Cash Mergers and the Volatility Smile

C. Alan Bester, Victor H. Martinez, and Ioanid Roşu*

April 13, 2019

Abstract

In an empirical study of cash mergers, we find that the equity options on target firms display a pronounced smile pattern in their implied volatilities which gets more pronounced when the merger success probability gets higher. We propose an arbitrage-free model to analyze option prices for firms undergoing a cash merger attempt. Our formula matches well the observed merger volatility smile. Furthermore, as predicted by the model, we show empirically that the merger volatility smile has a kink at the offer price, and that the magnitude of the kink is proportional to the merger success probability.

JEL Classification: G13, G34, C58.

Keywords: Mergers and acquisitions, Black–Scholes formula, success probability, fallback price, Markov Chain Monte Carlo.

*Bester is with the University of Western Ontario; Martinez is with State Street; and Roşu is with HEC Paris. We thank two anonymous referees, Fabio Trojani (the editor), Malcolm Baker, John Cochrane, George Constantinides, Doug Diamond, Pierre Collin-Dufresne, Charlotte Hansen, Ulrich Hege, Steve Kaplan, Jun Pan, Monika Piazzesi, Ľuboš Pástor, Ajay Subramanian, Pietro Veronesi; finance seminar participants at Chicago Booth, Courant Institute, CUNY Baruch, Princeton, Toronto, HEC Lausanne; and conference participants at the American Finance Association 2010 meetings, and the ESSEC-HEC-INSEAD-PSE 2011 meetings in Paris, for valuable comments. We acknowledge financial support from the Investissements d'Avenir Labex (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

1 Introduction

One of the most common violations of the Black and Scholes (1973) formula is the volatility smile (or smirk, or sneer), which is a pattern whereby at-the-money equity options have lower implied volatilities than either in-the-money or out-of-the-money options.[1] In the case of European options on S&P 500 futures, Rubinstein (1994) and Jackwerth and Rubinstein (1996) show that the volatility smile has become economically significant only after the 1987 market crash, period for which they also record strong violations of the assumption of a log-normal distribution for the underlying equity prices. The volatility smile has been documented in many other option markets, e.g., for options on currencies or fixed income securities, and various explanations have been proposed for the volatility smile, including stochastic volatility or jumps in the underlying prices, fear of crashes, etc.[2]

If changes in the probability of a stock index crash can affect the volatility smile, then one would expect certain extreme events in the life of a firm to also affect the volatility smile. A natural candidate for an extreme event is the firm being the target of a merger attempt.[3] Compared with studying market crashes, mergers have the advantage that many of the variables involved in a merger deal are observable, e.g., the offer price, or the effective date (when the merger is expected to be completed). To simplify our analysis, we consider only mergers for which the offer is made entirely in cash.[4]

FIGURE 1 ABOUT HERE

We then analyze all cash mergers announced between January 1996 and December 2014, and study the effect of the merger on the implied volatility smile for the target firm. Indeed, we find that the options on the target company undergoing a cash merger display

[1]Black and Scholes (1973) assume that the underlying equity price follows a log-normal distribution with constant volatility, which means that the implied volatilities should be the same, irrespective of the strike price of the option.

[2]See Hull and White (1987), Heston (1993), Bates (1996), Jackwerth and Rubinstein (1996), Bakshi, Cao, and Chen (1997), Duffie, Pan, and Singleton (2000), Jackwerth (2000), Kou (2002), Ziegler (2007), and Yan (2011).

[3]Black (1989) points out that the Black and Scholes (1973) formula is unlikely to work when the company is the subject of a merger attempt.

[4]When studying options on the target company of a cash merger, as a first approximation we can ignore the stock price of the acquirer, which leads to a simpler model.

a pronounced volatility smile, that we call the *merger volatility smile*. Furthermore, we show that the shape of the merger volatility smile depends crucially on the probability of success of the cash merger: the smile is more pronounced when the probability of success is higher. To illustrate this, Figure 1 shows the difference between the median merger volatility smile for cash mergers that eventually succeeded (which on average should have a higher success probability) and the median merger volatility smile for cash mergers that eventually failed (which on average should have a lower success probability). In addition, for the successful deals we notice a pronounced *kink* in the volatility smile that occurs when the option strike price is close to the target offer price. These findings are qualitatively the same whether as call option prices we use the end-of-day bid price, ask price, or bid-ask mid-quote.

To analyze further the merger volatility smile and understand the source of the kink in Figure 1, we propose a theoretical no-arbitrage model that prices options on the equity of a firm subject to a merger offer. The model is perhaps the simplest extension of the Black and Scholes (1973) model adapted to cash mergers. The model predicts that option prices on the equity of the target firm should exhibit a merger volatility smile, with a kink at the offer price. The size of the kink, i.e., the difference between the tangent slopes at the offer price, should be proportional to the success probability of the merger. An empirical test for cash mergers with options traded on the target company strongly supports the model.[5]

FIGURE 2 ABOUT HERE

Our model also produces stochastic volatility for the underlying equity price. But, rather than assuming stochastic volatility as an exogenous process, we show that it arises endogenously as a function of the probability of success and the other variables in the structural model. Figure 2 shows the stark difference between the median Black and Scholes (1973) at-the-money implied volatility for cash mergers that eventually succeed versus the median implied volatility for cash mergers that eventually fail. In particular, when then merger is close to a successful completion, the implied volatility is low. This

[5]A placebo test shows that the volatility smile of the target company before before the merger announcement date is essentially flat, in stark contrast to Figure 1 (see Figure 6 in Appendix C).

is in line with intuition: when the merger is close to being successful, the equity price is close to the cash offer, thus has low volatility.

More specifically, in our theoretical model we consider a cash merger, in which a company A, the acquirer, makes a cash offer to a company B, the target. The deal is expected to complete (succeed or fail) by a deterministic date T_e, the effective date of the merger. In practice, the offer is usually made at a significant premium compared with B's pre-announcement stock price, about 35% in our sample. Therefore, the distribution of the stock price of B is not log-normal, but usually bi-modal: if the deal is successful, the price rises to the offer price, B_1; if the deal is unsuccessful, the price reverts to a *fallback* price, B_2, that we assume to have a log-normal distribution.[6] We also model the success probability of the deal as a stochastic process, similar to a log-normal process, but constrained to be in $[0, 1]$. As in the martingale approach to the Black–Scholes formula, instead of using the actual success probability, we focus on the *risk neutral* probability, q.[7] We provide formulas for both the target stock price, $B(t)$, as well as for call options on the target, $C^{K,T}(t)$ with strike price K and expiration date T. The formulas have a particularly simple form when the success probability q and the fallback price B_2 are uncorrelated: the stock price of B is a mixture of B_1 and B_2 with weights given by q and $1-q$, and similarly for the prices of European options on B with expiration date T past the merger effective date T_e.

Both the success probability q and the fallback price B_2 are latent (unobserved) variables, and thus, in the absence of options traded on B, one cannot use only the stock price of B to identify both q and B_2. Therefore, our estimation method also uses call options traded on B, and thus has two (or more, if we use more than one option) observed variables to identify two latent variables.[8]

We apply the option formula to the cash mergers announced between January 1996

[6] The fallback price reflects the value of the target firm B based on fundamentals, but also based on other potential merger offers. The fallback price therefore should not be thought as some kind of fundamental price of company B, but simply as the price of firm B if the *current* deal fails.

[7] In the absence of time discounting, the risk neutral probability $q(t)$ would be equal to the price at t of a digital option that offers one if the deal is successful and zero otherwise.

[8] Since the Black–Scholes formula is non-linear in the stock price, we need a statistical technique that deals with non-linear formulas and identifies both the values of the latent variables and the parameters that generate the processes. The method we use is the Markov Chain Monte Carlo (MCMC); see, e.g., Johannes and Polson (2003), or Jacquier, Johannes, and Polson (2007).

and December 2014, for which there are options traded on the target company. We test our model in three different ways. First, we compare the model-implied option prices to those coming from the Black–Scholes formula, and we investigate the volatility smile. Since our estimation method uses one option each day, we check whether the prices of the other options on that day (with different strike prices) line up according to our formula. Second, we explore whether the success probabilities uncovered by our approach predict the actual deal outcomes we observe in the data. Third, we explore the implications of our model for the volatility dynamics and risk premia associated with mergers.

In comparison with the Black–Scholes formula with constant volatility, our option formula does significantly better. Indeed, the average percentage error is 9.39% for our model compared to an error of 19.49% in the case of the Black–Scholes model. In both cases, the error is larger than the average percentage bid-ask spread for options in our sample, which is 27.46%. In the same cross-section of firms, the 95% quartile of the percentage error for our model is 27.69%, compared to 68.61% for the Black–Scholes model.

Our model predicts a kink in the volatility smile, whose magnitude, as can already be observed from Figure 1, increases with the success probability. For a more formal test, instead of looking at the implied volatility plot, we consider plotting the call option price against the strike price. Then the model predicts that the magnitude of the kink normalized by the time discount coefficient should be precisely equal to the success probability. A regression of the normalized kink on the estimated success probability supports the prediction does indeed produce an intercept that is close to zero and a slope that is close to one. Remarkably, in our estimation procedure we use only one option each day, yet we match well the whole cross section of options for that day, and in particular, as already mentioned, we also match the magnitude of the kink.

Our model further predicts that the return volatility of the target is stochastic and is proportional to one minus the success probability. This is illustrated in Figure 2, which shows that when the success probability is close to one, the implied volatility of call options on the target, as well as the return volatility of the underlying, is close to zero. A formal regression test confirms this prediction, as well as several additional predictions.

We show that the probabilities estimated using our formula predict the outcomes of merger deals in the data. In particular, this method does significantly better than the "naive" method widely used in the mergers and acquisitions literature (see for instance Brown and Raymond 1986), which estimates the success probability based on the distance between the current stock price and the offer price in comparison to the pre-announcement price.

Another implication of our model is that the merger risk premium may be estimated to be proportional to the drift coefficient in the diffusion process for the success probability. This is noisy at the individual deal level, but over the whole sample the average merger risk premium is significantly positive, at an 122.05% annual rate.[9] One explanation for this large number is that betting on merger success is a leveraged, option-like bet on the market index. Indeed, Mitchell and Pulvino (2001) find that the performance of mutual funds involved in merger arbitrage (also called "risk arbitrage") is equivalent to writing naked put options on a market index. Nevertheless, it is possible that the merger risk premium is too high to be justified by fundamentals. If this is the case, one should expect the merger risk premium to decrease over time, as the result of more investors taking bets on mergers. In agreement with this intuition, we find that over the last five years of our sample (January 2010 to December 2014) the average annual merger risk premium has decreased to 86.47%.

Our paper is, to our knowledge, the first to study option pricing on mergers by allowing the success probability to be stochastic. This has the advantage of being realistic. Indeed, many news stories before the resolution of a merger involve the success of the merger. A practical advantage is that by estimating the whole time series of success probabilities, we can estimate for instance the merger risk premium. Also, our model is well suited to study cash mergers, which are difficult to analyze with other models of option pricing. Subramanian (2004) proposes a jump model of option prices on *stock-for-stock* mergers. According to his model, initially the price of each company involved in a merger follows a process associated to the success state, but may

[9] Dukes, Frolich and Ma (1992) consider the 761 cash mergers between 1971 and 1985 and report returns to merger arbitrage of approximately 0.47% daily, or approximately 118% annualized. See also Jindra and Walkling (2004), who obtain similar numbers but also take into account transaction costs.

jump later at some Poisson rate to the process associated to the failure state.[10] This approach cannot be extended to cash mergers: when the deal is successful, the stock price of the target becomes equal to the cash offer, which is essentially constant; thus, the corresponding process has no volatility. In our model, the price of the target is volatile: this is due to both a stochastic success probability and a stochastic fallback price.

The literature on option pricing for companies involved in mergers is scarce, and, with the exception of Subramanian (2004), mostly on the empirical side: see Barone-Adesi, Brown, and Harlow (1994) and Samuelson and Rosenthal (1986).[11] The latter paper is close in spirit to ours. They start with an empirical formula similar to our theoretical result, although they do not distinguish between risk neutral and actual probabilities. Assuming that the success probability and fallback prices are constant (at least on some time-intervals), they develop an econometric method of estimating the success probability.[12] The conclusion is that market prices usually reflect well the uncertainties involved, and that the market's predictions of the success probability improve monotonically with time.

Our paper is also related to the literature on pricing derivative securities under credit risk. The similarity with our framework lies in that the processes related to the underlying default are modeled explicitly, and their estimation is central in pricing the credit risk securities. See, e.g., Duffie and Singleton (1997), Pan and Singleton (2008), Berndt et al. (2005). Similar ideas to ours, but involving earning announcements can be found in Dubinsky and Johannes (2005), who use options to extract information regarding earnings announcements.

[10] An implication of Subramanian (2004) is that the prices of the acquirer and the target companies are perfectly correlated, which is not realistic when the merger has a low success probability. Moreover, his model implies that the success probability of a merger decreases deterministically with time, even when the merger is likely to succeed. Samuelson and Rosenthal (1986) find empirically that the success probability usually increases over time.

[11] Several papers show that options can be useful for extracting information about mergers, although the variable of interest in many of these paper are the merger synergies; see Hietala, Kaplan, and Robinson (2003), Barraclough, Robinson, Smith, and Whaley (2013). Cao, Chen, and Griffin (2005) observe that option trading volume imbalances are informative prior to merger announcements, but not in general.

[12] They estimate the fallback price by fitting a regression on a sample of failed deals between 1976–1981. The regression is of the fallback price on the offer price and on the price before the deal is announced.

The paper is organized as follows. Section 2 describes the model, and derives our main pricing formulas, both for the stock prices and the option prices corresponding to the stocks involved in a cash merger. Section 3 presents the data and methodology, as well as the empirical tests of our model. Section 4 discusses the assumptions of the model and the robustness of our empirical results. Section 5 concludes. All proofs are in Appendix A.

2 Model

2.1 Setup

We use a continuous-time framework as in Duffie (2001, Part II). Let $W(t)$ be a 3-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. The market has a risk-free security with instantaneous rate r, and three risky securities: B_1, the *offer price*; B_2, the *fallback price*; and p_m, the *bet price*, where the bet is placed on the success of the merger. Consider $T_e > 0$, the *effective date* of the merger. Then the prices of the three risky securities are Itô processes that satisfy for all $t \in [0, T_e)$,

$$\begin{aligned} \mathrm{d}B_i(t) &= \mu_i\big(B_i(t), t\big)\mathrm{d}t + \sigma_i\big(B_i(t), t\big)\mathrm{d}W_i(t), \quad i = 1, 2, \\ \mathrm{d}p_m(t) &= \mu_3\big(p_m(t), t\big)\mathrm{d}t + \sigma_3\big(p_m(t), t\big)\mathrm{d}W_3(t), \end{aligned} \tag{1}$$

such that μ_i and σ_i satisfy regularity conditions as in Duffie (2001), and $B_i > 0$ and $p_m \in (0, 1)$ almost surely. With this specification, the processes B_1, B_2 and p_m are independent. The case when some of these processes are correlated is discussed in Section 4.4.

The methodology of this section works with general processes B_1, B_2, p_m. A particular example is given by the processes:

$$\begin{aligned} \frac{\mathrm{d}B_i(t)}{B_i(t)} &= \mu_i \mathrm{d}t + \sigma_i \mathrm{d}W_i(t), \quad i = 1, 2, 3, \\ p_m(t) &= \mathrm{e}^{-r(T_e - t)}\, q(t), \quad \text{with} \quad q(t) = \Phi\big(d_-^{B_3, T}\big), \end{aligned} \tag{2}$$

where μ_i and σ_i are constants, $\Phi(\cdot)$ is the standard normal cumulative density, and

$d_-^{B_3,T}$ is the corresponding Black–Scholes term for pricing options on B_3 with expiration $T \geq T_e$ and strike $K = 1$.[13] Note that in this example $p_m(t)$ is the price of a digital option expiring at T which pays one if $B_3(T) \geq 1$ and zero otherwise.

To model cash mergers, consider a company A (the acquirer) which announces at $t = 0$ that it wants to merge with a company B, the target. The acquisition is to be made with B_1 dollars in cash per share, a quantity which is not necessarily known at $t = 0$. At the end of the effective date T_e the uncertainty about the merger is resolved, and the following quantities also become known: the offer B_1 and the fallback B_2.[14] If the merger succeeds, B_1 is the amount that the target's shareholders receive per share. If the merger fails, B_2 is the market value that the target has as an independent firm. Both values B_1 and B_2 become known at T_e whether the merger is successful or not.

At each date t between 0 and T_e, we assume that the process $p_m(t)$ in (1) is the price of a contract that pays 1 if the merger succeeds or 0 if the merger fails.[15] Define the *risk neutral success probability*, or simply the *success probability*, to be the process

$$q(t) \;=\; p_m(t)\; \mathrm{e}^{r(T_e-t)}\,. \tag{3}$$

Note that, despite the fact that we use continuous processes, we do not require the success probability on the effective date to converge either to 0 or 1. We thus allow for the possibility of a last-minute surprise at the effective date.[16] We define the time T_e' to be the instant after T_e when the uncertainty is resolved. We extend q at T_e' as follows:

[13]When $K = 1$, equation (A6) in Appendix A implies that $d_-^{B_3,T} = \frac{\ln(B_3(t))+(r-\frac{1}{2}\sigma_3^2)(T-t)}{\sigma_3\sqrt{T-t}}$.

[14]In this section, the effective date is assumed fixed and known in advance by all market participants. Later, in Section 4.3 we analyze the case when T_e can change before the deal is completed.

[15]In practice, this type of contract exists in betting markets that wager on the outcome of political elections or sports games, but, to our knowledge, not on mergers (probably because it would create opportunities for illegal insider trading). Even if this contract is not actually traded, we consider p_m as the price of a contract "as if" it was traded. Note that we are in effect assuming that that there is a complete set of traded securities contingent on the uncertainty associated with cash merger deals. However, even if markets are incomplete, there is a set of equivalent martingale measures such that the discounted prices of traded securities are martingales. Then, the actual market can be viewed as picking one measure from this set, and our estimation approach essentially infers what this measure is using data on stock and option prices. We thank an anonymous referee for making this point.

[16]One possibility is to model q as a process with jumps, but to estimate the jump parameters we would need a longer time series than we typically have for merger deals. Hence, we only allow a jump at T_e. In practice, mergers are often decided before the effective date, and as a result in some cases the target stops trading before the effective date (in about 3% of the merger deals in our sample). In that case, in our empirical analysis we redefine the effective date as the last actual trading date. We discuss the issue of a random effective date in Section 4.3.

$q(T_e') = 1$ if the merger is successful, or $q(T_e') = 0$ if the merger fails. We also extend B_1 and B_2 at T_e' by continuity: $B_i(T_e') = B_i(T_e)$ for $i = 1, 2$.

Denote by Q the equivalent martingale measure associated to B_1 B_2 and p_m, such that these processes are Q-martingales after discounting at the risk-free rate r (see Duffie, 2001, Chapter 6).[17] Equation (3) then implies that the success probability q is a Q-martingale. To include the final resolution of uncertainty, we extend the probability space Ω on which Q is defined by including the binomial jump of q at T_e'. This defines a new equivalent martingale measure Q' and a new filtration $\mathcal{F}'$, such that B_1 and B_2, q are processes on $[0, T_e] \cup \{T_e'\}$, and q is a Q'-martingale.

2.2 Option Prices and the Volatility Smile

We now compute the stock price of the target company B, as well as the price of a European call option traded on B with strike price K and expiration after the effective date, i.e., $T \geq T_e$.[18] Recall that $B_i(t)$, $i = 1, 2$, is the market price of a security that pays $B_i(T_e)$ on the effective date, where $B_1(T_e)$ is the offer price and $B_2(T_e)$ is the fallback price. We also must specify what happens if the owner of an option receives cash before the expiration date, which is the case for instance if the merger is successful and the offer price B_1 is above the strike price K. Then, we assume that the owner invests the cash proceeds in a money market account at the risk-free r.

Proposition 1. *If B_1, B_2 and q are independent processes, then the target's stock price satisfies*

$$B(t) \;=\; q(t)B_1(t) + (1 - q(t))B_2(t), \quad t \in [0, T_e]. \tag{4}$$

The price of a European call option on B with strike price K and expiration $T \geq T_e$ is

$$C^{K,T}(t) \;=\; q(t)C_1^{K,T_e}(t) + (1 - q(t))C_2^{K,T}(t), \quad t \in [0, T_e], \tag{5}$$

where $C_i^{K,\tilde{T}}(t)$ is the price of the European call option with payoff $\big(B_i(\tilde{T}) - K\big)_+$ at $\tilde{T}$. The prices of American and European call options on B with the same strike price and

[17]The equivalent martingale measure is associated to fairly general processes, and it is not to be confused with the Black–Scholes risk-neutral probability. In the example in (2), the two notions happen to coincide, but in general this is not true.

[18]The complementary case $T < T_e$ is discussed in Section 4.3.

expiration are equal.

In particular, if the offer price B_1 is constant, Proposition 1 implies that

$$B(t) \;=\; q(t)B_1\,\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t). \tag{6}$$

$$C^{K,T}(t) \;=\; q(t)(B_1-K)_+\,\mathrm{e}^{-r(T_e-t)} + (1-q(t))C_2^{K,T}(t) \quad \text{if } T \geq T_e. \tag{7}$$

Thus, when the sources of merger uncertainty are uncorrelated, the target stock price has a particularly simple formula. The same is true for European option prices on the target if in addition the option expires *after* the effective date. If instead the option expires before the effective date, the formula is more involved (see Section 4.3). When the sources of merger uncertainty are correlated, even the formula for the stock price becomes more complicated (see Section 4.4).

We now study the *volatility smile*, which is the Black–Scholes implied volatility curve when the stock and option prices are computed according to our model. The volatility curve plots the Black–Scholes implied volatility of the call option price against the strike price K. If the Black–Scholes model were correct, the curve would be a horizontal line, indicating that the implied volatility should be a constant: the true volatility parameter. But in practice, as observed by Rubinstein (1994), the plot of implied volatility against K is convex, first going down until the strike price is approximately equal to the underlying stock price (the option is at-the-money), and then going up. This phenomenon is called the volatility "smile," or, if the curve is always decreasing, the volatility "smirk."

Corollary 1. *In the context of Proposition 1, suppose the offer price B_1 is constant and that B_2 follows an exponential Brownian motion. Then, a European call option with strike K and expiration $T \geq T_e$ exhibits a kink at $K = B_1$:*

$$\left(\frac{\partial C}{\partial K}\right)_{K\downarrow B_1} - \left(\frac{\partial C}{\partial K}\right)_{K\uparrow B_1} \;=\; \mathrm{e}^{-r(T_e-t)}\,q(t). \tag{8}$$

The implied volatility also exhibits a kink at $K = B_1$:

$$\left(\frac{\partial \sigma_{\mathrm{impl}}}{\partial K}\right)_{K\downarrow B_1} - \left(\frac{\partial \sigma_{\mathrm{impl}}}{\partial K}\right)_{K\uparrow B_1} \;=\; \frac{\mathrm{e}^{-r(T_e-t)}\,q(t)}{\nu(B,K,r,\tau,\sigma_{\mathrm{impl}})}, \tag{9}$$

where $\nu = \frac{\partial C}{\partial \sigma}$ *is the call option vega, and* σ_{impl} *is the Black–Scholes implied volatility. Moreover, for* $q(t)$ *sufficiently close to 1 and* T_e *sufficiently close to* T*, the slope* $\left(\frac{\partial \sigma_{\text{impl}}}{\partial K}\right)_{K \uparrow B_1}$ *is negative and the slope* $\left(\frac{\partial \sigma_{\text{impl}}}{\partial K}\right)_{K \downarrow B_1}$ *is positive.*

Corollary 1 shows that a volatility smile arises naturally for of options on cash mergers. More precisely, the volatility curve has a kink at $K = B_1$ (the offer price). The magnitude of the kink (the difference between the slope of the curve on the right and left of $K = B_1$) is equal to the time-discounted success probability, divided by the option vega. Moreover, the left slope is negative and the right slope is positive for most parameter values. Thus, we obtain the familiar shape for the volatility smile around $K = B_1$ (see also Figure 1).

Next, we show that the stock price computed in (6) exhibits stochastic volatility. We define the instantaneous volatility of a positive process $B(t)$ as the number $\sigma_{\scriptscriptstyle B}(t)$ that satisfies $\frac{\mathrm{d}B}{B}(t) = \mu_{\scriptscriptstyle B}(t)\mathrm{d}t + \sigma_{\scriptscriptstyle B}(t)\mathrm{d}W(t)$, where $W(t)$ is a standard Brownian motion. In Corollary 2, we compute the instantaneous volatility $\sigma_{\scriptscriptstyle B}(t)$ when the company B is the target of a cash merger. We also show that the instantaneous volatility and the Black–Scholes implied volatility of $B(t)$ approach zero when the success probability $q(t)$ approaches one.

Corollary 2. *Suppose* B_1 *is constant, and* $B_2(t)$ *and* $q(t)$ *satisfy* $\frac{\mathrm{d}B_2}{B_2} = \mu_2 \mathrm{d}t + \sigma_2 \mathrm{d}W_2$*, and* $\frac{\mathrm{d}q}{q(1-q)} = \mu_q \mathrm{d}t + \sigma_q \mathrm{d}W_q$*, where* $W_2(t)$ *and* $W_q(t)$ *are IID standard Brownian motions. Then, the instantaneous volatility of* B *satisfies*

$$\begin{aligned}
\sigma_{\scriptscriptstyle B}(t) &= \big(1 - q(t)\big) \left(\left(\frac{B(t) - B_2(t)}{B(t)}\, \sigma_q \right)^2 + \left(\frac{B_2(t)}{B(t)}\, \sigma_2 \right)^2 \right)^{1/2} \\
&= \big(1 - q(t)\big) \frac{B_2(t)}{B(t)} \left(\left(\frac{B_1\, \mathrm{e}^{-(T_e - t)} - B_2(t)}{B_2(t)} q(t) \right)^2 \sigma_q^2 \;+\; \sigma_2^2 \right)^{1/2} .
\end{aligned} \tag{10}$$

If $T_e = T$*, the Black–Scholes implied volatility* $\sigma_{\text{impl}}(t)$ *of a European call option with strike price* K *and expiration* T *approaches zero when* $q(t)$ *approaches one.*

A consequence of this corollary is that the instantaneous volatility vanishes when $q(t)$ approaches one. This is intuitive, because when the success of the merger is assured, the stock price of the target equals the cash offer, which is assumed constant for a cash merger. In Section 3, we test the implication of Corollary 2 that σ_B is proportional to $1 - q$ (see Table 8).

Corollary 2 explains why, as seen in Figure 2, the implied volatility of the target company B in a merger tends to be lower for merger deals that are eventually successful. Indeed, in these successful mergers the probability q is likely to be closer to one, and therefore the implied volatility is likely to be closer to zero.

Note that Corollary 2 implies that the volatility of the target company in a merger is naturally stochastic. This is not obtained simply by assumption as in other studies, but it is a result of a model which provides economic underpinnings for stochastic volatility.

3 Empirical Analysis

3.1 Sample of Cash Mergers

We build a sample of all the cash mergers that were announced between January 1st, 1996 and December 31st, 2014, and have options traded on the target company. Merger data, e.g., company names, offer prices and effective dates, are from Thomson Reuters SDC Platinum. Option data are form OptionMetrics, which reports daily closing prices starting from January 1996. We use OptionMetrics also for daily closing stock prices, and for consistency we compare them with data from CRSP.

Specifically, we start by running the following session in SDC platinum: we search for all domestic mergers (deal type 1, 2, 3, 4, 11) with the following characteristics: *M&A Type* equal to "Disclosed Dollar Value"; target publicly traded; consideration offered in cash (category 35, 1) with *Consideration Structure* equal to "CASHO" (cash only); *Percent of Shares Acquiror is Seeking to Own after Transaction* between 80 and 100; *Percent of Shares Held at Announcement* less than 20 (including empty); *Status* either "Completed" or "Withdrawn".[19] This initial search produces 3298 deals. After removing deals with no option information in OptionMetrics before Announcement, there are 973 deals left. We further remove the deals with 10 options or less traded during deal (12

[19] This condition is not too restrictive, since the deal announcement date must before December 31st, 2014, which should leave enough time for most mergers to be completed by October 27, 2016, the date of the SDC query.

deals) and deal period equal to two days or less (4 deals). After reading in detail the description of the mergers in this list, we further remove the deals for which (i) the offer is not pure cash (includes the acquirer's stock), and (ii) the target company is subject to another concurrent merger offer. There are now 843 deals left: 736 completed, and 107 withdrawn.

To create our final sample of mergers, we remove the deals for which we cannot run our estimation procedure: with deal duration of 4 days or less (2 deals); at least one day with no underlying stock prices traded on the target (1 deal); at least one day with no options quoted (14 deals); no options quoted with expiration date past the current effective date of the merger (13 deals); non-converging estimation procedure (2 deals).[20] The final sample contains 812 deals: 711 completed, and 101 withdrawn.

For the final sample, we analyze in more detail the deals in which the offer price changed during the deal. The SDC field *Price Per Share* records only the last offer price. To determine the other offer prices, we examine the following fields: *Consideration* which gives a short list of the offer prices; *Synopsis* which describes this list in more detail; *History File Event* and *History File Date* which together give a list of events, including dates when the offer was sweetened (offer price went up) or amended (offer price changed, either up or down).

We also analyze the deals for which the effective date changed during the deal. The effective date of a merger is defined by SDC as the date when the merger is completed, or when the acquiring company officially stops pursuing the bid. To determine the mergers whose effective date changed during the deal, we analyze the SDC field *Tender Offer Extensions* and then read the *History File Event* and *History File Date* to extract the dates at which the tender offer got extended (even if the number of tender offers is blank, we search for "effective" or "extended").[21]

[20] The non-convergence in these two cases comes from the scarcity of call option trades throughout the deal (the selection procedure is based on call option volume *before* the deal announcement). As a result, the call prices are almost constant (the quotes are stale) throughout the deal, and thus the estimation procedure fails to properly identify the state variables.

[21] Useful information is found in the fields *Tender Offer Original Expiration Date* (which is self-explanatory) or *Tender Offer Expiration Date* (which is the last effective date on record). The *History File Event* field usually says what the new Effective Date is, but sometimes we guess it based on the following rule: if the effective date changes at a later date T without any current effective date being previously announced, we consider T as the earlier effective date. This makes sense, because in principle a current effective date should always be filed with the SEC.

TABLE 1 ABOUT HERE

Table 1 shows summary statistics for our merger sample. For instance, the average deal duration (the number of trading days until the deal either succeeds or fails) is 67 days, while the maximum deal duration is 402 days. We define the *offer premium* as the percentage difference between the (cash) offer price, and the target company stock price on the day before the merger announcement. The average offer premium in our sample is 33.5%, and its standard deviation is 37.5%. The offer price changes rarely in our sample: the median number of changes is zero, and the mean is well below one (0.13). Even the 95% percentile is 1 (one change for the duration of the deal), with a maximum of 5. The effective date changes more frequently in our sample, but the median is still zero and the mean is still below one.

Table 1 includes statistics about how often options are traded on the target company. The fraction of trading days when there exists at least one option with positive trading volume is 65% for the average deal, indicating that options in our sample are illiquid. Table 2 gives additional evidence that for the illiquidity of options. For instance, the average percentage bid-ask spread for all call options is 27.5%.

TABLE 2 ABOUT HERE

For stock prices we use closing daily prices, and for option prices we use the closing bid-ask mid-quote, i.e., the average bettwee closing ask and bid prices. For deals that are successful, the options traded on the target company are converted into the right to receive: (i) the cash equivalent of the offer price minus the strike price, if the offered price is larger than the strike price; or (ii) zero, in the opposite case.

3.2 Methodology

For each of the cash merger deals in our sample, we record the following observed variables for the target company: (i) the effective date of the merger, T_e, measured as the number of trading days from the announcement; (ii) the risk-free interest rate, r; (iii) the cash offer price, B_1; (iv) the stock price of the target company, $B(t)$, on trading day t; (v) the price of the call options, $C^{K,T}(t)$, traded on the target company with a strike price of K and expiration date T. When the offer price or the effective date changes, we simply change the value of B_1 or T_e in the formula, thus essentially assuming that these changes are unanticipated.[22]

The latent variables in this model are the success probability $q(t)$ and the fallback price $B_2(t)$, which are assumed to evolve according to:[23]

$$\frac{\mathrm{d}q(t)}{q(t)\big(1-q(t)\big)} = \mu_1 \mathrm{d}t + \sigma_1 \mathrm{d}W_1(t), \tag{11}$$

$$\frac{\mathrm{d}B_2(t)}{B_2(t)} = \mu_2\, \mathrm{d}t + \sigma_2\, \mathrm{d}W_2(t), \tag{12}$$

with independent increments $\mathrm{d}W_1(t)$ and $\mathrm{d}W_2(t)$, and constant coefficients μ_i and σ_i. Alternative specifications that include the case of correlated $\mathrm{d}W_1(t)$ and $\mathrm{d}W_2(t)$ are discussed in Section 4.

With this parametrization of the success probability, the drift μ_1 has a particularly useful interpretation in relation to the merger risk premium. Indeed, recall that for a price process that satisfies $\mathrm{d}S/S = \mu(S,t)\mathrm{d}t + \sigma(S,t)\mathrm{d}W(t)$ the instantaneous risk premium is given by $\mathsf{E}_t(\mathrm{d}S/S) - r\mathrm{d}t = (\mu(S,t) - r)\mathrm{d}t$. In the case of a merger, the merger risk premium is associated to the price $p_m(t) = q(t)\,\mathrm{e}^{-r(T_e - t)}$ of a digital option that pays 1 if the merger is successful and 0 otherwise. The instantaneous merger risk premium is then:[24]

$$\mathsf{E}_t\left(\frac{\mathrm{d}p_m}{p_m}\right) - r\mathrm{d}t = \mathsf{E}_t\left(\frac{\mathrm{d}q}{q}\right) = (1-q)\mu_1 \mathrm{d}t. \tag{13}$$

To use the formulas in Section 2, we consider at each date t only the call options with expiration date T larger than the current merger effective date T_e. Proposition 1 then implies that the stock price $B(t)$ and the call option price $C^{K,T}(t)$ satisfy, respectively,

[22] Table 1 shows that changes in B_1 and T_e are rare. Section 4 discusses these issues in more detail.

[23] The specification for q is similar to the Black–Scholes for the stock price: $\frac{\mathrm{d}S}{S} = \mu \mathrm{d}t + \sigma \mathrm{d}W(t)$, which ensures that satisfies $S > 0$. In our case we want $q \in (0,1)$, hence the introduction of the term $(1-q)$ in the denominator. Note that this specification is different from the example given in (2). In that example q is the price of a digital option on a log-normal stock. The new specification, however, works slightly better empirically, possibly because the price of a digital option at expiration converges to either to zero or one almost surely.

[24] Equation (13) implies that the merger risk premium is near zero when q is near one. This comes from the functional specification of q in (11), which implies that $\mathrm{d}q/q$ has almost zero volatility (and hence it is almost risk free) when q approaches one.

equations (6) and (7). In our empirical specification, however, we assume that these equation hold only approximately:

$$B(t) = q(t)B_1\,\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t) + \varepsilon_B(t), \tag{14}$$

$$C^{K,T}(t) = q(t)(B_1-K)_+\,\mathrm{e}^{-r(T_e-t)} + (1-q(t))C_{\mathrm{BS}}(B_2(t),K,T-t) + \varepsilon_C(t), \tag{15}$$

where $C_{\mathrm{BS}}(S,K,T-t)$ is the Black–Scholes formula (A6) with arguments r and σ_2 omitted. The errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$ are IID bivariate normal:

$$\begin{bmatrix} \varepsilon_B(t) \\ \varepsilon_C(t) \end{bmatrix} \sim \mathcal{N}(0,\Sigma_\varepsilon), \quad \text{where} \quad \Sigma_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}. \tag{16}$$

In the main empirical specification, on each day t we select only the option with maximum trading volume on that day.[25]

Equations (11), (12), (14), (15) and (16) define a state space model with observables $B(t)$ and $C(t)$, latent (state) variables $q(t)$ and $B_2(t)$, and model parameters μ_1, σ_1, μ_2, σ_2, $\sigma_{\varepsilon,B}$ and $\sigma_{\varepsilon,C}$. We adopt a Bayesian approach and conduct inference by sampling from the joint posterior density of state variables and model parameters given the observables.

Specifically, we use a Markov Chain Monte Carlo (MCMC) method based on a state space representation of our model. In this framework, the state equations (11) and (12) specify the dynamics of latent variables, while the pricing equations (14) and (15) specify the relationship between the latent variables and the observables. The addition of errors with distribution described by (16) in the pricing equations is standard practice in state space modeling; this also allows us to easily extend the estimation procedure to multiple options and missing data. This approach is one of several (Bayesian or frequentist) suitable for this problem and is not new to our paper. For a discussion, see, e.g., Johannes and Polson (2003) or Koop (2003). The resulting estimation procedure is described in detail in Appendix B.[26] All priors used in our estimation are flat.

[25] If all option trading volumes are zero on that day, use the option with the strike price K closest to the strike price for the most currently traded option with maximum volume. Finally, if none of these criteria are met, we simply choose the at-the-money option on that day.

[26] As noted by Johannes and Polson (2003), equations of the type (14) or (15) are a non-linear filter. The problem is that it is difficult to do the estimation using the actual filter. Instead MCMC is a much cleaner estimation technique, but it does smoothing, because it uses all the data at once.

TABLE 3 ABOUT HERE

3.3 Empirical Results

As described in the data section, our sample contains 812 cash mergers during 1996–2014. Table 3 shows information on the ten largest merger deals sorted on the offer value, five that succeeded and five that failed.[27] Figure 3 shows the estimated success probability q for these ten deals. We see that large q corresponds to merger success (the five deals in the left column), while small q corresponds to merger failure (the five deals in the right column).

FIGURE 3 ABOUT HERE

Before we discuss in more detail the success probability estimates, we analyze how well our pricing formulas (14) and (15) work. Recall that the pricing formulas for the stock price $B(t)$ and option price $C(t)$ hold with errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$, respectively. The fitted values are our estimates for the stock price $\hat{B}(t)$ and the option price $\hat{C}(t)$. Table 4 shows summary statistics for the average stock pricing error $\frac{1}{T_e}\sum_{t=1}^{T_e}\left|\frac{\hat{B}(t)-B(t)}{B(t)}\right|$ over the duration of the deal. The stock pricing errors are in general very small, with a median error of 3.2 basis points.

TABLE 4 ABOUT HERE

We illustrate graphically the performance of the option pricing model, by selecting the deal with the largest offer premium (75.44%) among the ten mergers in Table 3.[28] The target company of this deal is AT&T Wireless, with ticker AWE. Figure 4 illustrates for the company AWE how our model fits the call option prices, including the kink in the implied volatility curve. This is remarkable, as our estimation method only uses one option per day, yet the model is capable of accurately predicting the whole cross section of call option prices for each day.

[27] The offer value is defined as the offer price multipled by the target's number of shares outstanding.

[28] The offer premium is defined as the ratio $\frac{B_1-B(0)}{B(0)}$, where B_1 is the offer price and $B(0)$ is the target's stock price one trading day before the merger announcement date. For two target companies in our sample, the price one day before the merger announcement is missing, and therefore the offer premium is also missing.

FIGURE 4 ABOUT HERE

Table 5 compares the percentage errors in the call option prices computed according to two models. The first model is the one described in this paper, denoted "BMR" for short. The second model, denoted by "BS," is the Black and Scholes (1973), with the volatility parameter $\sigma = \bar{\sigma}^i_{ATM}$ computed as the average implied volatility for the at-the-money (ATM) call options over the duration of the deal.

TABLE 5 ABOUT HERE

The error is computed by restricting the sample of call options according to the moneyness of the option, i.e., the ratio of the strike price K to the underlying stock price $B(t)$. We consider the following moneyness categories: (i) all call options; (ii) deep-in-the-money (Deep-ITM) calls, with $K/B < 0.9$; (iii) in-the-money (ITM) calls, with $K/B \in [0.9, 0.95)$; (iv) near-in-the-money (Near-ITM) calls, with $K/B \in [0.95, 1]$; (v) near-out-the-money (Near-OTM) calls, with $K/B \in [1, 1.05)$; (vi) out-of-the-money (OTM) calls, with $K/B \in [1.05, 1.1]$; and (vii) deep-out-of-the-money (Deep-OTM) calls if $K/B > 1.1$.[29]

As we are interested in comparing the volatility smile for the two models, we ensure that each day the moneyness categories have sufficiently many options. Thus, suppose there are $N_{m,t}$ options in the moneyness category m that are quoted on day t. Then, to include these options in our calculations, we require $N_{m,t} \geq 8$ when m is the first category (all calls), and $N_{m,t} \geq 2$ for all the other categories (calls of a particular moneyness). Otherwise, the data corresponding to these options on day t is considered as missing.[30]

To understand how the relative pricing errors are computed in the table, consider for instance the results corresponding to the third moneyness category (ITM calls). By looking at the number of observations, we see that there are only 58 target stocks for

[29] The moneyness intervals are chosen following Bakshi, Cao, and Chen (1997), except that we use a larger step (0.05) than their step (0.03). The reason is that they study S&P 500 index options, which are much more liquid than the options on the individual stocks in our sample.

[30] As we see in Table 5, imposing $N \geq 2$, e.g., for ITM calls reduces to 58 (out of 812) the number of firms for which there is at least one day t with two or more quoted ITM calls. The higher daily threshold $N \geq 6$ for all calls is chosen so that there are enough firms (374) satisfying this restriction. Choosing a different threshold produces similar results.

which there is at least one day when two or more ITM calls are quoted. Then, for a target stock j and for a call option on j with (bid-ask mid-quote) price $C(t)$ quoted on day t on a stock with (closing) price $B(t)$ and with strike K, we compute the percentage pricing error by $\left|\frac{C_M(t)-C(t)}{C(t)}\right|$, where $C_M(t)$ is the model-implied option price, where the model M can be either the theoretical model in this paper, denoted by "BMR;" or the Black–Scholes model, denoted by "BS" for which the volatility parameter is the average at-the-money implied volatility. We then compute the average error e_j over this particular category of options on stock j. The table then shows various statistics of e_j over the cross section of 58 target stocks with non-missing ITM option data.

Overall, our model does significantly better than the Black–Scholes model, for the whole sample, as well as for the various subsamples of options. The average pricing error for all call options is 9.39% for the BMR model, and 19.49% for the BS model. The average pricing error for ITM call options is 6.97% for the BMR model, and 9.26% for the BS model. The average pricing error for ITM call options is 66.90% for the BMR model, and 84.17% for the BS model.

Rather than comparing the pricing errors using different models, we ask a different question: are the pricing errors in our model small when compared with the observed bid-ask spread? Indeed, the bid-ask spread is a well-known measure of uncertainty regarding the price of a financial security, and therefore errors that are smaller than the bid-ask spread can be considered "small."

TABLE 6 ABOUT HERE

Table 6 shows summary statistics for the ratio of the absolute pricing error of a call option, relative to the observed bid-ask spread, i.e., $\left|\frac{C_{\text{BMR}}(t)-C(t)}{s(t)}\right|$, where $C_{\text{BMR}}(t)$ is the model-implied option price, $C(t)$ is the call option bid-ask mid-quote, and $s(t)$ is its quoted bid-ask spread. According to Table 6, the 75% percentile of this relative pricing error is less than one for all moneyness categories. Moreover, we compute that the 97% percentile of the relative pricing error for all call options is 0.85. Thus, for at least 97% of all options in our sample, the pricing errors are within the bid-ask spread. In this sense, our model performs fairly well.

We next test several empirical implications of our model. Corollary 1 shows that the theoretical call price of a target company in a cash merger has a kink at $K = B_1$, i.e., when the strike price equals the offer price. Recall that a kink indicates a difference in slope above and below $K = B_1$. Moreover, Corollary 1 shows that there is a kink in the volatility smile, i.e., in a plot of the implied volatility against the strike price.[31] Formally, the magnitude of the call price kink (the difference in slope above and below the strike), normalized by the time discount coefficient, is equal to the success probability q. Thus, if we perform a OLS regression of the normalized kink on the estimated success probability, we should find that the slope coefficient is equal to one. In practice, as the independent variable q is estimated with error, the standard errors-in-variables (EIV) problem implies that the regression coefficient is usually less than one.

TABLE 7 ABOUT HERE

Table 7 displays results of panel regressions of the estimated kink (the difference in slopes above and below the strike price closest to the offer price) on the estimated success probability using our methodology which uses only one option each day. We also regress a modified kink $\tilde{C}^{\text{kink}}$ on a modified success probability $\tilde{q}(t)$, where the modification involves truncating the values of the kink to be in $(0, 1)$ and inverting them via the standard normal cumulative density function to be a variable on the whole real line.[32] If we want to use time fixed-effects, one problem is that there is usually no overlap between the time periods when two different merger deals are ongoing. Nevertheless, there might be still time trends that can occur during the lifetime of the deal, which means that using fixed-effects in the regression would be useful. Therefore, for each deal, we divide the period between the merger announcement date (denoted by $t = 1$) and the effective date (denoted by $t = T_e$) into ten periods. Namely, for each date $t = 1, \ldots, T_e$, we define its corresponding period by $\lceil \frac{10t}{T_e} \rceil$, where $\lceil x \rceil$ denotes the (integer) ceiling of the real number x. We then use time fixed-effects at the period level, and cluster standard errors by firm (clustering by firm and time produces very similar results).

As the model predicts, Table 7 confirms that there is a positive relation between the call price kink and the success probability. The coefficient on the success probability

[31] Figure 4 shows this kink for the target company with ticker "AWE" during a cash merger.

[32] We perform this modification so that their distribution is closer to being normal.

is between zero and one, and significant in all specifications. The constant coefficient is statistically significant, indicating that there is a positive bias in estimating the call price kink. Indeed, as observed in Figure 4, when one estimates a linear kink with concave price curves above and below $K = B_1$, the estimated kink is larger than the actual kink.[33]

Corollary 2 shows that the return volatility of the target company B in a cash merger, $\sigma_{B}(t)$, is proportional to $1 - q(t)$, where $q(t)$ is the success probability. To test this result, we need to find a good estimate of the time-varying return volatility at each date t. This could be done using a GARCH model or one of its related models. However, to avoid parametric assumptions, we divide the time between the announcement date ($t = 1$) and the effective date ($t = T_e$) in five periods, and estimate a return standard deviation for each period.[34] To define the five periods, first let the interval length $\tau = \lfloor \frac{T_e}{5} \rfloor$, where $\lfloor x \rfloor$ denotes the (integer) floor of the real number x. Then, divide the integer interval $[1, T_e]$ into five intervals: $I_1 = [1, \tau]$, $I_2 = [\tau + 1, 2\tau]$, $I_3 = [2\tau + 1, 3\tau]$, $I_2 = [3\tau + 1, 4\tau]$, $I_5 = [4\tau + 1, T_e]$. For each target company i, and for each period (time interval) $k = 1, \ldots, 5$, we construct the return volatility $\sigma_{B,i,k}$ as the standard deviation of i's stock return over I_k, and the success probability $q_{i,k}$ as the average $q_i(t)$ over $t \in I_k$. Thus, if we perform an OLS regression of the return probability on $1 - q$, we should find that the slope coefficient is positive.

TABLE 8 ABOUT HERE

Table 8 (Panel A) shows the results of panel regressions of the return volatility on one minus the estimated success probability using our methodology which uses only one option each day. We use firm (i) and period (k) fixed-effects, and standard errors are clustered by firm (clustering by firm and time produces very similar results). As the model predicts, Table 8 confirms that there is a positive relation between the return volatility and one minus the success probability.

[33] To reduce this bias, one could estimate the kink using a spline, rather than linear, approximation. For this method, however, besides a call with $K \approx B_1$, one needs to have at least two calls with $K > B_1$ and two calls with $K < B_1$. This severely reduces the sample, as equity options are not very liquid.

[34] In Table 7, we use ten periods rather than five. This is because the variables there are already estimated at each date t. In Table 8, we use only five periods because we also need to estimate the return volatility over the whole period, and therefore we want the typical period to be large enough to produce good volatility estimates.

By visual inspection of equation (10), we obtain additional implications of Corollary 2: deals with higher q volatility (σ_q) and higher fallback volatility (σ_2) tend to have targets with higher return volatility (σ_B). Moreover, we predict that deals with higher offer premium, $\frac{B_1 - B(0)}{B(0)}$, tend to have higher σ_B: a higher offer premium suggests a higher term $\frac{B_1 \, \mathrm{e}^{-r(T_e - t)} - B_2(t)}{B_2(t)}$ in equation (10), hence a higher σ_B. This argument, however, depends on the identification of the price before the announcement, $B(0)$, with the fallback price, $B_2(t)$, which is not true even on average if, e.g., the fallback price jumps after being the subject of a merger offer. As some independent variables are constant across time, we run cross-sectional regressions, where we collapse the variable q along its mean, and we estimate σ_B as the standard deviation of the whole return time series of $B(t)$. As the terms in equation (10) appear multiplicatively, we use as dependent variable the natural logarithm of σ_B.

Table 8 (Panel B) confirms our predictions, except that in one specification the offer premium $\frac{B_1 - B(0)}{B(0)}$ has the wrong sign. A possible explanation, as discussed above, is the fact that the positive effect of the offer premium depends on how close the fallback price $B_2(t)$ is to $B(0)$. This identification is clearly sensitive in the fallback volatility parameter σ_2, especially since by running cross-sectional regressions, we average out σ_B across the time dimension. And indeed, as long as we omit σ_2 from the regressions, the coefficient on the offer premium is positive.

By comparing the regressions in Panel B and Panel A, we see that in Panel A a large increase in R^2 comes from introducing the firm fixed-effects, while in Panel B a similar increase in R^2 comes from controlling for the (constant) estimates of σ_q and σ_2 across firms. This suggests that the firm fixed-effects are driven largely by these volatility estimates.

We now test whether the success probability estimated by our model predicts the outcome of the corresponding merger deal in our sample. Figure 3 illustrates the results for the ten largest deals from Table 3, five of which succeeded, and five of which failed. Figure 3 displays the time series of the posterior median together with a 90% credibility interval (i.e., the 5^{th}, 50^{th}, and 95^{th} percentiles of the posterior) for the time series of the state variable $q(t)$. The estimates of $q(t)$ for the five deals that succeeded (on the left column) are overall much higher than for the five deals that failed (on the right column).[35]

TABLE 9 ABOUT HERE

Table 9 shows the results of probit regressions of the deal outcome (a dummy variable equal to one if successful or zero if failure) on two measures of the success probability. The first measure is the estimated $q(t)$ using our methodology which uses only one option each day.[36] The second measure is $q_{\text{naive}}(t)$, the "naive" method of Brown and Raymond (1986), which is used widely in the merger literature. Formally, $q_{\text{naive}}(t) = \frac{B(t)-B_0}{B_1-B_0}$ if $B_0 < B(t) < B_1$, where B_1 is the offer price, B_0 is the target stock price one day before the announcement, and $B(t)$ is the current target stock price. If $B(t) < B_0$ (or $> B_1$), q_{naive} is set equal to zero (one).[37] Note that the naive method essentially sets the fallback price equal to the pre-announcement price, while our method estimates a separate fallback price, $B_2(t)$.[38]

To address the fact that the time period for merger deals are heterogenous, we choose ten evenly spaced days during the period of the merger deal: for $n = 1, \ldots, 10$, let t_n be the closest integer strictly smaller than $n\frac{T_e}{10}$. Thus, we run probit regressions of the deal outcome on the success probability at t_n for $n = 1, 2, \ldots, 10$. As the independent variable does not vary with time, it is similar to running a cross-sectional regression, where the independent variable is averaged out over time. For each of the two success probability measures, we run two probit regressions: one for the whole time period, and one only for the second half (at t_n for $n = 6, 7, \ldots, 10$). We expect that both success probability measures are better at predicting the deal outcome if we restrict to the second half-time period.

Table 9 confirms that our measure q is better than q_{naive} at predicting the deal outcome: the pseudo-R^2 increases from 20.01% for the naive measure to 32.33% for our

[35] This is the only place we use all the available options to estimate the state variables. If instead we select one option each day (the call with the maximum trading volume on that day), then the results still hold but the error bars are wider, and the contrast between the two groups is not as strong.

[36] Recall that $q(t)$ is the risk-neutral probability, and not the physical probability. However, this difference biases the results against us, therefore if q is a good predictor of merger success, the physical probability should be an even better predictor.

[37] If $B(t)$ is close to the offer price B_1, the naive probability is high. If $B(t)$ is close to the pre-announcement stock price $B_0(t)$, the naive probability is low.

[38] Our method continues to perform better than a modified naive method which replaces B_0 by aB_0, where a is any constant in $[0.5, 2]$.

measure. Our measure predicts the deal outcome better in the second half-time period: the pseudo-R^2 increases from 32.33% (for the whole time period) to 46.20% (for the second half). The same is true for the naive measure: the pseudo-R^2 increases from 20.01% (for the whole time period) to 29.92% (for the second half).

We next explore the possibility to estimate the merger risk premium using the drift coefficient in the diffusion process for the success probability (11). According to equation (13), the instantaneous merger risk premium equals $(1-q)\mu_1 \mathrm{d}t$. In practice, we take the merger risk premium over by averaging out $1-q$ over the life of the deal: $(\overline{1-q})\mu_1$. The individual estimates for $(\overline{1-q})\mu_1$ are very noisy, but over the whole sample the average merger risk premium is significantly positive, and the annual figure is 122.05%. This number is very large but is consistent with the existing literature.[39] One potential explanation for the large estimate is that merger arbitrage (which is also called "risk arbitrage") is a leveraged, option-like bet on the market index (this intuition is consistent with Mitchell and Pulvino 2001). Nevertheless, it is possible that the merger risk premium is too high to be justified by fundamentals. If this is the case, some investors who are aware of an excessive risk premium should have entered the merger arbitrage business, and as a result they should have brought down the premium. To check whether this is the case, we restrict our merger sample to the last five years, when the announcement date is between January 2010 and December 2014. For this subsample we see that indeed the estimated premium is significantly lower: 86.47%.

4 Discussion and Robustness

After considering various alternative specifications of our baseline model, our model appears to be essentially robust. One explanation for this robustness is that only a simple model can address the large amount of noise present in option prices on individual companies. Table 2 shows that the average percentage bid-ask spread of call options written on the target companies in our cash merger sample is 27.46%. Thus, attempts

[39]See, e.g., Dukes, Frolich and Ma (1992), who report an average daily premium of 0.47% (approximately 118% annualized), over 761 cash mergers between 1971 and 1985. See also Jindra and Walkling (2004), who confirm the results for cash mergers, but also take into account transaction costs; and Mitchell and Pulvino (2001), who consider the problem over a longer period of time, and for all types of mergers.

to impose any additional structure on our baseline model typically result in increasing the noise in our estimates, but do not significantly change our main results.

4.1 Assumptions on Observed Variables

In our baseline model, we assume that the effective date T_e is known from the beginning. As a consequence, in our empirical tests we have taken T_e to be the date when the merger is either successful or fails. But, in reality it is often the case that the effective date of the merger subsequently changes from the initial date reported on the merger announcement day. (See Table 1 for some statistics regarding the number of effective date changes.) We address this concern in several ways. First, in contrast with our baseline specification in which we consider options with shortest expiration date after T_e, we change our specification to include call options with longer maturity. When we do this, the results stay essentially the same. Second, in Section 4.3 we show that by considering options with expiration date before T_e, the errors are not very large.

In our empirical methodology, we also assume that when the cash offer price B_1 changes, it gets simply replaced in the formulas with the new value. That is, if we denote by $\tilde{B}_1(t)$ the time series of the offer price, we use equation (6) to write the stock price as $B(t) = q(t)\tilde{B}_1(t)\, e^{-r(T_e-t)} + (1 - q(t))B_2(t)$. Alternatively, if that the market knows that B_1 is stochastic (but independent from the other processes), Proposition 1 implies that $B(t) = q(t)B_1(t) + (1 - q(t))B_2(t)$, where $B_1(t)$ is the market price at t of a contingent security that pays the (random) offer price on the effective date T_e. Thus, as long as $B_1(t)$ stays close to the discounted value of the current offer price ($e^{-r(T_e-t)}\, \tilde{B}_1(t)$), the model error is small. In the end, the validity of our approach depends on the perceived volatility of $\tilde{B}_1$ being close to zero. Table 1 shows that this is a plausible assumption, since the number of offer price changes is usually zero.[40]

Other observed variables are the target stock price $B(t)$, and the call option price $C(t)$. In an alternative specification, instead of the usual bid-ask mid-quote, we use the ask price. This does not change our results, except that the option pricing errors in

[40] In principle, we can modify our model in order to be able to estimate the volatility of B_1. We have not pursued this avenue, however, since B_1 does not change often enough to allow proper statistical identification.

Table 5 corresponding to our model become somewhat larger, although still significantly smaller than the errors corresponding to the Black–Scholes model.

4.2 Alternative Specifications for the Latent Variables

Alternative specifications for the success probability $q(t)$ show that our model is robust. If $X_1(t)$ is the Itô process $\mathrm{d}X_1(t) = \mu_1 \mathrm{d}t + \sigma_1 \mathrm{d}W_1(t)$ with constant coefficients μ_1 and σ_1, we consider the following functional forms for the success probability: (i) $q(t) = \frac{\mathrm{e}^{X_1(t)}}{1+\mathrm{e}^{X_1(t)}}$, (ii) $q(t) = \mathsf{E}_t^Q\big(\mathbf{1}_{X_1(T)>0}\big) = \Phi\Big(\frac{X_1(t)+(r-\frac{1}{2}\sigma_1^2)(T-t)}{\sigma_1\sqrt{T-t}}\Big)$, which is the price of a digital option on X_1, i.e., the bet that $X_1(T) > 0$. We further consider the specification (i) but when $X_1(t)$ a jump process of the form $\mathrm{d}X_1(t) = \mu_1 \mathrm{d}t + \sigma_1 \mathrm{d}W_1(t) + Z\mathrm{d}J$, where Z is the jump size and J is a Poisson process. Furthermore, our estimation method allows direct constraints on $X_1(t)$. For instance, if we require $X_1(t) \in [-1, 1]$, we consider the following functional forms: (iii) $q(t) = \frac{1}{2}\big(1 + \frac{X_1}{(1-a)+a|X_1|}\big)$, with $a \in (0, 1)$, (iv) $q(t) = \frac{1}{2} + X_1 - (\frac{a}{4} + \frac{1}{2})X_1|X_1| + \frac{a}{4}X_1^3$, with $a \in (0, 1)$.[41] Using these alternative specifications, our results do not change significantly, although in some cases such as (i) the algorithm converges much less often, and thus no estimates are obtained.[42]

One important issue is that the success probability and the fallback price might in reality be correlated. To address this issue, we consider the usual functional specifications for $q(t)$ and $B_2(t)$, except that the corresponding Itô increments $\mathrm{d}W_1$ and $\mathrm{d}W_2$ are no longer independent, but instead have a bivariate normal distribution, with instantaneous correlation $\rho \in [-1, 1]$. Then, as we see in Section 4.4, the formulas for the stock and option prices become more complicated and involve numerical integration. This slows down the algorithm considerably, by a factor of at least 100. Nevertheless, when we perform the estimation for the 10 companies in Table 3, we note that in all cases the parameter ρ is poorly identified, i.e., its estimated posterior likelihood is essentially flat. We interpret this result as indicating that *a priori* it is not even clear what sign the correlation ρ should have. For instance, under normal circumstances, small changes in

[41] E.g, the function $f(x) = \frac{1}{2}\big(1 + \frac{x}{(1-a)+a|x|}\big)$ satisfies $f(-1) = 0$, $f(0) = \frac{1}{2}$, $f(1) = 1$, $f'(-1) = f'(1) = \frac{1-a}{2}$. Thus, when $a \approx 1$, $f'(-1) = f'(1) \approx 0$, which has the practical consequence that, once $q(t)$ is estimated to be close to 0 or 1, it is more likely to stay near those values.

[42] This is because $q(t) = \frac{\mathrm{e}^{X_1(t)}}{1+\mathrm{e}^{X_1(t)}}$ is nearly constant for X_1 large. Thus, when X_1 by chance drifts towards large values, the estimated posterior density is essentially flat and cannot distinguish between different values. Then, with large probability the drift continues and the algorithm does not converge.

B_2 are not likely to affect q, while larger changes in B_2 in either direction may actually decrease q, thus pointing to a potentially non-linear relationship between q and B_2. We are skeptical of pursuing such extensions, since they are likely to lead to poor statistical identification.

Finally, because stock price errors are relatively small (see Table 4), we may assume that equation (14) holds without error, i.e., $\varepsilon_B(t) = 0$. In that case the success probability $q(t)$ can be expressed as a function of the fallback price $B_2(t)$ and substituted in (15). To simplify formulas, denote by $B_1(t) = B_1\, \mathrm{e}^{-r(T_e - t)}$, $C_1(t) = (B_1 - K)_+\, \mathrm{e}^{-r(T_e - t)}$, and $C_2(t) = C_{\mathrm{BS}}(B_2(t), K, r, T - t, \sigma_2)$, where C_{BS} is the Black–Scholes formula (A6). Then equations (14) and (15) imply

$$q(t) \;=\; \frac{B(t) - B_2(t)}{B_1(t) - B_2(t)}, \quad C(t) \;=\; C_2(t) + \frac{B(t) - B_2(t)}{B_1(t) - B_2(t)}\big(C_1(t) - C_2(t)\big) + \eta_C(t). \quad (17)$$

By focusing only on the second equation, we avoid choosing a functional specification for $q(t)$, such as (11). But we must satisfy the constraint $q(t) \in [0, 1]$, which means that $B_2(t)$ must be selected so that $\frac{B(t)-B_2(t)}{B_1(t)-B_2(t)} \in [0, 1]$. The results that use this specification are much noisier than those obtained under our baseline specification, while they do not significantly change our results.

4.3 Options Expiring Before the Effective Date

In this section we use the same setup as in Section 2.1, except that that we consider options that expire before the effective date. To simplify the presentation, we assume that (i) B_1 is constant, (ii) $B_2(t)$ is a log-normal process with constant coefficients, and (iii) $p_m(t)$ is the price of a digital option that pays one if $B_3(T_e)$ is above K_3 and zero otherwise, where $B_3(t)$ is a log-normal process with constant coefficients. Specifically, $B_2(t)$ and $B_3(t)$ satisfy

$$\mathrm{d}B_i(t) \;=\; \mu_i B_i(t)\mathrm{d}t + \sigma_i B_i(t)\mathrm{d}W_i(t), \quad i \;=\; 2, 3. \qquad (18)$$

Using the risk neutral Black–Scholes formalism we write for any $T \geq t$

$$B_i(T) \;=\; B_i(t)\exp\Big(\big(r - \tfrac{\sigma_i^2}{2}\big)(T-t) + \sigma_i\sqrt{T-t}\,\varepsilon_i\Big), \quad i \;=\; 2,3, \tag{19}$$

where ε_2 and ε_3 have independent standard normal distributions. Also, the risk neutral probability q satisfies for any $t < T < T_e$

$$q(T) \;=\; \Phi\left(\frac{\sqrt{T_e - t}\,\Phi^{-1}\big(q(t)\big) + \sqrt{T-t}\,\varepsilon_3}{\sqrt{T_e - T}}\right), \tag{20}$$

where ε_3 has a standard normal distribution.[43]

The next result computes the price of European calls that expire before T_e.

Proposition 2. *Suppose B_1 is constant, and B_2 and q are independent processes. Define*

$$B_1(T) \;=\; B_1\,\mathrm{e}^{-(T_e - T)}, \quad \bar{\varepsilon} \;=\; \begin{cases} \frac{\sqrt{T_e - T}}{\sqrt{T-t}}\,\Phi^{-1}\left(\frac{K}{B_1(T)}\right) - \frac{\sqrt{T_e-t}}{\sqrt{T-t}}\,\Phi^{-1}\big(q(t)\big), & \textit{if } K < B_1(T), \\ +\infty, & \textit{if } K \geq B_1(T) \end{cases} \tag{21}$$

Then the price of a European call option on B with strike K and expiration $T < T_e$ is

$$\begin{aligned} C(t) \;=\; & \int_{-\infty}^{\bar{\varepsilon}} \cdot\Big[\big(q(T)B_1(T) - K\big)\,\mathrm{e}^{-r(T-t)} \;+\; \big(1 - q(T)\big)\;B_2(t)\Big]\phi(\varepsilon_3)\;\mathrm{d}\varepsilon_3 \\ & + \int_{\bar{\varepsilon}}^{+\infty} \Big[\big(q(T)B_1(T) - K\big)\,\mathrm{e}^{-r(T-t)}\;\Phi\Big(d_-\big(B_2(t), \tfrac{K - q(T)B_1(T)}{1-q(T)}, T-t\big)\Big) \\ & \qquad + \big(1 - q(T)\big)\;B_2(t)\;\Phi\Big(d_+\big(B_2(t), \tfrac{K-q(T)B_1(T)}{1-q(T)}, T-t\big)\Big)\Big]\;\phi(\varepsilon_3)\;\mathrm{d}\varepsilon_3. \end{aligned} \tag{22}$$

where $q(T)$ is the function of ε_3 described in (20), and $d_\pm(S,K,T-t)$ are as in the Black–Scholes formula (A6) with arguments r and σ_2 omitted.

FIGURE 5 ABOUT HERE

In Figure 5 we compare the correct price in (22) to the price from equation (7), which is correct only if $T > T_e$. The purpose of this exercise is to see whether a stochastic effective date significantly affects the option price. Suppose at $t = 0$ the current effective

[43] See the proof of Proposition 2.

date of the merger is in $T_e = 80$ days, and we consider a European call option with strike $K = 95$ and expiration in $T = 90$ days.[44] Then, formula (7) implies that the "simple" call price is $C_{\text{simple}} = 6.0408$. If the effective date suddenly changes to $T_e = 100$, the option expires before T_e, and equation (22) implies that $C_{\text{correct}} = 5.4906$. If instead we use the formula (7), which is no longer correct, we have $C_{\text{simple}} = 6.0360$, which is slightly smaller than before (by 0.08%) because the time value of the cash offer, which is now expected later. The corresponding pricing error is $(C_{\text{simple}} - C_{\text{correct}})/C_{\text{correct}} = 9.9336\%$, which in Figure 5 corresponds to the point on the curve $K = 95$ with x-coordinate $T = 90$. This error, however, assumes that the effective date changes with certainty at $t = 0$. If instead the probability of an effective date change is smaller, e.g., 1% per day, then the actual pricing error is smaller than 9.9336%.[45]

4.4 Correlated Latent Variables

In this section we use the same setup as in Section 2.1, except that the success probability and the fallback price are correlated. To simplify the presentation, we assume that (i) B_1 is constant, (ii) $B_2(t)$ is a log-normal process with constant coefficients, (iii) $p_m(t)$ is the price of a digital option that pays one if $B_3(T_e)$ is above K_3 and zero otherwise, where $B_3(t)$ is a log-normal process with constant coefficients, and (iv) the contemporaneous increments of $\ln\big(B_2(t)\big)$ and $\ln\big(B_3(t)\big)$ have a bivariate normal distribution with instantaneous correlation $\rho \in (-1, 1)$. Specifically, $B_2(t)$ and $B_3(t)$ are defined as in (18), and therefore satisfy the same formulas as in (19). The difference is that now ε_2 and ε_3 are no longer independent but have a bivariate normal distribution with density

$$f_\rho(x, y) \;=\; \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right). \tag{23}$$

The next result extends the formulas (6) and (7) to the new context.

Proposition 3. *If B_1 is constant, and B_2 and q are correlated processes as above, then*

[44] The parameters values are: $B_1 = 100$, $B_2(0) = 90$, $q(0) = 0.5$, $r = 0.0.5$. The number of days is annualized by division with 252, which by convention is the number of trading days in one year.

[45] Table 1 shows that the effective date changes on average 0.74 times for a deal duration which on average is 66.6 trading days, hence the probability of an effective date change during any particular day is $0.74/66.6 \approx 1.11\%$.

the target's stock price satisfies for $t < T_e$

$$B(t) \;=\; q(t) B_1 \,\mathrm{e}^{-r(T_e - t)} \;+\; \left(1 - \Phi\Big(\Phi^{-1}\big(q(t)\big) + \rho\sigma_2\sqrt{T-t}\Big)\right) B_2(t). \qquad (24)$$

The price of a European call option on B with strike K and expiration $T \geq T_e$ is

$$C^{K,T}(t) \;=\; q(t)(B_1 - K)_+ \,\mathrm{e}^{-r(T_e - t)} \;+ C_{\mathrm{BS}}\big(B_2(t), T - t\big) \qquad (25)$$

$$- \int_{-\infty}^{+\infty} \Phi\left(\frac{\Phi^{-1}(q(t)) + \rho\varepsilon_2}{\sqrt{1-\rho^2}}\right) C_{\mathrm{BS}}\left(B_2(t) \exp\Big(\big(r - \tfrac{\sigma_2^2}{2}\big)(T_e - t) + \sigma_2\sqrt{T_e - t}\,\varepsilon_2\Big), T - T_e\right) \phi(\varepsilon_2)\,\mathrm{d}\varepsilon_2,$$

where $C_{\mathrm{BS}}(S, T-t)$ is the Black–Scholes formula (A6) with arguments K, r and σ_2 omitted.

5 Conclusion

In an empirical study of cash mergers since 1996, we find that equity options on firms that are the target of a cash merger display a pronounced pattern in their implied volatility smile. We call this pattern the *merger volatility smile*, and we find that it is more pronounced when the merger is close to being successful. To address this empirical regularity, we propose an arbitrage-free option pricing formula on companies that are subject to merger attempts.

Theoretically, our formula predicts a *kink* in the implied volatility curve at the cash offer price, or equivalently a kink in the call option price. Furthermore, the magnitude of the price kink equals essentially equals the (risk neutral) success probability. Empirically, we find essentially a one-to-one relationship between the magnitude of the price kink and the merger's success probability, which confirms our theoretical prediction.

Our option formula matches option prices significantly better than the standard Black–Scholes formula. In addition, we use the resulting formula to estimate several variables of interest in a cash merger, i.e., the success probability and the fallback price. The estimated success probability turns out to be a good predictor of the deal outcome, and it does better than the naive method which identifies the success probability solely based on how the current target stock price is situated between the offer price and the pre-merger announcement price. Besides the success probability itself, we also estimate its drift parameter, which turns out to be related to the merger risk premium. The average merger risk premium in our sample is about 122% annually, which is consistent with the cash mergers literature, although there is evidence the premium has decreased substantially in the past five years.

Our methodology is flexible enough to incorporate other existing information, such as prior beliefs about the variables and the parameters of the model. It can also be used to compute option pricing for "stock-for-stock" mergers or "mixed-stock-and-cash" mergers, where the offer is made using the acquirer's stock, or a combination of stock and cash. In that case, it can help estimate the synergies of the deal. The method can in principle be applied to other binary events, such as bankruptcy or earnings announcements (matching or missing analyst expectations), and is flexible enough to incorporate other existing information, such as prior beliefs about the variables and the parameters of the model.

Appendix

Appendix A. Proofs of Results

Proof of Proposition 1. Recall that T_e' is the instant after T_e when the merger uncertainty is resolved, and Q' is the extension of the equivalent martingale measure Q to $[0, T_e] \cup \{T_e'\}$. As $q(T_e')$ is either 1 or 0 depending on whether the merger is successful or not, the target's payoff at T_e' is

$$B(T_e) \;=\; q(T_e')B_1(T_e') + (1 - q(T_e'))B_2(T_e'), \tag{A1}$$

We apply Theorem 6J in Duffie (2001) for redundant securities. Markets are dynamically complete before T_e, because the uncertainty stems from the three Brownian motions involved in the definition of the securities B_1, B_2, and p_m. Moreover, at T_e, the stock price has a binary uncertainty that can be spanned only by the bond and p_m. Then, in the absence of arbitrage, any other security whose payoff depends on B_1, B_2, p_m is a discounted Q'-martingale. In particular, the price of the target company $B(t)$ is a discounted Q'-martingale, that is,

$$\begin{aligned} B(t) &= \mathrm{e}^{-r(T_e-t)}\ \mathsf{E}_t^{Q'}\Big(q(T_e')B_1(T_e') + (1-q(T_e'))B_2(T_e')\Big) \\ &= \mathsf{E}_t^{Q'}\big(q(T_e')\big)\,\mathrm{e}^{-r(T_e-t)}\,\mathsf{E}_t^{Q'}\big(B_1(T_e')\big) + \mathsf{E}_t^{Q'}\big(1-q(T_e')\big)\,\mathrm{e}^{-r(T_e-t)}\,\mathsf{E}_t^{Q'}\big(B_2(T_e')\big) \\ &= q(t)B_1(t) + \big(1-q(t)\big)B_2(t), \end{aligned} \tag{A2}$$

where for the second equation we use the independence of B_1, B_2 and q. This proves (4).

Consider a European call option on B with strike K and maturity $T \geq T_e$. Denote by $C^{K,T}(t)$ its price at $t \leq T_e$, and by $X_+ = \max\{X, 0\}$. Denote by $C_i^{K,T}(t)$ the price of a European call option on B_i with strike K and maturity T. Consider the payoff of the call at $t = T \geq T_e'$:

$$q(T_e')\big(B_1(T_e') - K\big)_+\,\mathrm{e}^{T-T_e} + (1-q(T_e'))\big(B_2(T) - K\big)_+. \tag{A3}$$

where by assumption all cash obtained after the effective date T_e is invested at the risk-free rate r. By a similar calculation as above, the call price at t is

$$\begin{aligned} C^{K,T}(t) &= \mathrm{e}^{-r(T-t)}\ \mathsf{E}_t^{Q'}\Big(q(T_e')\big(B_1(T_e') - K\big)_+\,\mathrm{e}^{r(T-T_e)} + (1-q(T_e'))\big(B_2(T) - K\big)_+\Big) \\ &= q(t)C_1^{K,T_e}(t) + \big(1-q(t)\big)C_2^{K,T}(t), \end{aligned} \tag{A4}$$

which proves (5).

We show that the prices of an American and European call option on B with the same strike K and expiration T are equal. For this, it is sufficient to prove that European call option price, $C^{K,T}(t)$, is always larger than the exercise value, $B(t) - K$. Equations (4) and (5) imply that

$$C^{K,T}(t) - \big(B(t) - K\big) = q(t)\Big(C_1^{K,T_e}(t) - \big(B_1(t) - K\big)\Big) + (1-q(t))\Big(C_2^{K,T}(t) - \big(B_2(t) - K\big)\Big). \tag{A5}$$

But C_1^{K,T_e} and $C_2^{K,T}$ are prices of European call options on B_1 and B_2, respectively, hence the right hand side in (A5) is positive. □

Proof of Corollary 1. From (7) the call price is $C^{K,T}(t) = q(t)\,\mathrm{e}^{-r(T_e-t)}(B_1-K)_+ + (1-q(t))C_2^{K,T}(t)$. Since B_2 follows a log-normal process, the option price $C_2^{K,T}(t)$ satisfies the Black–Scholes equation and is thus differentiable in K (see more details below). We now differentiate $C^{K,T}(t)$ with respect to K to the left and to the right of $K = B_1$: $\big(\frac{\partial C}{\partial K}\big)_{K\uparrow B_1} = -q(t)\,\mathrm{e}^{-r(T_e-t)} - \big(1-q(t)\big)\frac{\partial C_2^{K,T}(t)}{\partial K}$ and $\big(\frac{\partial C}{\partial K}\big)_{K\downarrow B_1} = -\big(1-q(t)\big)\frac{\partial C_2^{K,T}(t)}{\partial K}$. The kink is the difference $\big(\frac{\partial C}{\partial K}\big)_{K\downarrow B_1} - \big(\frac{\partial C}{\partial K}\big)_{K\uparrow B_1} = q(t)\,\mathrm{e}^{-r(T_e-t)}$. This proves (8) if we can show that $C_2^{K,T}(t)$ is differentiable in K. To see this, consider the particular case when $B_2(t)$ is an exponential Brownian motion with drift μ_2 and volatility σ_2. Then according to the Black–Scholes formula, we have

$$\begin{aligned} C_2^{K,T}(t) &= C_{\mathrm{BS}}\big(B_2(t),K,r,T-t,\sigma_2\big) = B_2(t)\Phi(d_+) - K\,\mathrm{e}^{-r(T-t)}\,\Phi(d_-), \\ d_\pm &= \frac{\ln(B_2(t)/K) + \big(r \pm \frac{1}{2}\sigma_2^2\big)(T-t)}{\sigma_2\sqrt{T-t}}. \end{aligned} \tag{A6}$$

One now verifies that $\frac{\partial C_2^{K,T}(t)}{\partial K} = -\,\mathrm{e}^{-r(T-t)}\,\Phi(d_-)$, which is continuous in K.

Let $\tau_e = T_e - t$ and $\tau = T - t$. By definition, the Black–Scholes implied volatility σ_{impl} of $C^{K,T}(t)$ satisfies $C^{K,T}(t) = C_{\mathrm{BS}}(B(t),K,r,\tau,\sigma_{\mathrm{impl}})$, where C_{BS} is the Black–Scholes formula (A6). By differentiating $C(t) = C_{\mathrm{BS}}(B(t),K,r,\tau,\sigma_{\mathrm{impl}})$ with respect to K, we get

$$\left(\frac{\partial\sigma_{\mathrm{impl}}}{\partial K}\right)_{K\downarrow B_1} = \frac{\big(\frac{\partial C}{\partial K}\big)_{K\downarrow B_1} + \mathrm{e}^{-r\tau}\,\Phi\big(d_-(B(t),K,r,\tau,\sigma_{\mathrm{impl}})\big)}{\nu\big(B(t),K,r,\tau,\sigma_{\mathrm{impl}}\big)}, \tag{A7}$$

and a similar formula for $\left(\frac{\partial\sigma_{\mathrm{impl}}}{\partial K}\right)_{K\uparrow B_1}$. Taking the difference of these two formulas, we obtain (9).

Next, set $q(t) = 1$ and $T_e = T$. Then $\tau_e = \tau$, and

$$\begin{aligned} \left(\frac{\partial\sigma_{\mathrm{impl}}}{\partial K}\right)_{K\downarrow B_1} &= \frac{\mathrm{e}^{-r\tau}}{\nu(B,K,r,\tau,\sigma_{\mathrm{impl}})}\Big(-1 + \Phi\big(d_-(B,K,r,\tau,\sigma_{\mathrm{impl}})\big)\Big) < 0, \\ \left(\frac{\partial\sigma_{\mathrm{impl}}}{\partial K}\right)_{K\uparrow B_1} &= \frac{\mathrm{e}^{-r\tau}}{\nu(B,K,r,\tau,\sigma_{\mathrm{impl}})}\Big(\Phi\big(d_-(B,K,r,\tau,\sigma_{\mathrm{impl}})\big)\Big) > 0. \end{aligned} \tag{A8}$$

By continuity, we obtain that the two inequalities above also hold when $q(t)$ is sufficiently close to 1 and T_e is sufficiently close to T. □

Proof of Corollary 2. To obtain equation (10), one applies Itô calculus to differentiate equation (6). To prove the result regarding the implied volatility, let $T_e = T$. In what follows, we denote $X \approx Y$ to mean $\lim_{q(t)\to 0}(X - Y) = 0$. When the volatility is zero, the Black–Scholes formula implies $C_{\text{BS}}\big(S, K, r, \tau, 0\big) = (S - K\,\mathrm{e}^{-r\tau})_+$. When $q(t) \approx 1$, equations (6) and (7) imply that $B(t) \approx B_1\,\mathrm{e}^{-r\tau}$ and $C^{K,T}(t) \approx (B_1 - K)_+\,\mathrm{e}^{-r\tau}$. Thus, we obtain $C^{K,T}(t) \approx C_{\text{BS}}\big(B(t), K, r, \tau, 0\big)$. But by definition, $C^{K,T}(t) = C_{\text{BS}}\big(B(t), K, r, \tau, \sigma_{\text{impl}}\big)$, hence $\sigma_{\text{impl}} \approx 0$. This finishes the proof. □

Proof of Proposition 2. We start with the general case when B_1 is stochastic. The formula (4) for the stock price $B(t)$ does not depend on the expiration date T. In particular, we have $B(T) = q(T)B_1(T) + \big(1 - q(T)\big)B_2(T)$. If we denote by

$$a \;=\; q(T)B_1(T) - K, \qquad b \;=\; 1 - q(T), \qquad k \;=\; \frac{K - q(T)B_1(T)}{1 - q(T)} \;=\; -\frac{a}{b}, \tag{A9}$$

we can write $B(T) - K = b\big(B_2(T) - k\big)$. Note that each of the variables a, b and k is independent from B_2. We now compute the price of a European call with strike K that expires at $T < T_e$. Its payoff at T is

$$C^{K,T}(T) \;=\; \big(B(T) - K\big)_+ \;=\; b\big(B_2(T) - k\big)_+. \tag{A10}$$

Note that $B(T) > K$ is equivalent to $B_2(T) > k$. There are two cases: $k \le 0$ and $k > 0$. Then, the price of the call at $t < T$ is

$$\begin{aligned}
C(t) \;&=\; \mathsf{E}_t^Q\Big[\mathrm{e}^{-r(T-t)}\,b\,\big(B_2(T) - k\big)_+\Big] \;=\; \mathsf{E}_t^Q\Big[b\;\mathsf{E}_t^Q\Big(\mathrm{e}^{-r(T-t)}\big(B_2(T) - k\big)_+\Big)\Big] \\
&=\; \mathsf{P}(k \le 0)\cdot \mathsf{E}_t^Q\Big(b\,B_2(t) \;-\; \mathrm{e}^{-r(T-t)}\,b\,k \;\Big|\; k \le 0\Big) \\
&\quad+\; \mathsf{P}(k > 0)\cdot \mathsf{E}_t^Q\Big(b\,C_{\text{BS}}\big(B_2(t), k, r, T-t, \sigma_2\big) \;\Big|\; k > 0\Big) \\
&=\; \mathsf{P}(k \le 0)\cdot \mathsf{E}_t^Q\Big(a\,\mathrm{e}^{-r(T-t)} \;+ b\,B_2(t) \;\Big|\; k \le 0\Big) \\
&\quad+\; \mathsf{P}(k > 0)\cdot \mathsf{E}_t^Q\;\Big(a\,\mathrm{e}^{-r(T-t)}\;\Phi\Big(d_-\big(B_2(t), k, r, T-t, \sigma_2\big)\Big) \\
&\qquad\qquad\qquad + b\,B_2(t)\,\Phi\Big(d_+\big(B_2(t), k, r, T-t, \sigma_2\big)\Big) \;\Big|\; k > 0\Big).
\end{aligned} \tag{A11}$$

This proves equation (22) if we show that (i) $q(T) = \Phi\Big(\frac{\sqrt{T_e - t}\,\Phi^{-1}(q(t)) + \sqrt{T-t}\,\varepsilon_3}{\sqrt{T_e - T}}\Big)$ is as in (20), and (ii) the condition $k > 0$ is equivalent to $\varepsilon_3 < \bar\varepsilon$ where $\bar\varepsilon$ is given by (21).

Recall that $q(t)$ is the risk neutral probability associated to a digital option that pays 1 at T_e if a log-normal process B_3 is above a level K_3, or pays 0 otherwise. Then, $B_3(T_e) = B_3(t)\exp\big((r-\frac{\sigma_3^2}{2})(T_e-t)+\sigma_3\sqrt{T_e-t}\,\varepsilon_3\big)$, where $\varepsilon_3 \sim \mathcal{N}(0,1)$ has a standard normal distribution. The price of the digital option at t is $p_m(t) = \mathrm{e}^{-r(T_e-t)}\,\Phi\Big(\frac{\ln\big(\frac{B_3(t)}{K_3}\big)+(r-\frac{\sigma_3^2}{2})(T_e-t)}{\sigma_3\sqrt{T_e-t}}\Big)$, which implies

$$
\begin{aligned}
q(T) &= \Phi\left(\frac{\ln\big(\frac{B_3(T)}{K_3}\big)+\big(r-\frac{\sigma_3^2}{2}\big)(T_e-T)}{\sigma_3\sqrt{T_e-T}}\right) \\
&= \Phi\left(\frac{\ln\big(\frac{B_3(t)}{K_3}\big)+\big(r-\frac{\sigma_3^2}{2}\big)(T_e-t)+\sigma_3\sqrt{T-t}\,\varepsilon_3}{\sigma_3\sqrt{T_e-T}}\right) \qquad \text{(A12)} \\
&= \Phi\left(\frac{\sqrt{T_e-t}\,\Phi^{-1}\big(q(t)\big)+\sqrt{T-t}\,\varepsilon_3}{\sqrt{T_e-T}}\right).
\end{aligned}
$$

This proves (20). The condition $k > 0$ is equivalent to $q(T) > \frac{K}{B_1(T)}$, which from equation (A12) is equivalent to $\varepsilon_3 < \bar{\varepsilon}$, where

$$
\bar{\varepsilon} = \begin{cases} +\infty, \text{ if } K \geq B_1(T) \quad \text{or} \\ \frac{\sqrt{T_e-T}}{\sqrt{T-t}}\,\Phi^{-1}\left(\frac{K}{B_1(T)}\right) - \frac{\sqrt{T_e-t}}{\sqrt{T-t}}\,\Phi^{-1}\big(q(t)\big), \text{ if } K < B_1(T). \end{cases} \qquad \text{(A13)}
$$

which is the same as in (21). □

Proof of Proposition 3. The only non-trivial part of proving (24) is the computation of

$$
\mathsf{E}_t^Q\big(q(T_e)B_2(T_e)\big) = \mathsf{E}_t^Q\big(\mathbf{1}_{B_3(T_e)\geq K_3}B_2(T_e)\big), \qquad \text{(A14)}
$$

where $\mathbf{1}_{B_3(T_e)\geq K_3}$ is the indicator function which is one if $B_3(T_e) \geq K_3$, or zero otherwise. From (19) we get $\ln\big(\frac{B_3(T_e)}{K_3}\big) = \ln\big(\frac{B_3(t)}{K_3}\big) + \big(r-\frac{\sigma_3^2}{2}\big)(T_e-t)+\sigma_3\sqrt{T_e-t}\,\varepsilon_3$, which implies that the condition $B_3(T_e) \geq K_3$ is equivalent to $\varepsilon_3 \geq -\Phi^{-1}(q(t))$. From (19) we also write $B_2(T_e) = a_2\,\mathrm{e}^{b_2\varepsilon_2}$ for $a_2 = B_2(t)\exp\big((r-\frac{\sigma_2^2}{2})(T_e-t)\big)$ and $b_2 = \sigma_2\sqrt{T_e-t}$. But ε_2 and ε_3 are bivariately normal with correlation ρ. Hence, if we define $\bar{\varepsilon}_3 = -\Phi^{-1}(q(t))$, we compute

$$\mathsf{E}_t^Q\big(\mathbf{1}_{B_3(T_e)\ge K_3} B_2(T_e)\big) \;=\; a_2 \int_{\bar\varepsilon_3}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{b_2\varepsilon_2}\, f_\rho(\varepsilon_2,\varepsilon_3)\,\mathrm{d}\varepsilon_2\,\mathrm{d}\varepsilon_3$$
$$=\; a_2 \int_{\bar\varepsilon_3}^{+\infty} \mathrm{e}^{b_2^2/2}\,\phi(\varepsilon_3-\rho b_2)\,\mathrm{d}\varepsilon_3 \;=\; a_2\;\mathrm{e}^{b_2^2/2}\;\Phi(\rho b_2-\bar\varepsilon_3). \tag{A15}$$

But $a_2\,\mathrm{e}^{b_2^2/2} \;=\; B_2(t)\,\mathrm{e}^{r(T_e-t)}$, and $-\bar\varepsilon_3 \;=\; \Phi^{-1}(q(t))$, which implies $\mathsf{E}_t^Q\big(q(T_e)B_2(T_e)\big) \;=\; B_2(t)\,\mathrm{e}^{r(T_e-t)}\,\Phi\big(\Phi^{-1}(q(t))+\rho\sigma_2\sqrt{T-t}\big)$. The rest of (24) is straightforward to prove.

Similarly, the only non-trivial part of proving (25) is the computation of

$$\mathsf{E}_t^Q\Big[\mathrm{e}^{-r(T-t)}\,q(T)\big(B_2(T)-K\big)_+\Big] \;=\; \mathsf{E}_t^Q\Big[q(T_e)\;\mathrm{e}^{-r(T_e-t)}\;\mathsf{E}_{T_e}^Q\Big(\mathrm{e}^{-(T-T_e)}\big(B_2(T)-K\big)_+\Big)\Big]$$
$$=\; \mathsf{E}_t^Q\Big[\mathbf{1}_{B_3(T_e)\ge K_3}\;\mathrm{e}^{-r(T_e-t)}\;C_{\mathrm{BS}}\big(B_2(T_e),T-T_e\big)\Big], \tag{A16}$$

where $C_{\mathrm{BS}}(S,T-t)$ is the Black–Scholes formula (A6) with arguments K, r and σ_2 omitted. Here we can no longer integrate along ε_2 as before, because the function C_{BS} is too complicated. Nevertheless, we integrate along ε_3, using the formula

$$\int_{\bar\varepsilon_3}^{+\infty} f_\rho(\varepsilon_2,\varepsilon_3)\,\mathrm{d}\varepsilon_3 \;=\; \phi(\varepsilon_2)\Phi\left(\frac{\rho\varepsilon_2-\bar\varepsilon_3}{\sqrt{1-\rho^2}}\right). \tag{A17}$$

Substituting (A17) into (A16), we get the same integral as in (25). The rest of the formula (25) is straightforward to prove. □

Appendix B. MCMC Procedure for Cash Mergers

In this section we describe a Markov Chain Monte Carlo (MCMC) method based on the state space representation of our model. The goal is to use the time series of observed stock prices and prices of various call options on the target companies, and estimate the time series of the two latent variables: the success probability of the merger $q(t)$, and the fallback price of the target $B_2(t)$.

In order to simplify the presentation, we start with specifications for q and B_2 that are slightly different than the ones used in the empirical study.[46] Define the state variables X_1 and X_2 as Itô processes with constant drift and volatility:

$$\mathrm{d}X_i(t) \;=\; \mu_i \mathrm{d}t + \sigma_i \mathrm{d}W_i(t), \quad i = 1, 2, \tag{B1}$$

where $W_1(t)$ and $W_2(t)$ are independent standard Brownian motions. The latent variables q and B_2 are related to the state variables by the following equations:

$$q(t) \;=\; \frac{\mathrm{e}^{X_1(t)}}{1 + \mathrm{e}^{X_1(t)}}, \qquad B_2(t) \;=\; \mathrm{e}^{X_2(t)}. \tag{B2}$$

The latent variables and the observed variables are connected by the observation equation, which puts together equations (14) and (15):

$$\begin{aligned} B(t) &= q(t) B_1 \,\mathrm{e}^{-r(T_e - t)} + (1 - q(t)) B_2(t) + \varepsilon_B(t), \\ C(t) &= q(t)(B_1 - K)_+ \,\mathrm{e}^{-r(T_e - t)} + (1 - q(t)) C_{\mathrm{BS}}(B_2(t), K, T - t) + \varepsilon_C(t), \end{aligned} \tag{B3}$$

where $C_{\mathrm{BS}}(S, K, T - t)$ is the Black–Scholes formula (A6) with arguments r and σ_2 omitted. The errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$ are IID normal with zero mean, and independent from each other. If more than one call option is employed in the estimation process, $C(t)$ is multi-dimensional.

To simplify notation, we rename the observed variables: $Y_B = B$, and $Y_C = C$. The state variables are collected under $X = [X_1, X_2]^T$, and the observed variables are collected under $Y = [Y_B, Y_C]^T$. (The superscript T after a vector indicates transposition.) There are other observed parameters: the effective date (T_e), the interest rate (r), the cash offer (B_1), and the strike prices (K) and maturities (T) of various call options on the company B.

The vector of latent parameters is $\theta = [\mu_1, \mu_2, \sigma_1, \sigma_2]^T$. The observation equations (14) and (15) can be rewritten as $Y = f(X, \theta) + \varepsilon$, where $\varepsilon = [\varepsilon_B, \varepsilon_C]^T$ is the vector of model errors. The diagonal matrix of model error variances, $\Sigma_\varepsilon = \mathrm{diag}\big(\sigma^2_{\varepsilon_B}, \sigma^2_{\varepsilon_C}\big)$ is called the matrix of hyperparameters.

The Markov Chain Monte Carlo (MCMC) method provides a way to sample from

[46] Using Itô calculus, one sees that the specifications of q and B_2 given here in equation (B2) differ from the ones used in the empirical part in equations (11) and (12) only up to a drift term.

the posterior distribution with density $p(\theta, X, \Sigma_\varepsilon|Y)$, and then estimate the parameters θ, the state variables X, and the hyperparameters Σ_ε. Bayes' Theorem says that the posterior density is proportional to the likelihood times the prior density. In our case, $p(X, \Sigma_\varepsilon, \theta|Y) \propto p(Y|X, \Sigma_\varepsilon, \theta) \cdot p(X, \Sigma_\varepsilon, \theta) = p(Y|X, \Sigma_\varepsilon, \theta) \cdot p(X|\theta) \cdot p(\Sigma_\varepsilon) \cdot p(\theta)$. On the right hand side, the first term in the product is the likelihood for the observation equation; the second term is the likelihood for the state equation; and the third and fourth terms are the prior densities of the hyperparameters Σ_ε and the parameters θ. We obtain

$$p(Y|X, \Sigma_\varepsilon, \theta) = \prod_{t=1}^{T_e} \phi\big(Y(t)|f(X(t), \theta), \Sigma_\varepsilon\big), \quad p(X|\theta) = p(X(1)|\theta) \cdot \prod_{t=2}^{T_e} \phi\big(Z(t)|\mu, \Sigma_X\big), \tag{B4}$$

where $Z_i(t) = X_i(t) - X_i(t-1)$, $\mu = [\mu_1, \mu_2]^T$, $\Sigma_X = \text{diag}\big(\sigma_1^2, \sigma_2^2\big)$, and $\phi(x|\mu, \Sigma) = (2\pi)^{-n/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$ the n-dimensional multivariate normal density with mean μ and covariance matrix Σ

We now describe the MCMC algorithm.

STEP 0. Initialize $\theta^{(1)}$, $X^{(1)}$, $\Sigma_\varepsilon^{(1)}$. Fix a number of iterations M. Then for each $i = 1, \ldots M-1$ go through steps 1–3 below.

STEP 1. Update $\Sigma_\varepsilon^{(i+1)}$ from $p(\Sigma_\varepsilon|\theta^{(i)}, X^{(i)}, Y)$. Note that with a flat prior for Σ_ε,

$$p(\Sigma_\varepsilon|\theta^{(i)}, X^{(i)}, Y) \propto \prod_{t=1}^{T_e} \phi\big(Y(t)|f(X(t), \theta), \Sigma_\varepsilon\big). \tag{B5}$$

This implies that $\left(\sigma_{\varepsilon,j}^{(i+1)}\right)^2$, $j = B, C$, is sampled from an inverted gamma-2 distribution, $IG_2(s, \nu)$, where $s = \sum_{t=1}^{T_e}(Y_j(t) - f_j(X(t), \theta))^2$ and $\nu = T_e - 1$. The inverted gamma-2 distribution $IG_2(s, \nu)$ has log-density $\ln p_{IG_2}(x) = -\frac{\nu+1}{2}\ln(x) - \frac{s}{2x}$. One could also use a conjugate prior for Σ_ε, which is also an inverted gamma-2 distribution.

STEP 2. Update $X^{(i+1)}$ from $p(X|\theta^{(i)}, \Sigma_\varepsilon^{(i+1)}, Y)$. To simplify notation, denote by $\theta = \theta^{(i)}$, and $\Sigma_\varepsilon = \Sigma_\varepsilon^{(i+1)}$. Notice that $p(X|\theta, \Sigma_\varepsilon, Y) \propto p(Y|\theta, \Sigma_\varepsilon, X) \cdot p(X|\theta)$, assuming flat priors for X. Then, if $t = 2, \ldots, T_e - 1$,

$$p(X(t)|\theta, \Sigma_\varepsilon, Y) \propto \phi\big(Y(t)|f(X(t),\theta),\Sigma_\varepsilon\big) \cdot \phi\big(X(t) - X^{(i+1)}(t-1)|\mu,\Sigma_X\big) \cdot \phi\big(X^{(i)}(t) - X(t)|\mu,\Sigma_X\big). \tag{B6}$$

If $t = 1$, replace the second term in the product with $p(X(1)|\theta)$; and if $t = T$, drop the third term out of the product. This is a non-standard density, so we perform the Metropolis–Hastings algorithm to sample from this distribution. This algorithm is described at the end of next step.

STEP 3. Update $\theta^{(i+1)}$ from $p(\theta|X^{(i+1)}, \Sigma_\varepsilon^{(i+1)}, Y)$. To simplify notation, denote by $X = X^{(i+1)}$, and $\Sigma_\varepsilon = \Sigma_\varepsilon^{(i+1)}$. Assuming a flat prior for θ, $p(\theta|X,\Sigma_\varepsilon,Y) \propto p(Y|\theta,\Sigma_\varepsilon,X)\cdot p(X|\theta)$. Then, if we assume that $X(1)$ does not depend on θ,

$$p(\theta|X,\Sigma_\varepsilon,Y) \propto \prod_{t=1}^{T_e} \phi\big(Y(t)|f(X(t),\theta),\Sigma_\varepsilon\big) \cdot \prod_{t=2}^{T_e} \phi\big(X(t) - X(t-1)|\mu,\Sigma_X\big). \tag{B7}$$

Recall that $\theta = [\mu_1, \mu_2\sigma_1, \sigma_2]^T$. For μ_1, μ_2, and σ_1, we can drop the first product from the formula, since it does not contain those parameters. In that case, we have the following updates: $\mu_k^{(i+1)} \sim \Phi\Big(\frac{1}{T_e-1}\sum_{t=2}^{T_e}(X_k(t) - X_k(t-1))^2\ ,\ \frac{(\sigma_k^{(i)})^2}{T_e-1}\Big)$, $k = 1, 2$; $\big(\sigma_1^{(i+1)}\big)^2 \sim IG_2\Big(\sum_{t=2}^{T_e}\big(X_1(t) - X_1(t-1) - \mu_1^{(i+1)}\big)^2\ ,\ T_e - 2\Big)$. For the other parameters the density is non-standard, so we need to perform the Metropolis–Hastings algorithm.

METROPOLIS–HASTINGS. The goal of this algorithm is to draw from a given density $p(x)$. Start with an element X_0, which is given to us from the beginning. (E.g., in the MCMC case, X_0 is the value of a parameter $\theta^{(i)}$, while X is the updated value $\theta^{(i+1)}$). Take another density $q(x)$, from which we know how to draw a random element. Initialize $X_{CURR} = X_0$. The Metropolis–Hastings algorithm consists of the following steps:

(1) Draw $X_{PROP} \sim q(x|X_{CURR})$ (this is the "proposed" X).

(2) Compute $\alpha = \min\Big(\frac{p(X_{PROP})}{p(X_{CURR})}\ \frac{q(X_{CURR}|X_{PROP})}{q(X_{PROP}|X_{CURR})}\ ,\ 1\Big)$.

(3) Draw $u \sim U[0,1]$ (the uniform distribution on $[0,1]$). Then define $X^{(i+1)}$ by: if

$u < \alpha$, $X^{(i+1)} = X_{PROP}$ ("accept"); if $u \geq \alpha$, $X^{(i+1)} = X_{CURR}$ ("reject").

Typically, we use the "Random-Walk Metropolis–Hastings" version, for which $q(y|x) = \phi(x|0, a^2)$, for some positive value of a. Equivalently, $X_{PROP} = X_{CURR} + e$, where $e \sim \mathcal{N}(0\ ,\ a^2)$.

In our empirical study, we choose the number of iterations to be $M = 400,000$ and we observe that the algorithm typically converges (i.e., the estimated posterior density appears stationary) after an initial "burn-out" period of about $200,000$ iterations. With these numbers, it takes our algorithm on average about one day per firm to finish at our current computing speed. Since we use the Metropolis–Hastings algorithm described above, we follow standard procedure and require that the average acceptance ratios are between 0.04 and 0.96; otherwise, we modify the value of the random walk parameter a until we obtain acceptance ratio in this interval. If this step fails as well, we exclude the company from the sample. A total of 2 companies have been excluded for this reason.

REFERENCES

[1] Bakshi, Gurdip, Charles Cao, and Zhiwu Chen (1997): Empirical Performance of Alternative Option Pricing Models, *Journal of Finance*, **52**, 2003–2049.

[2] Barone-Adesi, Giovanni, Keith C. Brown, and W.V. Harlow (1994): On the Use of Implied Volatilities in the Prediction of Successful Corporate Takeovers, *Advances in Futures and Options Research*, **7**, 147–165.

[3] Barraclough, Kathryn, David T. Robinson, Tom Smith, and Robert E. Whaley (2013): Using Option Prices to Infer Overpayments and Synergies in M&A Transactions, *Review of Financial Studies*, **26**, 695–722.

[4] Berndt, Antje, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranz (2005): Measuring Default Risk Premia from Default Swap Rates and EDFs; BIS working paper.

[5] Black, Fischer (1989): How We Came Up with the Option Formula, *Journal of Portfolio Management*, **15**, 4–8.

[6] Black, Fischer, and Myron Scholes (1973): The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, **81,** 637–654.

[7] Brown, Keith C., and Michael V. Raymond (1986): Risk Arbitrage and the Prediction of Successful Corporate Takeovers, *Financial Management*, **15**, 54–63.

[8] Cao, Charles, Zhiwu Chen, and John M. Griffin (2005): Informational Content of Option Volume Prior to Takeovers, *Journal of Business*, , **78**, 1073–1109.

[9] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997): Measuring Mutual Fund Performance with Characteristic-Based Benchmarks *Journal of Finance*, **52**, 1035–1058.

[10] Dubinsky, Andrew, and Michael Johannes (2005): Earnings Announcements and Equity Options; working paper.

[11] Duffie, Darrell (2001): *Dynamic Asset Pricing Theory*, 3rd Edition, Princeton University Press.

[12] Duffie, Darrell, and Kenneth J. Singleton (1997): An Econometric Model of the Term Structure of Interest-Rate Swap Yields, *Journal of Finance*, **52**, 1287–1321.

[13] Dukes, William, Cheryl Frolich, and Christopher Ma (1992): Risk Arbitrage in Tender Offers: Handsome Rewards—and not for Insiders Only, *Journal of Portfolio Management*, **18**, 47–55.

[14] Fama, Eugene F., and Kenneth R. French (1993): Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, **33**, 3–56.

[15] Hietala, Pekka, Steven N. Kaplan, and David T. Robinson (2003): What Is the Price of Hubris? Using Takeover Battles to Infer Overpayments and Synergies, *Financial Management*, **32**, 5–31.

[16] Jackwerth, Jens C., and Mark Rubinstein (1996): Recovering Probability Distributions from Option Prices, *Journal of Finance*, **51**, 1611–1631.

[17] Johannes, Michael, and Nicholas Polson (2003): MCMC Methods for Continuous-Time Financial Econometrics; prepared for the *Handbook of Financial Econometrics*, December.

[18] Jacquier, Eric, Michael Johannes, and Nicholas Polson (2007): MCMC Maximum Likelihood for Latent State Models, *Journal of Econometrics*, **137**, 615–640.

[19] Jindra, Jan, and Ralph A. Walkling (2004): Speculation Spreads and the Market Pricing of Proposed Acquisitions, *Journal of Corporate Finance*, **10**, 495–526.

[20] Karatzas, Ioannis, and Steven E. Shreve (1998): *Methods of Mathematical Finance*, Springer.

[21] Koop, Gary (2003): *Bayesian Econometrics*, Wiley.

[22] Mitchell, Mark, and Todd Pulvino (2001): Characteristics of Risk and Return in Risk Arbitrage, *Journal of Finance*, **56**, 2135–2175.

[23] Pan, Jun, and Kenneth J. Singleton (2008): Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads, *Journal of Finance*, **63**, 2345–2384.

[24] Rubinstein, Mark (1994): Implied Binomial Trees, *Journal of Finance*, **49**, 771–818.

[25] Samuelson, William, and Leonard Rosenthal (1986): Price Movements as Indicators of Tender Offer Success, *Journal of Finance*, **41**, 481–499.

[26] Subramanian, Ajay (2004): Option Pricing on Stocks in Mergers and Acquisitions, *Journal of Finance*, **59**, 795–829.

Figure 1: Median Implied Volatility Smile for Successful and Failed Mergers. This figure plots Black–Scholes daily implied volatilities (σ_{implied}) for the call options on the target company of the 843 cash mergers announced between January 1996 and December 2014 with options traded on the target company. At each date between the announcement date and the effective date of the merger, we select the options with the earliest expiration date after the effective date. Option prices are either the bid price ('bid'), ask price ('ask'), or bid-ask mid-quote ('mid'). On the x-axis is the option moneyness m: if K is the strike price of the option and B_1 the cash offer price per share, for each $m \in 0.7, 0.75, \ldots, 1.2$ we consider the median implied volatility σ_{implied} corresponding to the call options for which K/B_1 rounds to m. For each m we plot two median implied volatilities: one computed only for the target companies in deals that eventually succeeded ('success'), and one computed only for the target companies in deals that eventually failed ('failure').

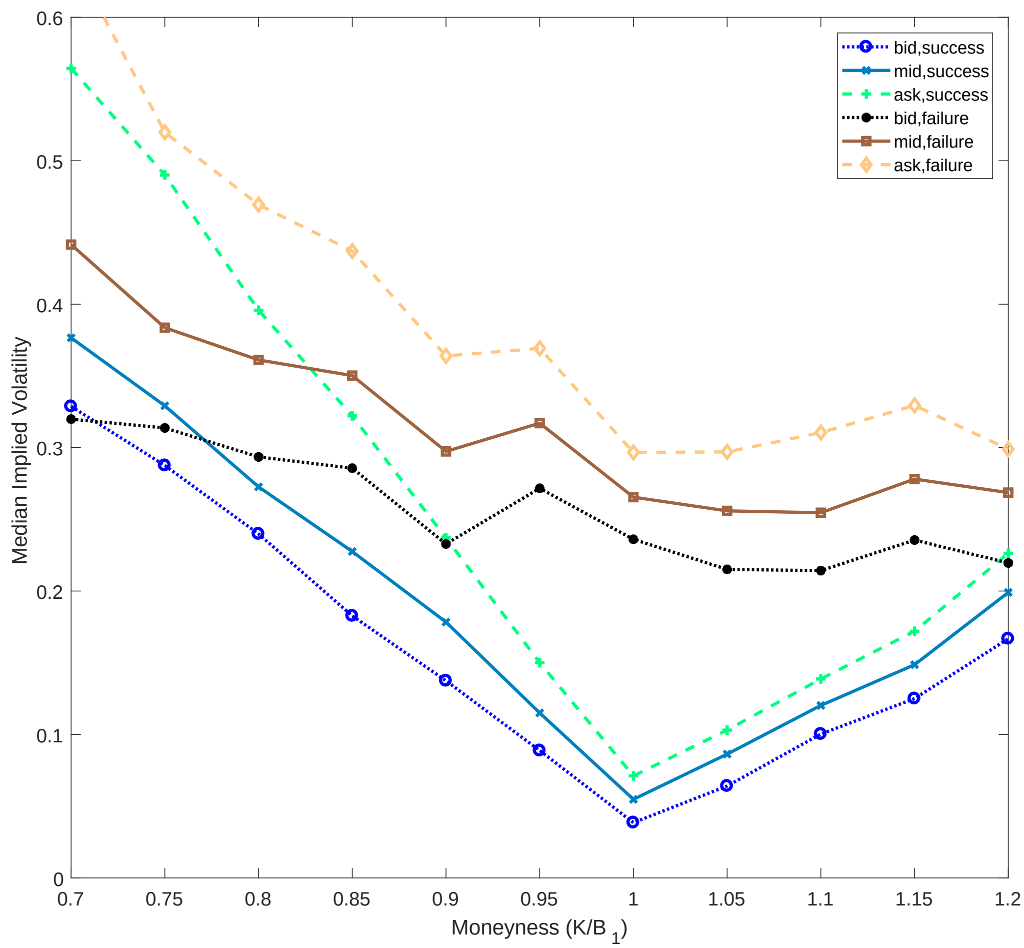

Figure 2: Time Series of Median Implied Volatilities for Successful and Failed Mergers. This figure plots the time series of median Black–Scholes implied volatilities (σ_{implied}) for at-the-money call options on the target company of the 843 cash mergers announced between January 1996 and December 2014 with options traded on the target company. At each date between the announcement date and the effective date of the merger, we select the options with the earliest expiration date after the effective date. Option and underlying prices are the average between quoted ask and bid prices. The time is from 20 days before to 20 days after the merger announcement date, and 20 days before the date by which the merger either succeeds or fails. We mark with "$*$" the median σ_{implied} for the deals that eventually succeeded, and with "$\circ$" the median σ_{implied} for the deals that eventually failed.

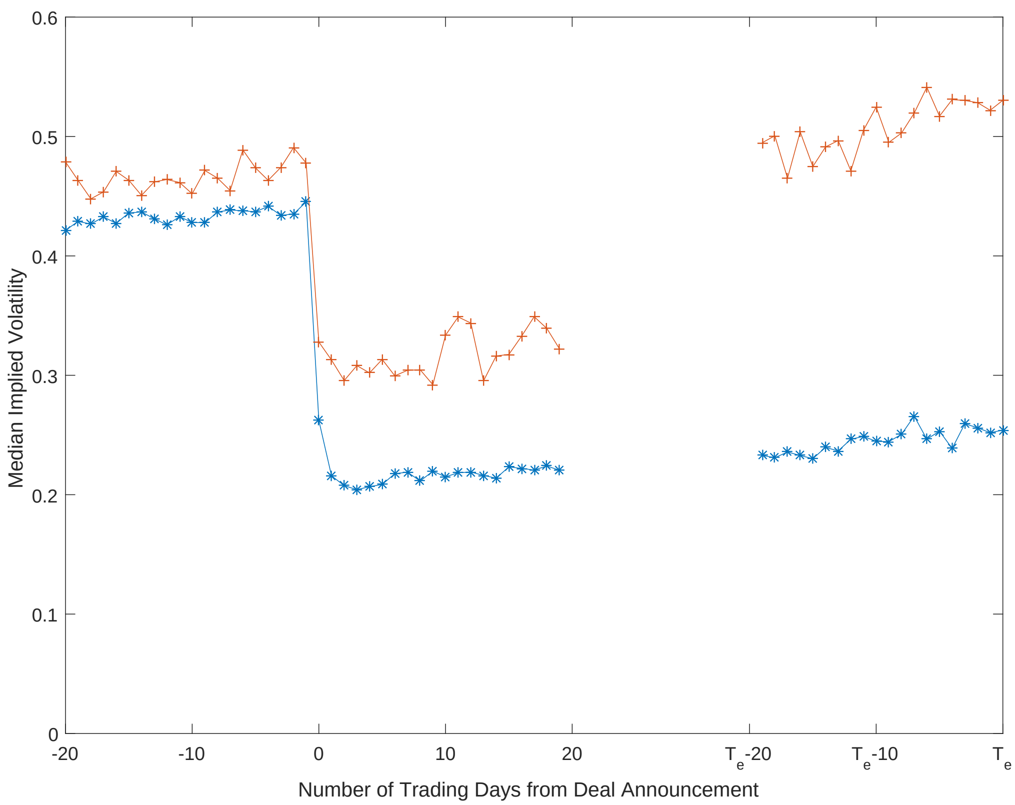

Figure 3: Success Probability Estimates for Ten Selected Cash Mergers. Using the methodology in the paper, this figure plots the success probability for the subsample of ten cash mergers described in Table 3. The deals corresponding to target tickers BUD, AWE, TXU, AT, FDC succeeded, while those for SLM, MCIC, UCL, HET, AVP failed. The dash-dotted lines represent the 5% and 95% error bands around the estimated median values. The estimates are obtained using the call options on the target company with the earliest expiration date after the effective date of the merger.

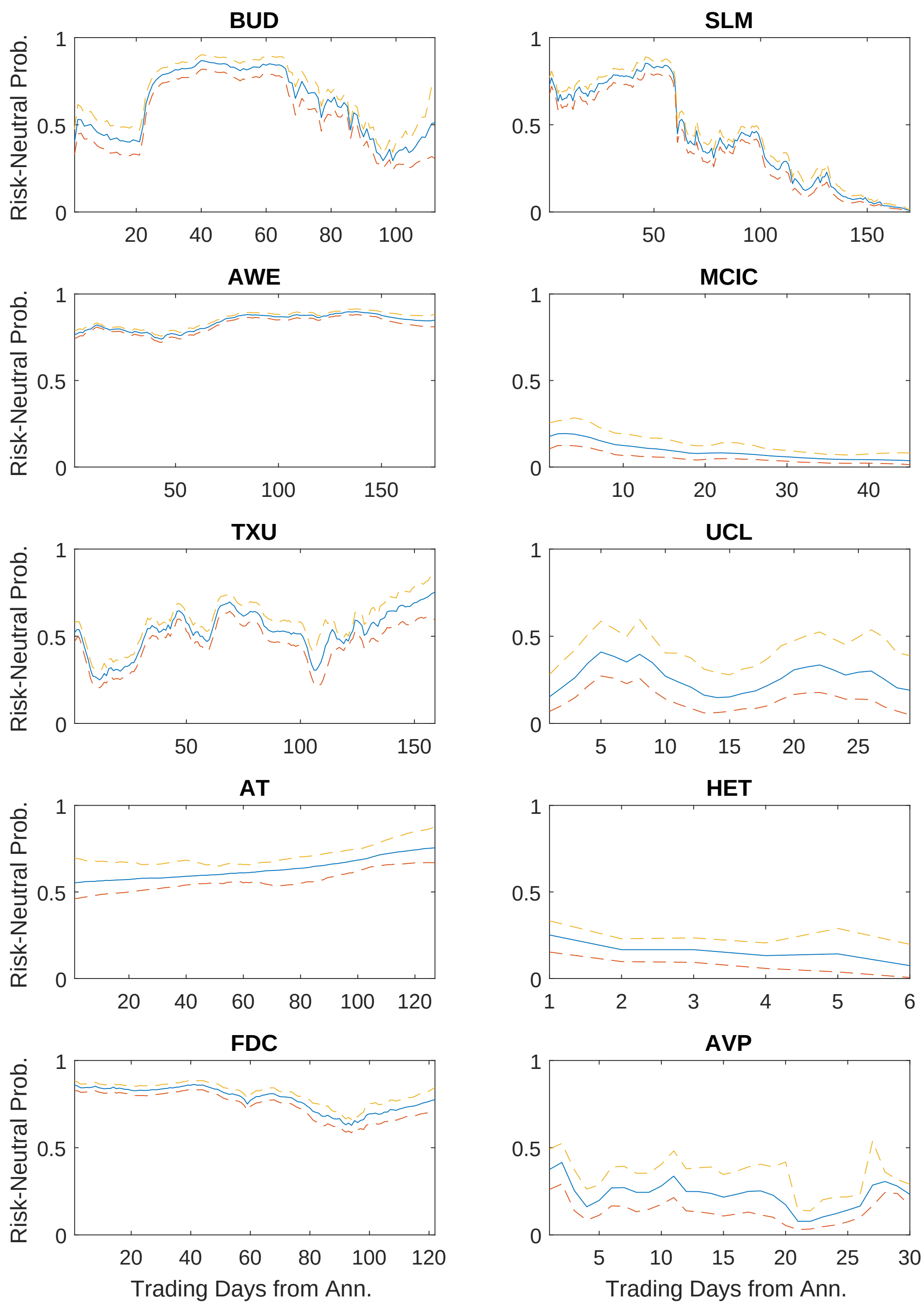

Figure 4: Comparison of Observed and Theoretical Volatility Smiles for AWE. From the sample of ten large cash merger deals in Table 3 we select the deal with the largest offer premium (75.44%). The deal has target company AT&T Wireless, with ticker AWE. The figure plots the observed and theoretical volatility smiles of call options traded on AWE, for equally spaced trading days during the merger deal. The option expiration date is the earliest date after the effective date of the merger. On the x-axis we plot the ratio K/B_1 (call strike price K to merger offer price B_1). On the y-axis we plot (i) the Black–Scholes implied volatility for the observed call price, using either a star or a dot: a star for an option with positive trading volume, or a dot for an option with zero volume; (ii) the Black–Scholes implied volatility for the theoretical call price (based on the model in this paper), using a continuous solid line. The parameters used to compute the ghest trading volume on that day.

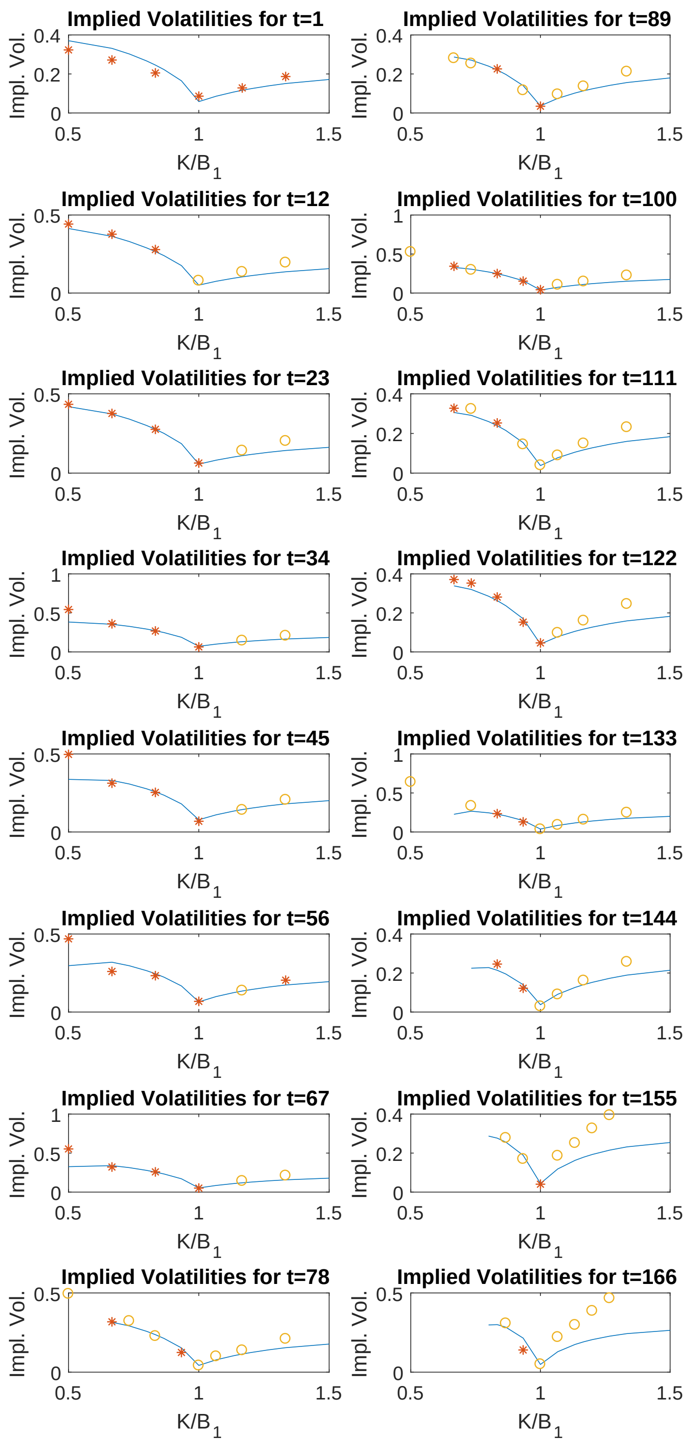

Figure 5: Pricing Calls with Expiration before the Effective Date. This figure displays pricing errors for European call options on the target company B of a merger, when the expiration date T is smaller than the merger effective date T_e. The figure compares prices compute at $t = 0$: (i) the correct price C_{correct} from equation (22), and (ii) the "simple" price C_{simple} from equation (7) which is correct only if T is larger than T_e. On the x-axis is the time T until expiration, and on the y-axis is the pricing error $(C_{\text{simple}} - C_{\text{correct}})/C_{\text{correct}}$. The parameters used in the formulas are: $T_e = 100$ days (effective date) divided by $252 =$ the number of trading days in 1 year; $B_1 = 100$ (offer price); $B_2(0) = 90$ (fallback price at $t = 0$); $q(0) = 0.5$ (success probability at $t = 0$); $r = 0.05$ (annual risk-free rate); $\sigma_2 = 0.4$ (annual fallback price volatility). The stock price at $t = 0$ corresponding to these parameters is $B(0) = 94.018$. The figure display the results when the expiration date T varies between 80 and 100 days, and the strike price K varies between 75 and 105 ($K = 95$ corresponds to the at-the-money call).

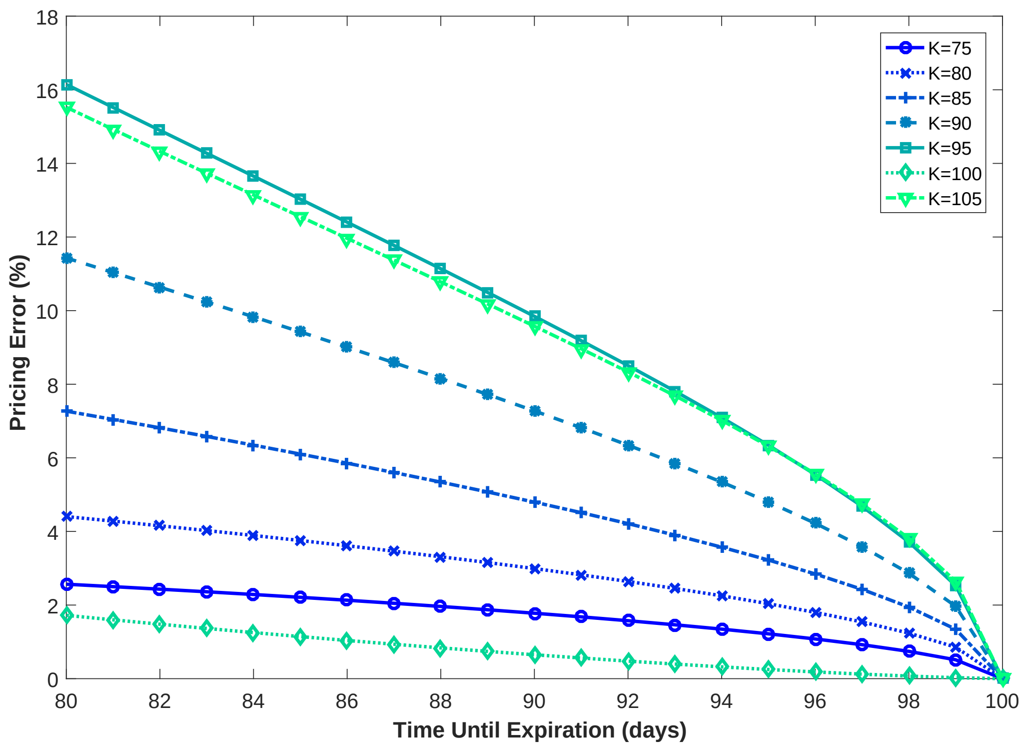

Table 1: Data Description. This table shows summary statistics for our sample of cash mergers from January 1996 to December 2014 for which there are option prices quoted on the target company. Included are the mean, standard deviation, and various percentiles for (i) the number of trading days between deal announcement and deal conclusion (Deal Duration); (ii) the percentage difference between the offer price per share and the share price for the target company one day before the deal was announced (Offer Premium); (iii) the percentage of trading days for which there is at least one option with positive trading volume (Frac. Days Calls Traded); (iv) the average number of call option contracts (1 contract = 100 options) traded on the target company (Ave. Call Volume); (v) the number of times the offer price changed during the merger deal (Offer Price Changes); (vi) the number of times the effective date of the merger was changed during the merger deal (Effec. Date Changes). N represents the number of firms with non-missing estimates.

			Percentile							
Variable	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
Deal Duration (trading days)	66.6	42.5	5	25	35.5	55	84	149.9	402	812
Offer Premium (%)	33.51	37.50	-95.13	3.27	16.50	28.07	43.42	78.93	729.15	810
Frac. Days Calls Traded (%)	65.04	28.66	0	15.02	41.38	68.29	92.82	100	100	812
Ave. Call Volume	1085.4	4389.5	0	5.9	30.5	113.8	467.8	4070.6	59143.2	812
Offer Price Changes	0.13	0.43	0	0	0	0	0	1	5	812
Effec. Date Changes	0.74	1.78	0	0	0	0	1	4	21	812

Table 2: Bid-Ask Spreads of Call Options. For target company i in our sample of cash mergers, we report statistics for μ_i, the average absolute bid-ask spread (Panel A) or relative bid-ask spread (Panel B). This average is computed over several categories of options that depend on the option's moneyness $m = K/S$ (K is the strike price and S is the underlying stock price): all calls; deep-in-the-money (Deep-ITM) calls, with $m < 0.9$; in-the-money (ITM) calls, with $m \in [0.9, 0.95)$; near-in-the-money (Near-ITM) calls, with $m \in [0.95, 1]$; near-out-the-money (Near-OTM) calls, with $m \in [1, 1.05]$; out-of-the-money (OTM) calls, with $m \in (1.05, 1.1]$; and deep-out-of-the-money (Deep-OTM) calls if $m > 1.1$. Entries with zero bid-ask spread or zero bid price are considered as missing. N represents the number of target firms with at least one non-missing estimate.

A. Absolute Bid-Ask Spread ($)

Selection	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
			Percentile							
All Calls	1.08	0.98	0.05	0.23	0.41	0.71	1.36	3.48	4.78	804
Deep-ITM Calls	1.26	1.09	0.15	0.29	0.47	0.84	1.66	3.78	4.83	790
ITM Calls	0.79	0.80	0.08	0.17	0.28	0.49	0.92	2.51	4.58	568
Near-ITM Calls	0.54	0.55	0.04	0.11	0.23	0.35	0.63	1.72	4.31	528
Near-OTM Calls	0.37	0.43	0.03	0.06	0.14	0.24	0.41	1.15	3.92	549
OTM Calls	0.36	0.44	0.03	0.06	0.14	0.23	0.40	1.00	4.95	465
Deep-OTM Calls	0.38	0.57	0.03	0.05	0.14	0.22	0.41	1.11	4.95	487

B. Relative Bid-Ask Spread (%)

Selection	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
			Percentile							
All Calls	27.46	20.00	3.76	8.64	14.79	22.06	33.28	66.22	195.70	804
Deep-ITM Calls	17.29	15.94	2.25	4.53	7.88	12.28	20.53	50.30	145.11	790
ITM Calls	28.50	23.78	4.54	7.51	13.51	21.04	35.58	74.73	163.64	568
Near-ITM Calls	41.22	29.84	6.13	11.09	19.87	31.61	53.81	106.37	160.00	528
Near-OTM Calls	74.92	35.67	8.09	21.06	49.36	72.22	97.19	137.73	181.81	549
OTM Calls	85.16	35.73	5.77	26.52	59.62	86.23	112.78	139.01	196.04	465
Deep-OTM Calls	94.93	34.49	5.37	39.01	67.87	93.15	119.23	154.81	196.04	487

Table 3: Ten Selected Cash Mergers. This table shows information on ten large deals sorted on offer value (offer price per share times the target's number of shares outstanding). From our sample of cash mergers with options traded on the target company, we select the five largest deals that succeeded, and the five largest deals that failed. Panel A shows the names of the acquirer and target company, the ticker of the target company, and the offer value in billion U.S. dollars. Panel B shows the target's ticker, the deal announcement date, the date when the deal succeeded or failed, the target's stock price one day before the announcement $B(0)$, the offer price B_1, and the offer premium which is the percentage change $\frac{B_1-B(0)}{B(0)}$.

A. List of Deals

Acquirer Name	Target Name	Tgt.Ticker	Offer Value ($ bn)
InBev NV	Anheuser-Busch Cos Inc	BUD	49.92
Cingular Wireless LLC	AT&T Wireless Services Inc	AWE	40.72
TXU Corp SPV	TXU Corp	TXU	31.80
Atlantis Holdings LLC	Alltel Corp	AT	25.76
Kohlberg Kravis Roberts & Co	First Data Corp	FDC	25.60
Investor Group	SLM Corp	SLM	24.53
GTE Corp	MCI Communications Corp	MCIC	22.24
China National Offshore Oil	Unocal Corp	UCL	18.20
Penn National Gaming Inc	Harrahs Entertainment Inc	HET	16.17
Coty US Inc	Avon Products Inc	AVP	10.67

B. Deal Information

Target Ticker	Date Announced	Date Ended	Deal Status	Price 1 Day. Before Ann.	Offer Price ($)	Offer Premium (%)
BUD	6/11/2008	11/18/2008	Completed	50.23	70	37.56
AWE	2/17/2004	10/26/2004	Completed	8.55	15	75.44
TXU	2/26/2007	10/10/2007	Completed	57.64	69.25	20.14
AT	5/20/2007	11/16/2007	Completed	58.19	71.5	22.87
FDC	4/2/2007	9/24/2007	Completed	26.90	34	26.39
SLM	4/16/2007	12/13/2007	Withdrawn	40.75	60	47.24
MCIC	10/15/1997	12/17/1997	Withdrawn	25.13	40	59.20
UCL	6/22/2005	8/2/2005	Withdrawn	44.34	67	51.11
HET	12/12/2006	12/19/2006	Withdrawn	78.46	87	10.88
AVP	4/2/2012	5/14/2012	Withdrawn	18.52	24.75	26.51

Table 4: Relative Pricing Errors for the Stock Price. For target company i in our sample of cash mergers, we report statistics for μ_i, the time series average of the relative stock pricing error. On trading day t, the relative pricing error is $(S_i^{\text{BMR}}(t) - S_i(t))/S_i(t)$, where $S_i(t)$ is the target company's observed stock price, and $S_i^{\text{BMR}}(t)$ is the theoretical stock price (based on the model in this paper). The parameters used to compute the theoretical price are estimated by using each day only the option with the highest trading volume on that day. We winsorize the pricing error at 100%: for one merger deal the average pricing error is 130.89%. N represents the number of firms with non-missing estimates.

			Percentile							
	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
Pricing Error (%)	0.149	0.358	0.001	0.003	0.009	0.032	0.122	0.653	3.511	811

Table 5: Pricing Errors for Call Options. This table shows relative pricing errors, measured in percentage points, for call options traded on the target company in our sample of cash mergers. We consider only call options with the earliest expiration date after the merger effective date. The theoretical price C^{theory} are computed according to two models: BMR (the model in this paper) and BS (the Black–Scholes model). For target company i we report statistics for μ_i, the average relative pricing error for call options traded on i. This average is computed over several categories of options that depend on the option's moneyness $m = K/S$ (K is the strike price and S is the underlying stock price): all calls; deep-in-the-money (Deep-ITM) calls, with $m < 0.9$; in-the-money (ITM) calls, with $m \in [0.9, 0.95)$; near-in-the-money (Near-ITM) calls, with $m \in [0.95, 1]$; near-out-the-money (Near-OTM) calls, with $m \in [1, 1.05]$; out-of-the-money (OTM) calls, with $m \in (1.05, 1.1]$; and deep-out-of-the-money (Deep-OTM) calls if $m > 1.1$. On each day t we require the first category (all calls) to have at least eight (quoted) options, and the other categories (calls of a given moneyness) to have at least two options, otherwise we declare as missing the pricing errors on that day. The (relative) error is defined as $|C^{\text{theory}}(t) - C(t)/C(t)|$, where $C^{\text{BMR}}(t)$ is the theoretical call price computed using the BMR model; $C^{\text{BS}}(t)$ is the theoretical call price computed using the BS model, where the volatility parameter is constant and equal to the average over t of the at-the-money Black–Scholes implied volatility corresponding to the observed underlying and call prices; and $C(t)$ is the observed call price (at the bid-ask mid-quote). The parameters used to compute the BMR theoretical price are estimated by using each day only the option with the highest trading volume on that day. N represents the number of firms with non-missing estimates.

				Percentile							
Selection	Model	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
All Calls	BMR	9.39	10.65	0.35	1.41	3.35	6.06	11.59	27.69	72.88	243
	BS	19.49	36.92	0.51	1.74	3.94	8.14	19.13	68.61	385.95	243
Deep-ITM Calls	BMR	2.61	10.63	0.20	0.46	0.88	1.42	2.25	5.68	268.44	712
	BS	2.83	10.54	0.27	0.49	0.98	1.58	2.84	6.72	270.32	712
ITM Calls	BMR	6.97	9.18	1.55	1.68	3.45	4.84	7.41	17.01	69.10	58
	BS	9.26	9.21	1.73	2.12	4.85	7.13	10.60	21.57	67.76	58
Near-ITM Calls	BMR	12.95	12.03	2.67	3.13	7.27	10.10	13.86	32.13	85.29	58
	BS	20.33	13.37	4.12	7.21	10.72	16.70	27.62	44.13	78.40	58
Near-OTM Calls	BMR	40.62	50.73	4.97	6.84	13.85	30.84	49.23	89.76	338.46	46
	BS	98.17	116.87	8.86	12.67	35.06	56.37	72.62	386.70	536.47	46
OTM Calls	BMR	41.66	29.18	10.47	12.17	19.80	30.81	57.59	99.01	100.92	27
	BS	95.08	129.40	11.61	17.34	34.12	65.80	92.07	472.63	600.32	27
Deep-OTM Calls	BMR	66.90	66.40	12.47	17.50	36.91	56.50	77.97	136.17	564.92	128
	BS	84.17	67.42	7.10	21.86	50.33	82.09	99.52	147.76	607.71	128

Table 6: Ratio of Absolute Call Option Pricing Errors to Bid-Ask Spread. For target company i in our sample of cash mergers, we report statistics for μ_i, the average ratio of the absolute option pricing error to the option bid-ask spread. This average is computed over several categories of options that depend on the option's moneyness $m = K/S$ (K is the strike price, and S is the underlying stock price): all calls; deep-in-the-money (Deep-ITM) calls, with $m < 0.9$; in-the-money (ITM) calls, with $m \in [0.9, 0.95)$; near-in-the-money (Near-ITM) calls, with $m \in [0.95, 1]$; near-out-the-money (Near-OTM) calls, with $m \in [1, 1.05]$; out-of-the-money (OTM) calls, with $m \in (1.05, 1.1]$; and deep-out-of-the-money (Deep-OTM) calls if $m > 1.1$. The absolute pricing error is the difference $|C^{\text{theory}}(t) - C^{\text{obs}}(t)|$ between theoretical and observed call prices. The parameters used to compute the theoretical price are estimated by using each day only the option with the highest trading volume on that day. N represents the number of firms with non-missing estimates.

			Percentile							
Selection	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
All Calls	0.37	0.50	0.07	0.16	0.24	0.30	0.39	0.72	11.55	794
Deep-ITM Calls	0.22	0.65	0.00	0.04	0.09	0.14	0.22	0.50	13.17	779
ITM Calls	0.47	0.93	0.00	0.08	0.17	0.30	0.49	1.26	18.54	561
Near-ITM Calls	0.64	1.08	0.00	0.13	0.26	0.43	0.69	1.86	21.31	550
Near-OTM Calls	0.72	1.06	0.01	0.15	0.33	0.48	0.70	2.08	16.28	567
OTM Calls	0.65	0.77	0.01	0.19	0.37	0.50	0.65	1.78	12.51	535
Deep-OTM Calls	0.55	0.31	0.13	0.34	0.46	0.50	0.50	1.02	3.81	709

Table 7: Call Price Kink and Success Probability. For target company i in our sample of cash mergers, consider the call options traded on date t which expire at the earliest date after the effective date of the merger. Let $C_i^{\text{kink}}(t)$ be the call price kink, i.e., the difference between the right slope and left slope of option prices corresponding to the strike prices K nearest to the merger offer price. Let $q_i(t)$ be the success probability estimated at t using only one option per day (with the highest trading volume). We write the same variables with a "tilde" ($\tilde{C}_i^{\text{kink}}(t)$ and $\tilde{q}_i(t)$) by truncating their values to be in the interval $(0, 1)$ and inverting them via the standard normal CDF function. The table shows the results of panel regressions of the call price kink on the success probability. Time fixed-effects are at the period level, where the period corresponding to date t is given by $\lceil \frac{10t}{T_e} \rceil$, where T_e is the effective date of the merger (there are ten periods for each firm). Standard errors are clustered by firm, and the corresponding t-statistics are in parentheses.

	C^{kink}	C^{kink}	C^{kink}	C^{kink}	$\tilde{C}^{\text{kink}}$	$\tilde{C}^{\text{kink}}$	$\tilde{C}^{\text{kink}}$	$\tilde{C}^{\text{kink}}$
q	0.705	0.547	0.693	0.527				
	(13.69)	(12.02)	(13.22)	(11.22)				
$\tilde{q}$					0.954	0.801	0.941	0.782
					(11.68)	(10.65)	(11.23)	(9.79)
const.	0.230	0.321	0.237	0.333	0.291	0.324	0.294	0.329
	(6.76)	(12.20)	(6.81)	(12.26)	(5.32)	(19.44)	(5.32)	(18.55)
Firm FE	NO	YES	NO	YES	NO	YES	NO	YES
Time FE	NO	NO	YES	YES	NO	NO	YES	YES
N	41,534	41,532	41,534	41,532	41,534	41,532	41,534	41,532
Clusters	683	681	683	681	683	681	683	681
R^2	8.85%	42.83%	9.06%	43.01%	9.12%	51.34%	9.33%	51.51%

Table 8: Return Volatility and Success Probability. For target company i in our sample of cash mergers, we divide the time interval between the merger announcement date ($t = 1$) and the effective date ($t = T_e$) into five integer periods as follows: if $\tau = \lfloor \frac{T_e}{5} \rfloor$, let $I_1 = [1, \tau]$, $I_2 = [\tau + 1, 2\tau]$, $I_3 = [2\tau + 1, 3\tau]$, $I_2 = [3\tau + 1, 4\tau]$, $I_5 = [4\tau + 1, T_e]$, where each integer in I_k is a trading day. Let $\sigma_{B,i,k}$ be the return volatility of firm i, measured as the standard deviation of i's daily stock return over the interval I_k, $k \in \{1, 2, 3, 4, 5\}$. If $q_i(t)$ be the success probability estimated at t using only one option per day (with the highest trading volume), denote by $q_{i,k}$ the average success probability over $t \in I_k$. Panel A shows the results of panel regressions of the return volatility σ_B on one minus the success probability. Time fixed-effects are at the period level. Standard errors are clustered by firm. Panel B shows the results of cross-sectional regressions of σ_B on the time-series average $1 - q$, the volatility estimates σ_q and σ_2 from, respectively, $\frac{\mathrm{d}q}{q(1-q)} = \mu_q \mathrm{d}t + \sigma_q \mathrm{d}W_q$ and $\frac{\mathrm{d}B_2}{B_2} = \mu_2 \mathrm{d}t + \sigma_2 \mathrm{d}W_2$, and the merger offer premium, $\pi = \frac{B_1 - B(0)}{B(0)}$. The t-statistics are in parentheses.

A. Panel Regressions

	σ_B (%)	σ_B (%)	σ_B (%)	σ_B (%)
$1 - q$	1.519	2.737	1.389	2.147
	(5.45)	(6.40)	(4.79)	(4.86)
const.	0.415	-0.167	0.465	0.169
	(3.43)	(-0.37)	(3.72)	(0.98)
Firm FE	NO	YES	NO	YES
Time FE	NO	NO	YES	YES
N	4,060	4,060	4,060	4,060
Clusters	812	812	812	812
R^2	1.47%	31.12%	5.49%	34.98%

B. Cross-Sectional Regressions

	$\ln(\sigma_B)$	$\ln(\sigma_B)$	$\ln(\sigma_B)$	$\ln(\sigma_B)$	$\ln(\sigma_B)$	$\ln(\sigma_B)$
$1 - q$	0.942		1.581		0.921	1.597
	(3.99)		(7.80)		(3.90)	(7.93)
σ_q		0.071	0.079			0.080
		(15.11)	(16.89)			(17.22)
σ_2		0.183	0.175			0.252
		(6.30)	(6.24)			(7.19)
π				0.255	0.243	-0.445
				(2.17)	(2.08)	(-3.61)
const.	-5.443	-5.482	-6.131	-5.162	-5.518	-6.026
	(-53.26)	(-121.17)	(-65.30)	(-87.28)	(-50.84)	(-61.61)
N	812	812	812	810	810	810
R^2	1.93%	25.49%	30.62%	0.45%	2.18%	31.68%

Table 9: Success Probability as Predictor of Deal Outcome. For target company i in our sample of cash mergers, let $q_i(t)$ be the success probability estimated at date t using only one option per day (with the highest trading volume). Also, let $q_i^{\text{naive}}(t)$ be the "naive" success probability, computed as the ratio $\frac{S_i(t)-S_i(0)}{S_i^{\text{offer}}-S_i(0)}$, where $S_i(t)$ is the current stock price, $S_i(0)$ is the stock price one day before the merger announcement, S_i^{offer} is the offer price, and the ratio is truncated to be in the interval $(0,1)$. To create a panel, we consider 10 equally spaced days, $t_1,\ldots,t_{10}$, throughout the life of the deal (between deal announcement and deal conclusion). We run probit regressions, where the dependent variable is the deal's outcome (success or failure), and the independent variable is either q for regressions (1) and (3), or q^{naive} for regressions (2) and (4). Regressions (1) and (2) are carried for all time values (t_1 to t_{10}), while regressions (3) and (4) are carried for the second half of the deal (t_6 to t_{10}). The t-statistics are in parentheses, and are obtained from standard errors clustered by firm.

	(1)	(2)	(3)	(4)
q	4.069		5.131	
	(13.04)		(13.79)	
q^{naive}		2.154		2.732
		(12.81)		(14.05)
const.	-0.937	-0.557	-1.450	-1.056
	(-6.24)	(-3.72)	(-8.38)	(-6.13)
N	8,120	8,100	4,060	4,050
Clusters	812	810	812	810
Pseudo-R^2	32.33%	20.01%	46.20%	29.92%

C Additional Figures and Tables

Figure 6: Median Implied Volatility Smile Before the Announcement Date. This figure plots Black–Scholes daily implied volatilities (σ_{implied}) for the call options on the target company of the 843 cash mergers announced between January 1996 and December 2014 with options traded on the target company. At each date between 30 days and 1 day before the announcement date of the merger, we select the options with the earliest expiration date after the effective date. Option prices are either the bid price ('bid'), ask price ('ask'), or bid-ask mid-quote ('mid'). On the x-axis is the option moneyness m: if K is the strike price of the option and B_1 the cash offer price per share, for each $m \in 0.7, 0.75, \ldots, 1.2$ we consider the median implied volatility σ_{implied} corresponding to the call options for which K/B_1 rounds to m.

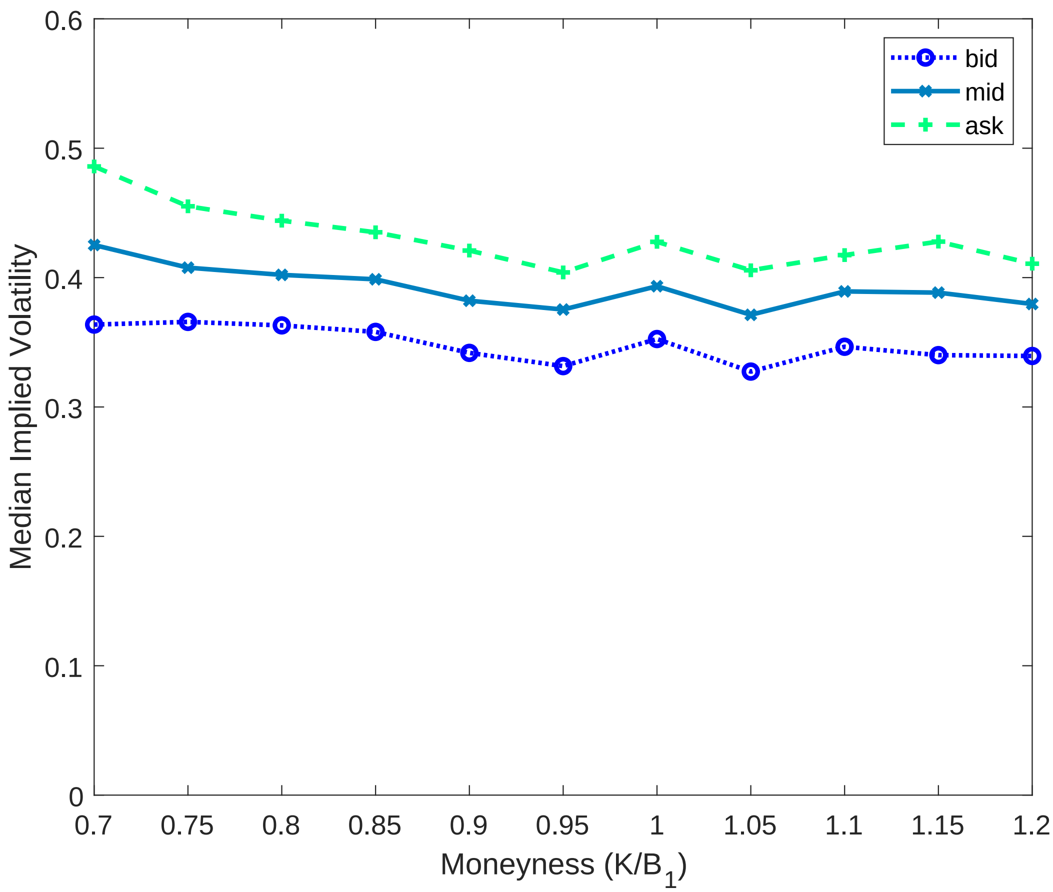

Table 10: Pricing Errors for Put Options. This table shows relative pricing errors, measured in percentage points, for put options traded on the target company in our sample of cash mergers. We consider only put options with the earliest expiration date after the merger effective date. The theoretical price P^{theory} are computed according to two models: BMR (the model in this paper) and BS (the Black–Scholes model). For target company i we report statistics for μ_i, the average relative pricing error for put options traded on i. This average is computed over several categories of options that depend on the option's moneyness $m = S/K$ (S is the underlying stock price, and K is the strike price): all puts; deep-in-the-money (Deep-ITM) puts, with $m < 0.9$; in-the-money (ITM) puts, with $m \in [0.9, 0.95)$; near-in-the-money (Near-ITM) puts, with $m \in [0.95, 1]$; near-out-the-money (Near-OTM) puts, with $m \in [1, 1.05]$; out-of-the-money (OTM) puts, with $m \in (1.05, 1.1]$; and deep-out-of-the-money (Deep-OTM) puts if $m > 1.1$. On each day t we require the first category (all puts) to have at least six (quoted) options, and the other categories (puts of a given moneyness) to have at least two options, otherwise we declare as missing the pricing errors on that day. The (relative) error is defined as $|P^{\text{theory}}(t) - P^{\text{mid}}(t)/P^{\text{mid}}(t)|$, where $P^{\text{BMR}}(t)$ is the theoretical European put price computed using the BMR model; $P^{\text{BS}}(t)$ is the theoretical European put price computed using the BS model, where the volatility parameter is constant and equal to the average over t of the at-the-money Black–Scholes implied volatility corresponding to the observed underlying and call prices; and $P^{\text{mid}}(t)$ is the observed put price at the bid-ask mid-quote. The parameters used to compute the BMR theoretical price are estimated by using each day only the option with the highest trading volume on that day. N represents the number of firms with non-missing estimates.

							Percentile				
Selection	Model	Mean	StDev	Min	5%	25%	50%	75%	95%	Max	N
All Puts	BMR	31.58	30.11	1.33	5.23	14.43	24.45	43.60	72.26	392.93	264
	BS	38.29	31.62	1.53	6.15	17.53	35.20	55.33	73.53	403.85	264
Deep-ITM Puts	BMR	5.35	27.22	0.16	0.40	0.97	2.22	4.41	13.62	551.61	434
	BS	5.35	27.26	0.16	0.41	0.95	2.35	4.52	15.15	558.82	434
ITM Puts	BMR	29.16	184.92	0.65	0.93	2.29	5.08	7.51	19.20	1484.73	64
	BS	30.53	191.99	0.71	0.95	2.46	4.84	8.11	20.54	1541.74	64
Near-ITM Puts	BMR	15.34	9.17	3.92	5.05	9.82	13.10	17.89	33.56	52.84	61
	BS	23.99	26.94	3.21	5.86	9.65	15.81	24.41	76.46	175.56	61
Near-OTM Puts	BMR	41.37	23.91	7.64	8.58	20.81	37.48	66.47	79.06	89.73	52
	BS	54.70	22.60	11.19	17.41	42.67	53.43	65.36	98.33	102.52	52
OTM Puts	BMR	53.60	32.54	9.77	9.84	26.70	49.09	82.78	100.00	111.23	34
	BS	77.23	28.95	14.20	24.26	60.84	85.03	96.99	100.00	153.66	34
Deep-OTM Puts	BMR	74.47	26.04	8.62	25.92	53.65	82.63	99.19	100.00	104.13	321
	BS	89.39	18.48	16.39	47.07	85.81	98.56	99.97	100.00	151.78	321