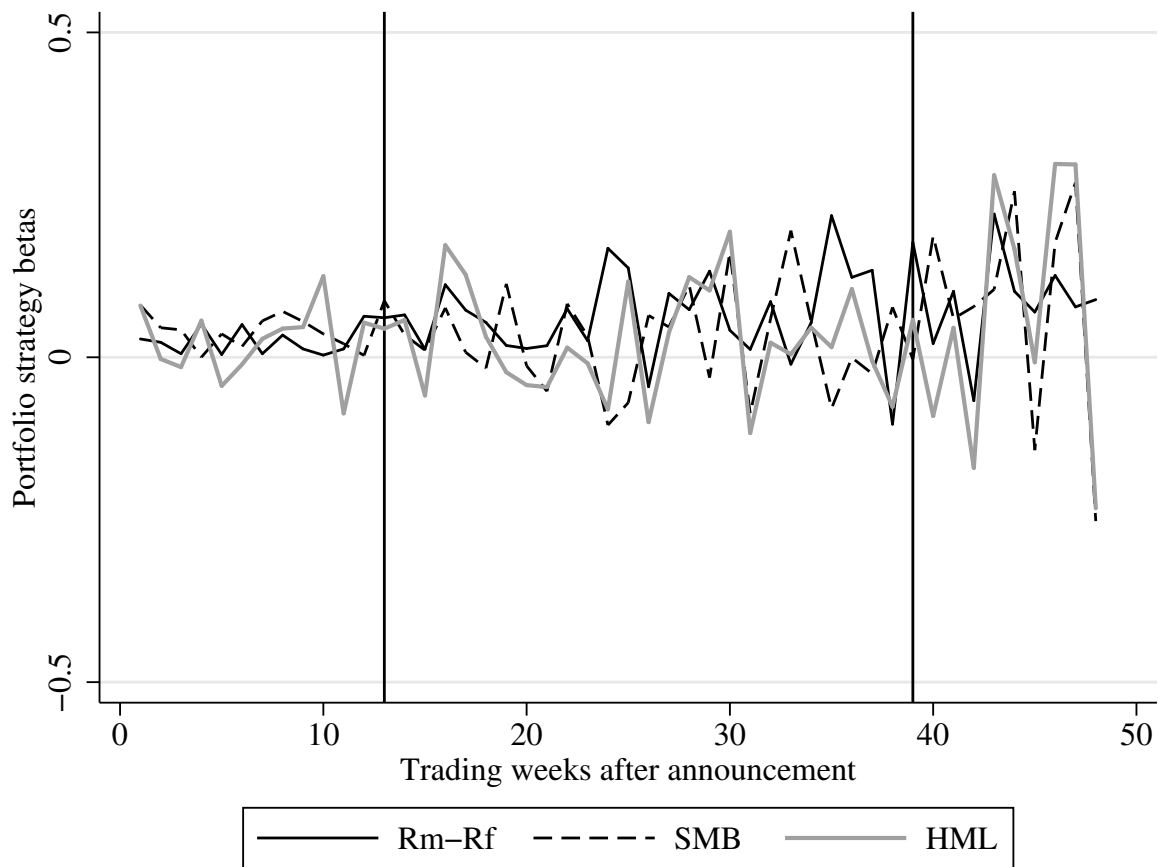
Figure 5. Event-time variation in risk

The figure reports betas with the three Fama-French factors separately for each event-week-specific trading strategy. The nth-week trading strategy only invests in deals that are active in the nth week after announcement. Betas are constructed using portfolio returns corresponding to calendar months for which at least one deal is present (this prevents betas corresponding to later event windows, in which fewer deals exist, from being biased toward zero). The vertical bars mark the three event windows for the Low Hazard 1, High Hazard, and Low Hazard 2 trading strategies described in Section 5.
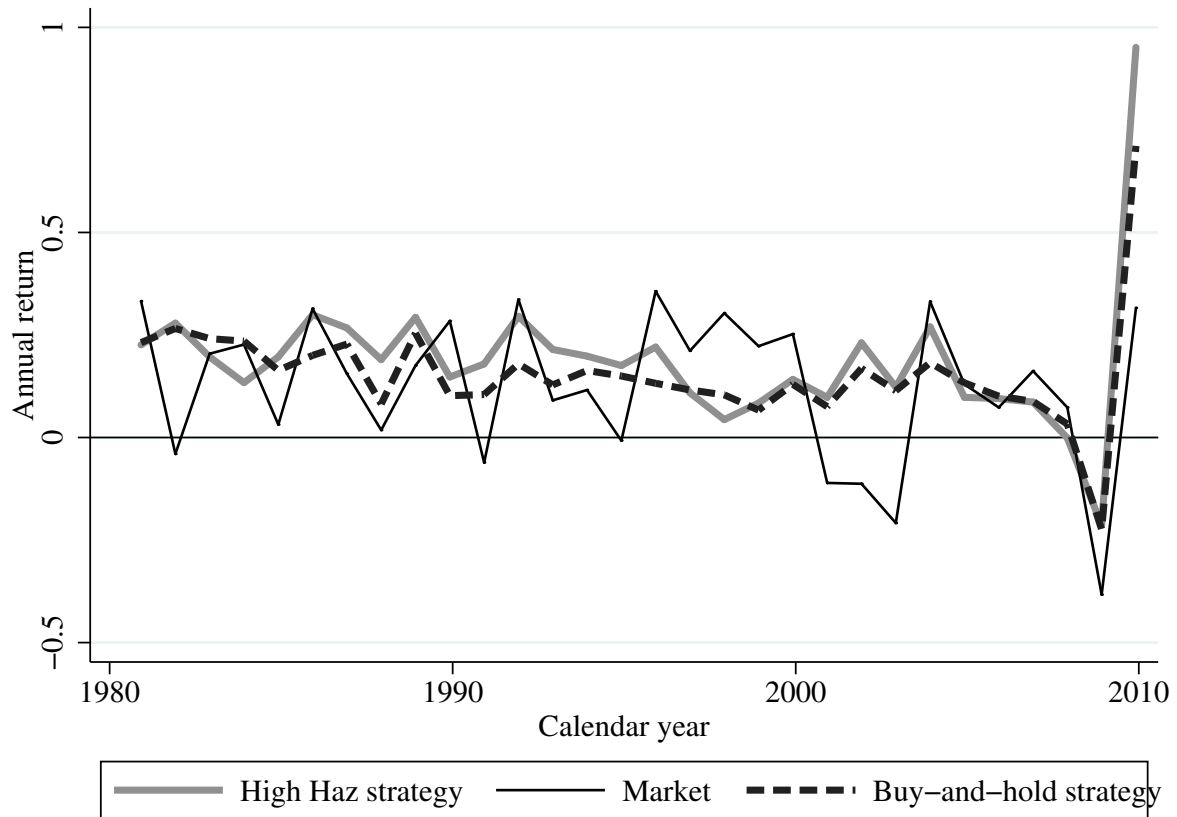
Figure 6. Portfolio returns comparisons
The figure plots the realized yearly returns of the High Hazard and Buy-and-Hold strategies and of the market return.
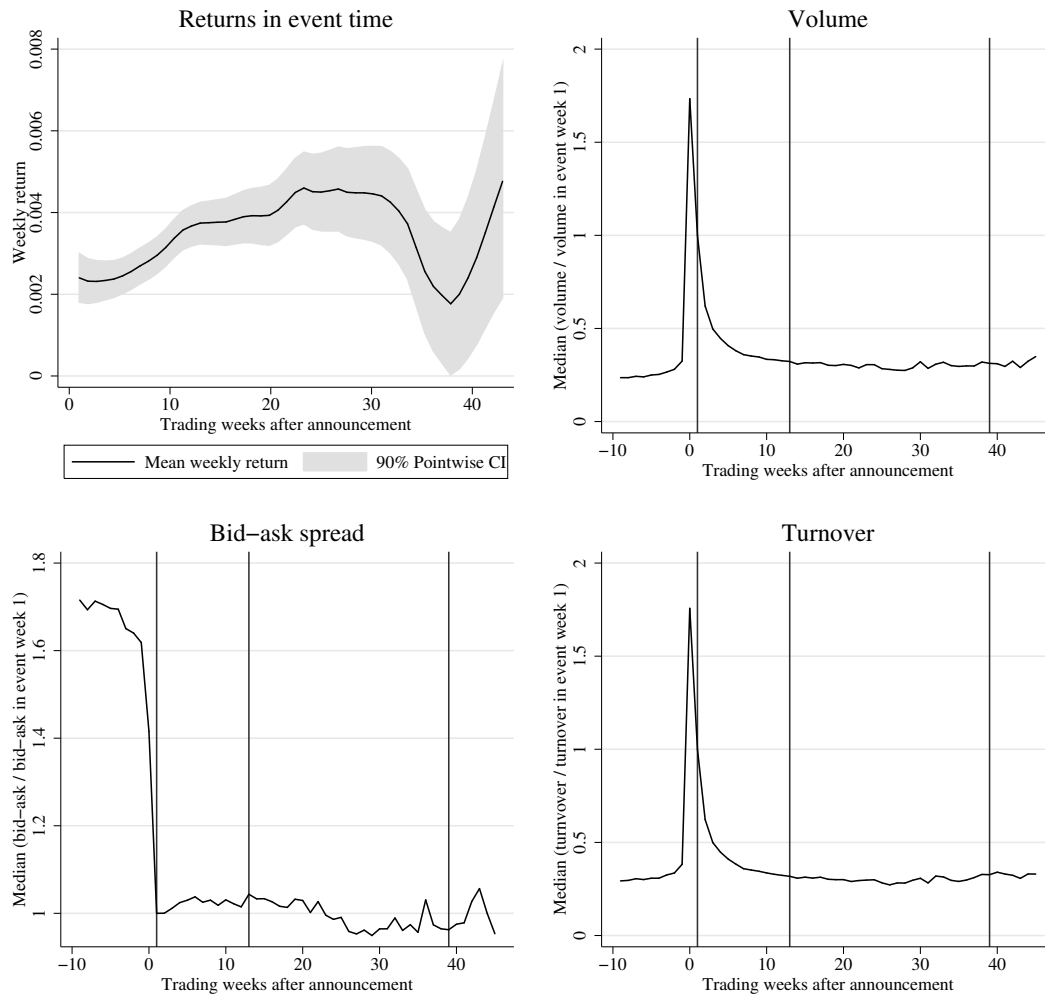
Figure 7. Event-time variation in volume, bid-ask spread, and turnover

The figure plots average returns together with median volume, median bid-ask spread, and median turnover, over event time. Median volume is the median of the target's average daily volume in each event week relative to the average daily volume in the week following the merger announcement. Median bid-ask spread is the median of the target's average daily bid-ask spread in each event week relative to the average daily bid-ask spread in the week following the merger announcement. Median turnover is the median of the target's average daily turnover in each event week relative to the average daily turnover in the week following the merger announcement.

	Mean	Median	SD
Number of deals	5,377		
Fraction equity financed	0.37		
% Completed within one year	72.8		
% Withdrawn within one year	21.0		
% Pending within one year	6.2		
Time to completion (trading days)	103.5	88.0	59.9
Time to withdrawal (trading days)	66.0	41.0	88.4
Premium	1.33	1.23	1.02
Size 1960–1979 ($mil)	63.7	23.4	132.7
Size 1980s ($mil)	231.9	50.7	756.3
Size 1990s ($mil)	472.2	100.9	2,077.7
Size 2000s ($mil)	1,033.4	187.7	3,346.5

Table 1. Summary statistics
This table reports summary statistics for the mergers in our dataset.

Dep var: Weekly return	Test returns flat over time: p-value	Hazard rate coefficient: Beta		Implied underreaction: Theta		Nonparametric test of implied underreaction: p-value	R^2	Obs.
All mergers	0.0010	0.0348 ***	(0.0077)	0.6013 ***	(0.0908)	0.0041	0.0121	90,044
Early sample	0.0000	0.0430 ***	(0.0110)	0.6198 ***	(0.1021)	0.0906	0.0143	38,760
Late sample	0.0000	0.0262 **	(0.0105)	0.5987 **	(0.1577)	0.0005	0.0106	51,284
Late sample, early hazards	0.0000	0.0248 **	(0.0100)	0.6186 **	(0.1512)	0.0179	0.0106	51,284
Cash	0.0797	0.0355 ***	(0.0103)	0.5360 ***	(0.1367)	0.0052	0.0178	62,587
Equity	0.0029	0.0390 ***	(0.0106)	0.6433 ***	(0.1013)	0.0920	0.0161	27,457
Small	0.0066	0.0653 ***	(0.0128)	0.3335 ***	(0.1361)	0.0005	0.0158	43,481
Large	0.2234	0.0107	(0.0084)	0.8516	(0.1100)	0.2987	0.0220	46,548
Diversifying	0.0009	0.0507 ***	(0.0184)	0.3502 ***	(0.2349)	0.0005	0.0224	35,312
Nondiversifying	0.0000	0.0241 *	(0.0143)	0.6339 *	(0.2088)	0.0630	0.0144	24,442
Low target price	0.0511	0.0675 ***	(0.0140)	0.3102 ***	(0.1461)	0.0128	0.0155	44,652
High target price	0.2123	0.0068	(0.0070)	0.9201	(0.0928)	0.0113	0.0271	45,377
Low news	0.0023	0.0442 ***	(0.0135)	0.4646 ***	(0.1667)	0.0005	0.0157	24,745
High news	0.0365	0.0140	(0.0132)	0.7930	(0.1731)	0.0865	0.0226	19,691

Table 2. Hazard rates versus returns

This table reports the results of an analysis of hazards and returns for all mergers combined (first row) and across groups of mergers sorted by characteristics as described in Section 4.3. Column (1) reports the p-value of a test that returns are constant in event time. Column (2) reports the results of regressions of weekly returns on hazard rates. The following specification is estimated: , where i indexes mergers; t is calendar time (year-month); and w is event week following merger announcement. γ_t represents a set of calendar time year-month fixed effects. h_t is the hazard rate, as estimated using the full aggregate sample or in the relevant subsample of mergers. Column (3) reports the estimate of the parameter θ of the underreaction model discussed in Section 4. Column (4) presents the p-value of a nonparametric test of the behavioral model, comparing the estimated true hazard curve with beliefs regarding hazard rates implied by the returns series using a Cramer-Von Mises test. All standard errors are allowed to be double clustered by calendar year-month and by merger. R-squareds correspond to the regression of returns on hazard rates presented in column (2).

53

	Strategy alpha		Test high > low: p-value
Panel A: Full sample			
Low hazard 1	0.0027 *	(0.0016)	
High hazard	0.0105 ***	(0.0016)	0.0004
Low hazard 2	0.0039 *	(0.0021)	
Buy and hold	0.0071 ***	(0.0016)	
Obs.	1,989		
R^2	0.0494		
Panel B: Late Sample, Early Hazards			
Low hazard 1	0.0016	(0.0025)	
High hazard	0.0071 ***	(0.0025)	0.0457
Low hazard 2	0.0016	(0.0031)	
Buy and hold	0.0051 **	(0.0025)	
Obs.	840		
R^2	0.0865		
Panel C: Tradable Strategy			
Low hazard 1	0.0027 *	(0.0015)	
High hazard	0.0103 ***	(0.0015)	0.0000
Low hazard 2	0.0022	(0.0015)	
Buy and hold	0.0071 ***	(0.0015)	
Obs.	2,256		
R^2	0.0461		

Table 3. Strategy alphas

Panel A of the table reports Fama-French alphas for the three portfolio strategies Low Hazard 1, High Hazard, Low Hazard 2, as well as for the Buy-and-Hold strategy. The "High > Low" column reports p-values for the null hypothesis that $(\alpha_{high} - \alpha_{low_1}) + (\alpha_{high} - \alpha_{low_2}) \leq 0$. We compute alphas using only calendar months in which a strategy invests in at least one active merger. We exclude months without active deals (rather than assuming that the strategy obtains the risk-free rate) to avoid biasing the alpha of the Low Hazard 2 strategy toward zero (Low Hazard 2 invests in deals toward the end of event life and therefore has fewer deals to invest in relative to the other two strategies). In panel B, we use hazard rates of completion estimated in the early period (prior to 1991) to determine the event window cutoffs for each of our three strategies. Then, we compute alphas using returns only during the later period (starting in 1991). In panel C, we compute the alphas and betas using all calendar month returns and assume that each strategy earns the risk-free rate in months in which there are no active deals in which to invest.

54

	Rm-Rf	SMB	HML
Panel A: Strategy betas			
Low hazard 1	0.0589	0.1262 **	0.0601
	(0.0379)	(0.0525)	(0.0569)
High hazard	0.1490 ***	0.1067 **	0.0442
	(0.0382)	(0.0527)	(0.0571)
Low hazard 2	0.1078 **	0.2797 ***	0.0846
	(0.0507)	(0.0678)	(0.0721)
Test high hazard = low hazard: p-value	0.1855	0.8435	0.7006
Panel B: Event-week betas and hazard rates			
Regression coefficient: weekly hazard	0.3860	-0.4860	-0.3758
	(0.4987)	(0.7187)	(0.8306)
Obs.	48	48	48
R^2	0.01	0.01	0.00

Table 4. Strategy betas

Panel A reports the betas with respect to the Fama-French factors for the different trading strategies, together with a test of the null hypothesis that betas do not vary over event time, separately for each of the three factors. The bottom panel reports the results of a regression of the Fama-French betas of each event-week-specific strategy on the estimated hazard rate of the corresponding event week. Betas and hazard rates are measured at the event week level.

Panel A: Rm < -3%	Average return	Alpha	Rm-Rf	SMB	HML	Obs	R2
Low hazard 1	0.0015 (0.0055)	0.0300 ** (0.0125)	0.3956 ** (0.1744)	0.1734 (0.1767)	0.0443 (0.1528)		
High hazard	0.0010 (0.0056)	0.0262 ** (0.0125)	0.3390 * (0.1744)	0.2415 (0.1770)	0.0588 (0.1529)	225	0.114
Low hazard 2	-0.0074 (0.0072)	0.0376 ** (0.0189)	0.5910 ** (0.2571)	-0.0397 (0.2162)	-0.2588 (0.1681)		

Panel B: Other risk factors	Alpha	UMD	OTM call	OTM put	Min return	Max return
Low hazard 1	0.0011 (0.0025)	-0.0094 (0.0504)	-0.0045 (0.0038)	-0.0020 (0.0043)	-0.11	0.13
High hazard	0.0091 *** (0.0025)	-0.1494 *** (0.0504)	-0.0062 (0.0038)	0.0031 (0.0043)	-0.111	0.41
Low hazard 2	0.0016 (0.0029)	-0.0386 (0.0559)	-0.0037 (0.0051)	-0.0052 (0.0049)	-0.24	0.62
Market return					-0.225	0.166

Table 5. Other risks

The top panel of the table reports the average return, alphas, and betas of the portfolio strategies conditional on restricting the sample to calendar months in which the market return is less than or equal to -3%. The bottom panel reports alphas and betas of a six-factor model that includes the three Fama-French factors (betas not reported in the table), the momentum factor UMD, and two option-based factors from Agarwal and Naik (2004). Note that the alphas and sample sizes are slightly different from our baseline case because these option factors are only available from 1983 to 2011.

hi>lo1				0	
hi>lo2				1	
hi>bh				0	

Joint	SD				Skewness
Lo1	0.0231	hi>lo1		0	-0.457
Hi	0.0387				3.754
Lo2	0.0679	hi>lo2		1	2.633
BH	0.0234	hi>bh		0	0.2

	Size		Volume		Turnover		Bid-ask spread		Time period	
	Small	Large	Low	High	Low	High	High	Low	Early	Late
Low haz 1	0.0028	0.0016	0.0020	0.0016	0.0021	0.0022	-0.0006	0.0026	0.0034	0.0016
	(0.0019)	(0.0019)	(0.0020)	(0.0020)	(0.0021)	(0.0020)	(0.0022)	(0.0022)	(0.0021)	(0.0025)
High haz	0.0141 ***	0.0063 ***	0.0119 ***	0.0051 **	0.0114 ***	0.0064 ***	0.0105 ***	0.0023	0.0128 ***	0.0070 ***
	(0.0020)	(0.0020)	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0022)	(0.0022)	(0.0021)	(0.0025)
Low haz 2	0.0028	0.0030	-0.0012	0.0024	-0.0008	0.0032	0.0028	0.0002	0.0055 *	0.0034
	(0.0031)	(0.0031)	(0.0032)	(0.0033)	(0.0032)	(0.0036)	(0.0030)	(0.0035)	(0.0032)	(0.0029)
Buy & hold	0.0081 ***	0.0048 **	0.0071 ***	0.0037 *	0.0071 ***	0.0047 **	0.0047 **	0.0033	0.0084 ***	0.0051 **
	(0.0019)	(0.0019)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0022)	(0.0022)	(0.0021)	(0.0025)
p-value: Illiquid > Liquid	0.0026		0.0105		0.0486		0.0048		0.0393	
p-value: High haz > Low haz	0.0000	0.0670	0.0000	0.1401	0.0001	0.1030	0.0006	0.3828	0.0034	0.0771
Obs.	3,713		3,665		3,655		2,171		1,989	
R^2	0.0447		0.0502		0.0449		0.0635		0.0677	

Table 6. Limits to arbitrage

The table reports alphas for trading strategies after dividing the sample in half by target characteristics (Columns 1–8) or by calendar time (Columns 9–10). The sample splits by target characteristics are executed separately within each calendar year using the median target characteristic in that year as measured two days following merger announcement. At the bottom of the table we report p-values for the test "High > Low," corresponding to the similar test in Table 3. We also report a test "High: Illiquid > Liquid", which tests the null hypothesis that the alphas in the environment in which arbitrage is more difficult (left column of each subdivision) is less than or equal to the corresponding alpha in the environment in which arbitrage is more favorable (right column of each subdivision).

57

Appendix

A.1 Data Filters

The following table reports the sample size at each step of the filtering, with the Thomson One (TO) dataset in the first column and the Mitchell and Pulvino (MP) dataset in the second column.

Step	# Obs. (TO)	# Obs. (MP)
Starting sample size	40,617	15,587
Keep only pure cash or equity mergers, without collar agreements	14,481	11,644
Exclude tender offers with known expiration dates	8,039	8,320
Exclude deals that compete with previous bids in the past 3 years	7,688	6,447
Target (and acquirer if an equity-financed deal) matches with CRSP returns	3,352	5,916
Keep deals for which offer price is above target price 2 days after deal announcement	3,051	4,974

Because MP covers a longer time series, while TO offers more comprehensive coverage of recent years, we use the MP dataset for years up to and including 1995 and the TO dataset afterward. After the combination, we are left with 5,377 deals in total. Our results are robust to using only MP or TO data.

A.2 Proposition 1

Proposition 1. Suppose that the true hazard rate for merger i is $h_i(t) = \alpha_i h(t)$, where $h(t)$ is an unobserved common component and α_i is a merger-specific unobservable parameter distributed in the cross-section according to the distribution function $G(\alpha)$ with mean equal to 1. Then we have

$$h(t) \geq h_\theta(t) \ \forall t$$

where $h_\theta(t)$ is the measured hazard rate that ignores the unobserved heterogeneity. In addition, if $h(0) = h(T) = 0$, $h(t)$ has to be more time-varying than $h_\theta(t)$.

Proof: Hazard rates are defined such that

$$h_\theta(t) = \frac{\int_0^\infty -\alpha S(t)^{\alpha-1} S'(t) g(\alpha) d\alpha}{\int_0^\infty S(t)^\alpha g(\alpha) d\alpha}$$

and

$$h(t) = E_\alpha(\alpha h(t)) = \int_0^\infty -\alpha S(t)^{-1} S'(t) g(\alpha) d\alpha.$$

Rearranging terms yields

$$E_\alpha(\alpha h(t)) = \int_0^\infty -\alpha S(t)^{-1} S'(t) g(\alpha) d\alpha = \int_0^\infty \frac{-\alpha S(t)^{\alpha-1} S'(t)}{S(t)^\alpha} g(\alpha) d\alpha.$$

For ease of notation call

$$X = -\alpha S(t)^{\alpha-1} S'(t)$$

$$Y = S(t)^\alpha$$

with $E[Y] \geq 0$ because $S(t) \geq 0$ and $\alpha \geq 0$. We have

$$E_\alpha(\alpha h(t)) = E[\frac{X}{Y}]$$

$$h_\theta(t) = \frac{E[X]}{E[Y]}.$$

Now define $Cov(\frac{X}{Y}, Y) = E[X] - E[\frac{X}{Y}]E[Y]$. It follows that

$$Cov(\frac{X}{Y}, Y) = Cov(-\alpha S(t)^{-1} S'(t), S(t)^\alpha) \leq 0$$

because $S(t)^{-1} \geq 0$, $S'(t) \leq 0$ and $Cov(\alpha, S(t)^\alpha) \leq 0$ ($S(t)^\alpha$ is decreasing in α because $S(t) < 1$).

This implies

$$E[\frac{X}{Y}]E[Y] \geq E[X]$$

and because $E[Y] > 0$ we can write:

$$E[\frac{X}{Y}] \geq \frac{E[X]}{E[Y]}$$

or:

$$E_\alpha(\alpha h(t)) \geq h_\theta(t).$$

Finally, if $h(0) = h(T) = 0$, then both $h(t)$ and $h_\theta(t)$ are 0 at 0 and T, so that $h(t)$ has to vary at least as much as $h_\theta(t)$ between 0 and T.

A.3 Hazard Heterogeneity

In this section we explore hazard rates of completion after accounting for additional observable deal heterogeneity. Note that our behavioral model focuses on the location of the hump in event time (e.g., early or late in event time) rather than the vertical scale of the hazard rate. A traditional Cox proportional hazard model cannot model this type of heterogeneity because the Cox model imposes that the shape of the underlying hazard curve is equal for all deals, and the individual characteristics only affect the scale of the whole curve proportionally. Therefore, we nonparametrically estimate the entire hazard curve for groups of deals subdivided according to their characteristics.

We collect information about twelve deal characteristics, all available at the time of announcement:

1. Financing: cash or equity-financed, as described in the text.

2. Size: small or large target size, as described in the text.

3. Calendar time period: early or late time period, as described in the text.

4. Merger arbitrage spread: the relative difference between the effective offer price and the target price two days after merger announcement. The spread is a proxy for the market's assessment of the probability that the merger will eventually complete. The spread for cash mergers is the ratio of the initial offer price at deal announcement to the price of the target two days after deal announcement. For equity mergers, the spread is defined as $\Delta * P_{t=2}^A / P_{t=2}^T$, where Δ is the exchange ratio, defined as the number of acquirer shares offered for each share of the target, and P^A and P^T are the acquirer's and target's share prices, respectively.

5. Premium: the relative difference between the effective offer price and the target price two days *before* merger announcement (so it reflects the offered price relative to the pre-announcement price). It is computed similarly to the merger arbitrage spread.

6. Drop-dead date: the date after which the merger may be terminated by any of the parties. In some cases, this date is changed during the merger, and it does not exactly correspond to an expected date for completion. Our main datasets (TO and MP) do not report drop-dead dates or expected completion dates. Therefore, we merge our data with the FactSet database, which contains drop-dead dates for a subset (12%) of mergers in our sample. We use this date to estimate a hazard rate curve for these deals that takes into account the passage of time *relative* to this drop-dead date. In particular, we estimate a hazard rate of completion after rescaling each deal's completion or withdrawal dates by the drop-dead date. In this rescaled time, the drop-dead date corresponds to event time=1, and the completion or withdrawal dates are expressed as a fraction of 1. For example, if a deal has a drop-dead date 150 days after announcement and it completes on day 100, in rescaled time, completion occurs at time= 2/3. After estimating the hazard curve in relative time, we compute the hazard curve for each deal in the original calendar time by inverting the time scaling by drop-dead date.

7. Merger wave: we categorize deals based upon whether they occur in calendar months in which many other deals or few of them occur. Besides being interesting in itself (do deals that occur during waves look different from others in terms of hazard rates?), calendar periods with many mergers also proxy for periods with regulatory changes that may have encouraged the prevalence of mergers (Andrade, Mitchell, and Stafford, 2001).

8. Diversifying mergers: we categorize as diversifying those mergers in which the acquirer and the target belong to different industry sector codes (58 industries) as categorized by Thomson One. In our sample, 57% of mergers are diversifying. Because data on the industry of the acquirer is not available for private acquirers in the MP sample, we use all available TO observations to estimate the hazard rates instead of data from our combined TO and MP sample.

9. Target share price: deals involving targets with above or below median price. The low-price sample includes penny stocks but our median cutoff offers a more equal split of the data into two samples.

10. Floating or fixed exchange ratio: whether equity-financed mergers use a fixed or floating exchange ratio. We find that almost all equity-financed mergers use a fixed exchange ratio (we match data on fixed or floating exchange ratio to 25% of our sample, and find that 96% of matched observations use fixed exchange ratios). We estimate the hazard rate restricted to equity-financed mergers known to use fixed exchange ratios.

11. Friendly or hostile: hostile takeovers represent a very small portion of our sample (7%), so we study the hazard rate for the sample restricted to friendly mergers.

12. Termination fee: whether there is a high or low termination fee for the target. Data on the termination fee comes from TO, which covers 44% of the TO sample. Because termination fee data is only available for SDC, we use all available SDC observations to estimate the hazard rates instead of data from our combined TO and MP sample.

For the continuous characteristics described above, we divide deals into two groups of equal size (with high and low values of the characteristic) and compute separate completion hazard rate curves. For binary measures, we plot completion hazard rates for the samples as described above. In unreported results, we also look at heterogeneity in withdrawal hazard rates and find little variation. As Figures A2 and A3 show, in many cases the difference in hazard rates across deals with different characteristics is a scale effect and the location of the hump does not differ significantly. Only with respect to the type of financing, size, diversification, target share price, and drop-dead dates does the location of the peak of the hump vary across groups.

We next study the relation between hazards and returns once we take into account heterogeneity. We focus on the characteristics for which there are significant differences in the location of the peaks between groups (financing, size, diversifying, target share price, and drop-dead date), and divide our set of deals into eight groups depending on the combined characteristics. We do not perform a full intersection of these characteristics (which would produce 2^n groups) because some of these intersections lead to groups that are too small to allow estimation of a within-group hazard rate. The eight groups are:

1. cash, small, diversifying, low price,

2. cash, small, diversifying, high price,

3. cash, small, nondiversifying, low price,

4. cash, small, nondiversifying, high price,

5. cash, large,

6. equity, small, low price,

7. equity, small, high price,

8. equity, large.

For each group, we estimate the hazard curve separately within the group. We then assign to each individual deal the hazard curve of the group it belongs to. Finally, for deals for which we have the drop-dead date, we assign to each deal the hazard rate computed as described above (using the timing of completion relative to the drop-dead date). At the end of this procedure, every deal has been assigned a hazard rate, computed as the hazard curve of the characteristics group or, if available, the one computed using the drop-dead date.

Table A1 reports the regression of returns on hazards as in Table 2. Contrary to Table 2, where each deal within a regression sample has the same hazard rate, here hazard rates are different across deals. To compare the magnitude and precision of the estimates, we standardize the hazard curve within each category. For comparison, column 1 reports a similar regression as that in Table 2, except

61

hazard rates have been standardized. Next, we report the regression with hazard rates obtained from the eight groups without using information from the drop-dead date (column 2). Column 3 uses the hazards obtained from the eight groups but substitutes them with the ones obtained using the drop-dead date for deals for which this is available. For each specification, we present results with calendar time (year x month) fixed effects.

The table shows that, after conditioning on multivariate deal characteristics, the relationship between hazards and returns becomes slightly stronger. These results suggest that imperfectly accounting for heterogeneity acts as a bias against our finding of a strong relationship between hazard rates and returns when we only condition on the form of financing. However, the gains from accounting for additional heterogeneity are very small. To account for heterogeneity, we must estimate the hazard rates within smaller and smaller subsamples. Some of the improvement in estimation is offset by the more noisy estimate of the hazard rate within each group.

A.4 Estimating θ

Under the assumption that beliefs about completion hazard rates are:

$$\hat{h}(t) = (1 - \theta)\overline{h} + \theta h(t),$$

the parameter θ can be estimated in the following way: First, substitute the expression for $\hat{h}$ into Equation (4) and assume that beliefs regarding withdrawal hazard rates are always correct:

$$E\left[ret_t\right] - rdt = \left(\frac{P_C}{\hat{P}(t)} - 1\right)(1 - \theta)[h(t) - \overline{h}]dt.$$

Second, consider the term

$$C_t = \left(\frac{P_C}{\hat{P}(t)} - 1\right)(h(t) - \overline{h}).$$

C_t depends on the average price of deals surviving until time t, $\hat{P}(t)$, and the hazard rate at time t, $h(t)$. Therefore, C_t can be constructed separately for each group of deals. We can then rewrite

$$E\left[ret_t\right] - rdt = (1 - \theta)C_t dt$$

so that $(1 - \theta)$ is the coefficient of a regression of returns on C_t.

A.5 Competing Bids

We study the arrival rate of competing bids from other potential acquirers. Figure A5 reports the hazard rate of competing bids, showing that it is slightly higher in the weeks immediately following merger announcement than in later weeks (although it is relatively flat compared with hazard rates

of completion).

A.6 Robustness to Cutoff Dates

We show that the results of Table 3 are robust to perturbations to the timing of the event-window cutoffs.

Figure A6 reports the difference between the alpha of a strategy that only invests in deals active between event weeks t_1 to t_2, and the alpha of the Buy-and-Hold strategy. t_1 can be read on the vertical axis, and t_2 on the horizontal axis.

For example, the Low Hazard 1 strategy invests in deals active between event weeks 1 and 10, so its alpha can be read as $(t_1 = 1, t_2 = 10)$, which corresponds to the circle at the bottom left. The circle in the middle corresponds to the High Hazard strategy and the circle on the top right corresponds to the Low Hazard 2 trading strategy. The bottom-right corner $(t_1 = 1, t_2 = 45)$ corresponds to the Buy-and-Hold strategy. Because the graph reports the alphas of the strategies relative to the Buy-and-Hold strategy, it is not surprising to find exactly 0 at (1,45), negative numbers for the Low Hazard strategies and a positive number for the High Hazard strategy.

Starting from the circles representing the cutoffs for our three trading strategies, it is straightforward to see that perturbations to the cutoff points in all directions do not dramatically affect the alphas. The Low Hazard strategies lie in areas with low alphas (relative to the Buy-and-Hold strategy). Meanwhile, the High Hazard strategy lies in an area with high alphas (relative to the Buy-and-Hold strategy). This shows that the strategy alphas do not strictly depend on the cutoff points but more generally align well with the high and low return event windows predicted by the underreaction hypothesis.

A.7 Downside Beta

In this section we replicate the estimation of downside beta proposed by Mitchell and Pulvino (2001). In particular, we estimate the coefficients of the regression:

$$R_{strat} - R_f = (1 - \delta)[\alpha_{mktlow} + \beta_{mktlow}(R_M - R_f)] + \delta[\alpha_{mkthigh} + \beta_{mkthigh}(R_M - R_f)] + \epsilon,$$

where δ is an indicator that the monthly market return is below a threshold (which we take to be -4% as do Mitchell and Pulvino 2001); R_{strat} is the return of the strategy (we perform the exercise for our Low Hazard 1, Low Hazard 2, and High Hazard strategies separately); and R_M is the market return. To ensure continuity of the relation between expected strategy returns and expected market returns, we constrain the coefficients to satisfy

$$\alpha_{mktlow} + \beta_{mktlow}(-4\%) = \alpha_{mkthigh} + \beta_{mkthigh}(-4\%).$$

Figure A7 plots the estimated piecewise linear functions for the three strategies. The figure shows that the three event-time strategies have similar exposures to downside risk.

A.8 Transaction Costs

We study the returns to trading strategies that take into account the transaction costs associated with the activity of a realistic merger arbitrage fund. Because it is difficult to precisely estimate trading costs, we present results using two different methods.

The first method closely follows the "RAIM" strategy presented by Mitchell and Pulvino (2001).

We simulate four funds that trade in the four main strategies considered in this paper (Low Hazard 1, High Hazard, Low Hazard 2, and Buy-and-Hold), starting with $1M of capital each in 1970. Every month, each fund invests equally in all available deals in their respective event windows subject to the following limits to their positions. At most 10% of the capital can be invested in any particular deal, and the total trade in any deal cannot produce price pressure of more than 5% of the price. We estimate price pressure following the procedure described in Breen, Hodrik, and Korajczyk (2002).

In addition to limiting the size of the position and the amount of trading that each fund can perform, we compute direct transaction costs the fund incurs (e.g., broker commissions).[24] We also compute indirect transaction costs, which capture lower returns because of price impact and illiquidity, using the estimates of Breen, Hodrik, and Korajczyk (2002).

We also present results using a second method to account for trading costs, which modifies the procedure described above by proxying for indirect transaction costs using the bid-ask spread. We obtain the bid-ask spread for each trade from CRSP, and when that is not available, we use the bid-ask spread estimated following Corwin and Schultz (2012). For firms for which neither are available, we set the bid-ask spread to the average among firms in our sample by year and size category.

Table A2 shows that the alphas of all strategies are noticeably reduced once we take into account transaction costs, with Method 2 reducing the alphas by more than Method 1. Although transaction costs reduce the alphas of all strategies, our main finding that the alpha of the High Hazard strategy is significantly higher than the alphas of the two Low Hazard strategies still holds under both methods.

---

[24]Following Mitchell and Pulvino (2001), we approximate direct trading costs by assuming a fixed dollar cost per share traded: $0.1 before 1980, $0.05 between 1980 and 1990, and $0.04 between 1990 and 1998. Given the advances in technology after 1998, trading costs have fallen further. For all deals after 1998, we use the conservative estimate of $0.03 per share.
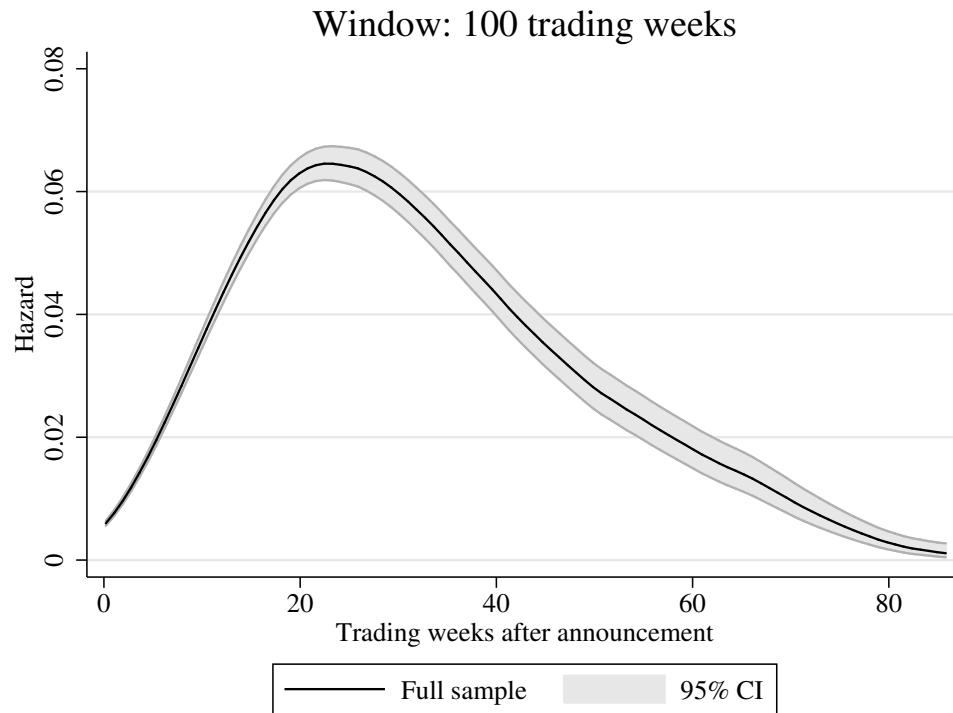
Appendix Figures and Tables

Figure A1. Hazard rates over a longer event window

This figure reports the estimated hazard rates of completion up to 100 weeks after announcement. The hazard rate of completion is as defined in Section 2.
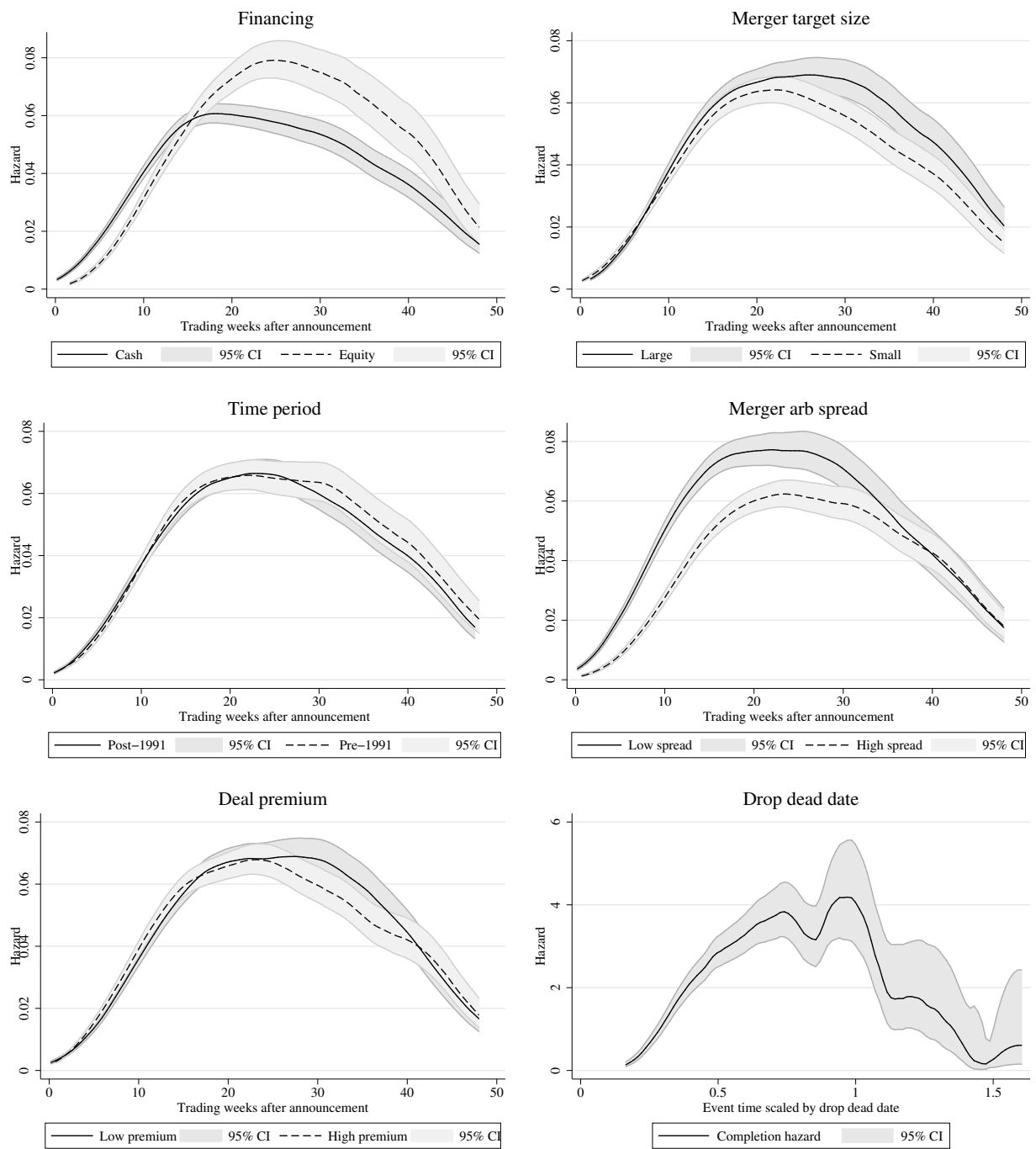
Figure A2. Heterogeneity (1)

This figure reports the estimated hazard rates of completion for each of the characteristics considered. The hazard rate of completion is as defined in Section 2.
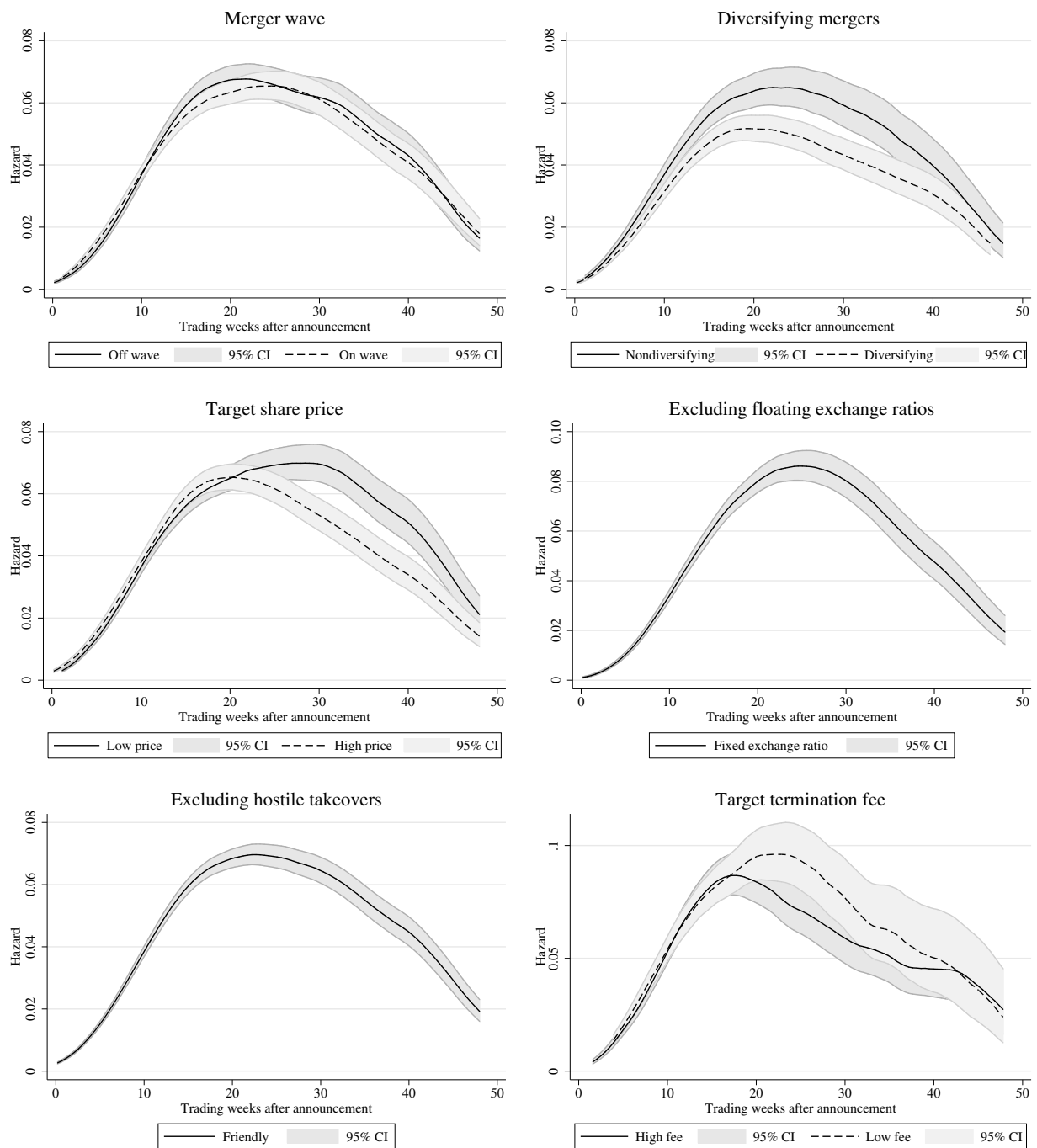
Figure A3. Heterogeneity (2)

This figure reports the estimated hazard rates of completion for each of the characteristics considered. The hazard rate of completion is as defined in Section 2.
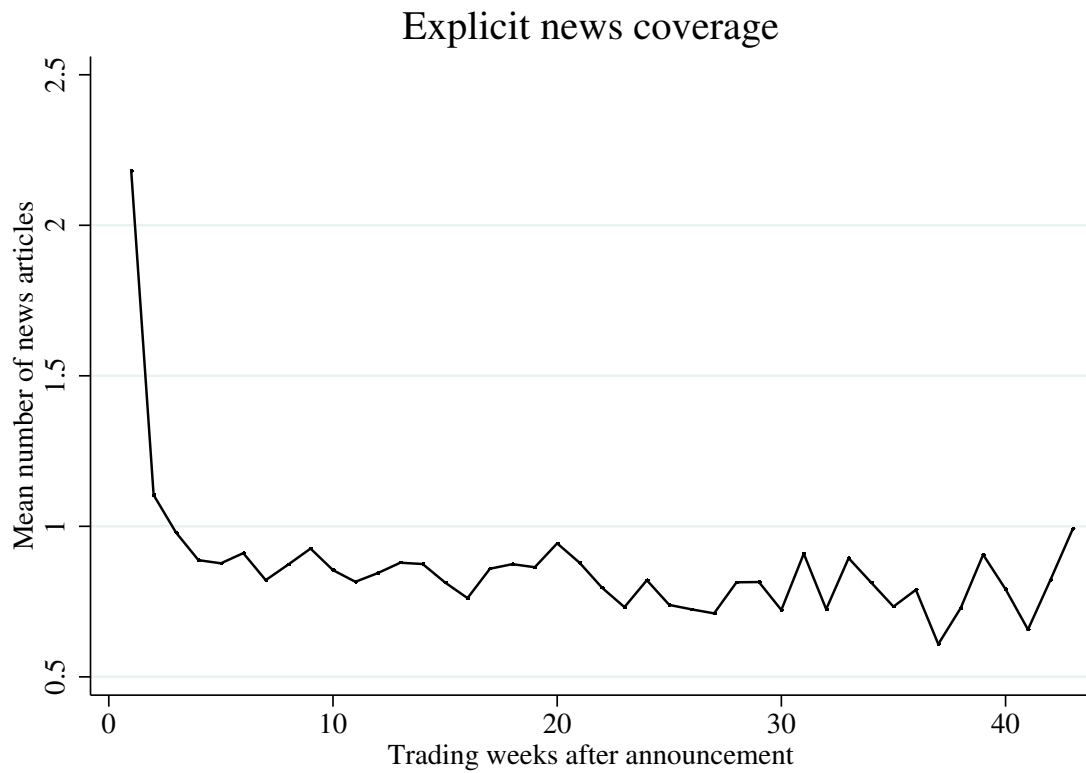
Figure A4. Explicit news coverage

This figure reports average prevalence of explicit news coverage of the target in each event week in the year after merger announcement.
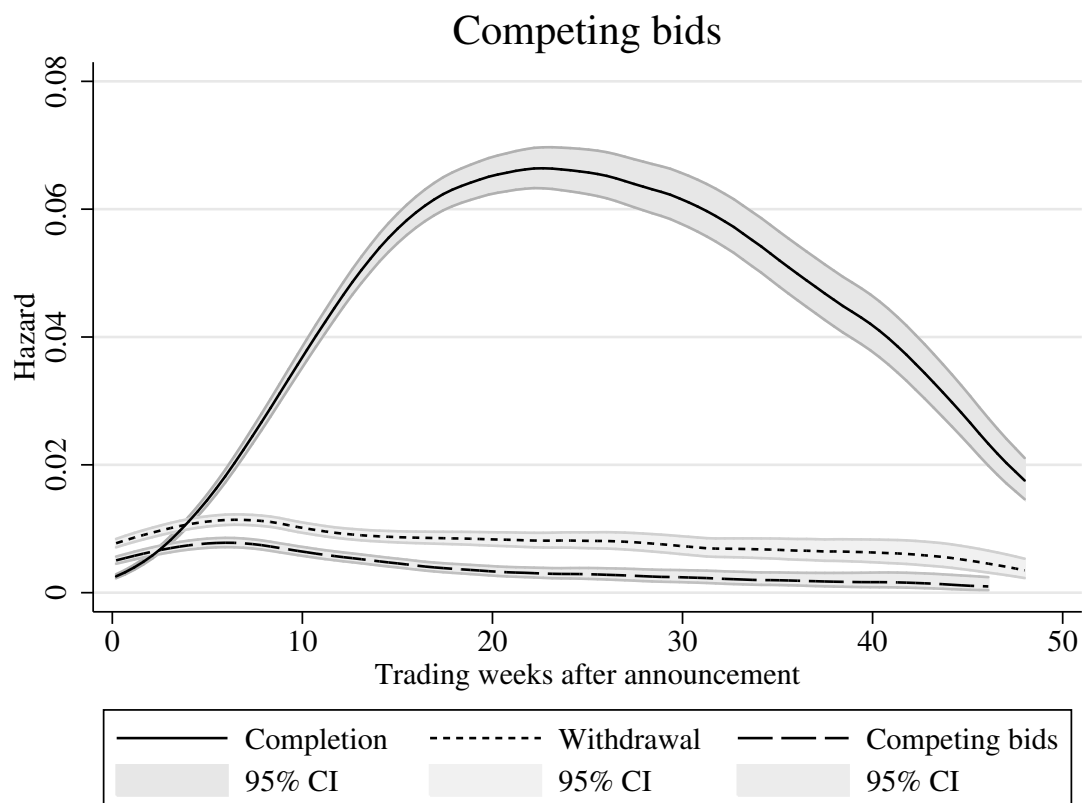
Figure A5. Hazard rates of completion, withdrawal, and competing bids
This figure reports the estimated hazard rates of completion, withdrawal, and of receiving a competing bid from another acquirer. The hazard rate of receiving a competing bid represents the probability of receiving at least one competing bid in event week t conditional on no completion or withdrawal prior to week t. The hazard rates of completion and withdrawal are as defined in Section 2.
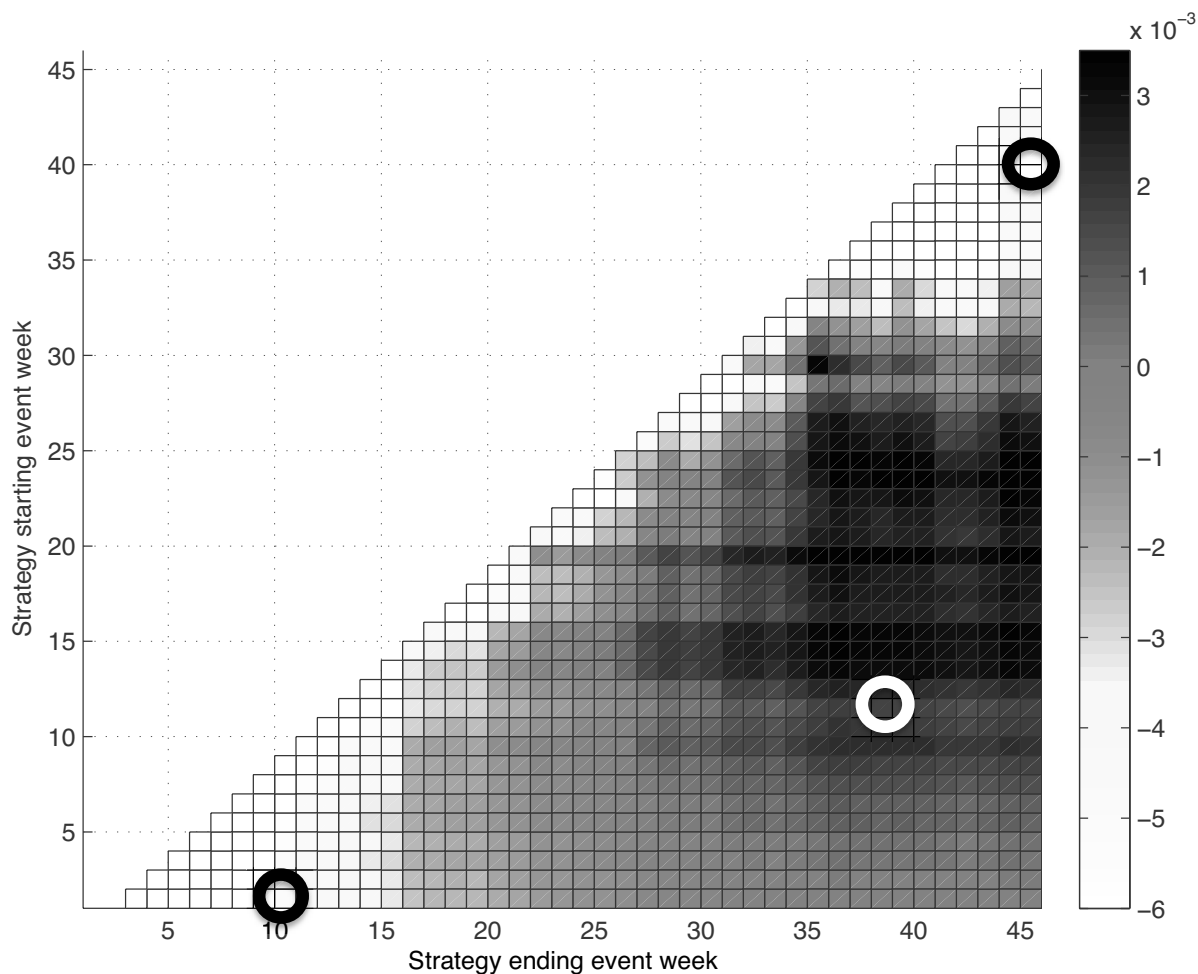
Figure A6. Robustness: Strategy alphas

The figure reports the Fama-French alphas of all possible event window trading strategies in excess of the alpha of the Buy-and-Hold strategy. On the y-axis, we report the first week of the strategy's event window. On the x-axis, we report the last week of the strategy's event window. The color at point (x, y) indicates the relative alpha for a strategy that invests in all deals active between event weeks y to x. The three circles correspond to the Low Hazard 1, High Hazard, and Low Hazard 2 strategies.
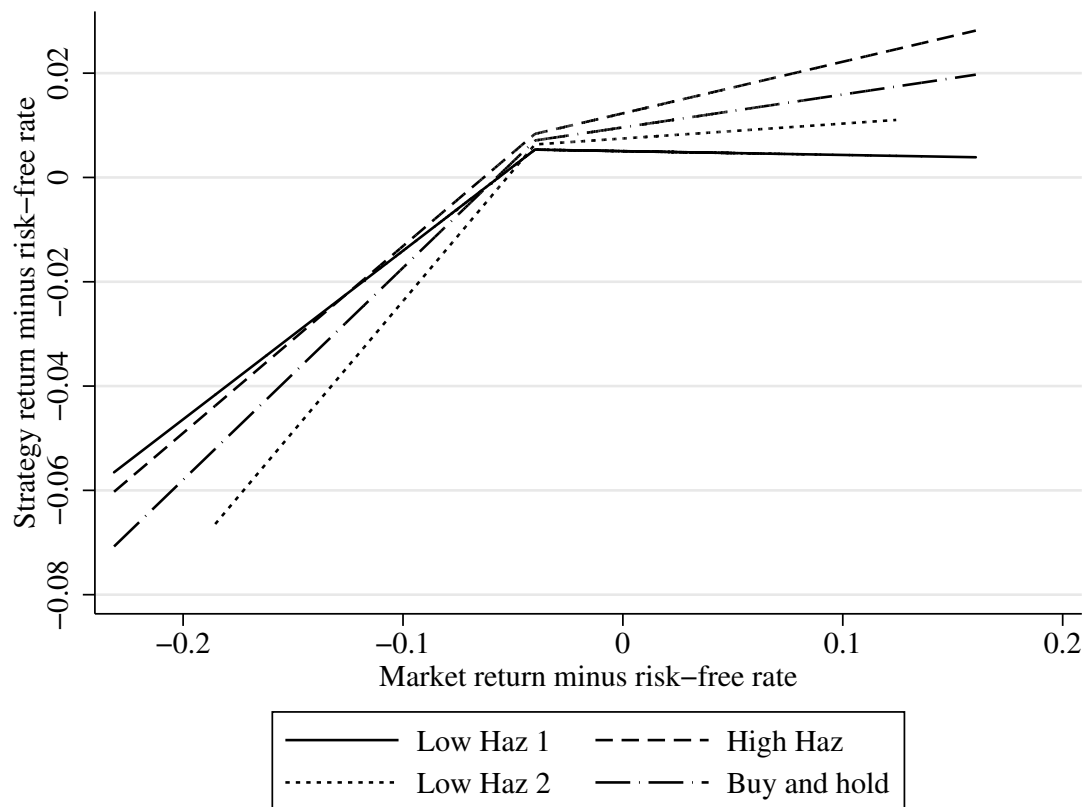
70

Figure A7. Downside beta

This figure explores event-time variation in downside beta. We estimate piecewise linear functions of the expected returns of the four trading strategies (Low Hazard 1, Low Hazard 2 and High Hazard, and Buy-and-Hold) against the market return. The specification is described in the Appendix.

71

Dep var: **Weekly return**	(1) Baseline	(2) Heterogeneous hazards	(3) Heterogenous hazards + dropdead dates
Weekly hazard (SD)	0.0659 ***	0.0660 ***	0.0719 ***
	(0.0147)	(0.0147)	(0.0151)
Calendar year x month FE	y	y	y
SD clustered by deal & (calendar year x month)	y	y	y
Obs.	90,044	90,044	90,038
R^2	0.0121	0.0125	0.0125

1 Standardized with only cash and equity heterogeneity

2 Standardized with 8 types of heterogeneity

3 Standardized with 9 types of heterogeneity (including dropdead)

Table A1: Heterogeneity

This table reports regressions of weekly returns on hazard rates as in Table 2, after accounting for additional hazard rate heterogeneity. We split deals into eight (including dropdead) deal characteristics as described in the Appendix, and for each group, we estimate a hazard rate curve. For deals for which we have a drop-dead date, we also compute a hazard curve by scaling time relative to the drop-dead date, as described in the Appendix. To make the results comparable across columns, we standardize hazard curves within each group, and also present for comparison in Column (1) the regression in which we do not account for heterogeneity across deals. Column (2) reports the regression of returns on hazards when the hazard rate for each deal is obtained from the corresponding group based on deal characteristics. Column (3) reports the same regression after replacing the characteristics-based hazard curve with the one obtained using the drop-dead date for all deals for which drop-dead date data are available. Note the sample size drops by six observations in Column 3 because these deals survive until more than two times the drop-dead date. We lack power to estimate the hazard curve so far after the drop-dead date. In each specification, we include calendar time fixed effects. All standard errors are allowed to be double clustered by calendar year-month and by merger.

	No trading costs alpha		Trading costs alpha (method 1)		Trading costs alpha (method 2)	
Low hazard 1	0.0022 *	(0.0011)	-0.0035 ***	(0.0010)	-0.0077 ***	(0.0009)
High hazard	0.0104 ***	(0.0018)	0.0044 ***	(0.0013)	-0.0009	(0.0012)
Low hazard 2	0.0022	(0.0032)	-0.0002	(0.0008)	-0.0011 *	(0.0006)
Buy and hold	0.0071 ***	(0.0011)	0.0025 **	(0.0010)	-0.0011	(0.0009)
p-value: High > low	0.0003		0.0000		0.0022	
Obs.	1,916		1,916		1,916	
R^2	0.0386		0.1221		0.1394	

Table A2. Accounting for transaction costs

This left column of the table reports the returns of strategies that invest equally in all available deals from 1970 to the present, ignoring transaction costs and limits to portfolio weights. If no deals are active during the relevant event window for each strategy, the return is equal to the risk-free rate. In the right columns, we follow Mitchell and Pulvino (2001) and Breen et al. (2002) in estimating the alphas for feasible trading strategies after accounting for transaction costs. Each strategy starts with \$1M of funds in 1970. Every month, each fund invests equally in all available deals in their respective event windows subject to the limitation that at most 10% of the capital can be invested in any particular deal. For each trade, we account for direct transaction costs of \$0.10 per share prior to 1980, \$0.05 per share from 1980 to 1989, \$0.04 per share from 1990 to 1998, and \$0.03 from 1999 onwards. In addition, we account for indirect transaction costs using two methods. Method 1 computes indirect transaction costs using the estimates of price pressure from Breen, Hodrik, and Korajczyk (2002). In addition, we impose the additional restriction that the total trading in any deal cannot produce price pressure of more than 5% of the price. Method 2 computes indirect transaction costs using the bid-ask spread. If data on bid-ask spread is unavailable from CRSP, we supplement with the estimated bid-ask spread from Corwin and Schultz (2012).