Option Pricing on Cash Mergers

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

September 30, 2009

Abstract

When a cash merger is announced but not completed, there are two main sources of uncertainty related to the target company: the probability of success and the price conditional on the deal failing. We propose an arbitrage-free option pricing formula that focuses on these sources of uncertainty. We test our formula in a study of all cash mergers between 1996 and 2008 which have sufficiently liquid options traded on the target company. The estimated success probability is a good predictor of the deal outcome. Our option formula for cash mergers does significantly better than the Black–Scholes formula and produces a volatility smile close to the one observed in practice. In particular, we provide an explanation for the kink in the volatility smile and show that the kink increases with the probability of deal success.

JEL Classification: G13, G34.

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

*Victor H. Martinez is with the Baruch College, CUNY, Zicklin School of Business, Ioanid Roşu and C. Alan Bester are with the University of Chicago, Booth School of Business. Emails: `victor_martinez@baruch.cuny.edu`, `irosu@chicagobooth.edu`, `cbester@chicagobooth.edu`. The authors are grateful for conversations with Malcolm Baker, John Cochrane, George Constantinides, Doug Diamond, Pierre Collin-Dufresne, Charlotte Hansen, Steve Kaplan, Monika Piazzesi, Ľuboš Pástor, Pietro Veronesi, and for helpful comments from participants in finance seminars at Chicago, CUNY, Princeton, Toronto.

1 Introduction

One of the most common violations of the Black–Scholes formula (see Black and Scholes (1973) is the *volatility smile*, which represents the tendency for the at-the-money option to have a lower implied volatility than the other options.[1] In the U.S. option markets this phenomenon was not observed before the 1987 market crash, but it appeared shortly thereafter (see Rubinstein (1994) or Jackwerth and Rubinstein (1996)). The emergence of a volatility smile is widely attributed to a change in traders' perceptions of a larger market crash, or equivalently to a change from the assumption of a log-normal distribution of equity prices to a bi-modal distribution due to the small probability of a large market crash.

A clear case where a company's stock price has a bi-modal distribution is when the company is the target of a merger attempt.[2] In a typical merger (or takeover, or acquisition, or tender offer), a company A, the acquirer, makes an offer to a company B, the target. The offer can be made with A's stock, with cash, or a combination of both. The offer is usually made at a significant premium from B's pre-announcement stock price. Therefore, the distribution of B's stock price can be thought as bi-modal: if the deal is successful, the target stock price rises to the offer price; if the deal is unsuccessful, the stock price reverts to a *fallback* price.[3]

In this paper we give an arbitrage-free formula that prices options on the target company of a cash merger by focusing on the main uncertainties surrounding the merger: the success probability and the fallback price. We test our formula in a study of all cash mergers during the 1996–2008 period with sufficiently liquid options traded on the target company. We find our model produces pricing errors 21% smaller on average than the Black–Scholes formula. We further find that the formula produces a volatility smile that is close to the one observed in practice.

In particular, the formula predicts a volatility smile with a kink at a strike price equal to the merger offer price. The magnitude of the kink (i.e., the difference between the slope of the volatility smile above and below the offer price) should be proportional to the discounted risk-

[1] Black and Scholes (1973) assume a constant volatility for the underlying stock price, which implies that the implied volatilities should be the same, irrespective of the strike price of the option.

[2] Other papers, including Black (1989), point out that the Black–Scholes formula is unlikely to work well when the company is the subject of a merger attempt.

[3] 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.

neutral probability that the deal is successful. Empirical results indicate that this prediction is strongly supported in the data. Our model can also explain why, as we observe in the data, the Black–Scholes implied volatility decreases when the deal is close to being successful: a success probability close to 1 leads to an implied volatility close to 0.

To set up the model, consider a company which is the target of a cash merger. If the deal is successful, the target company's shareholders receive a fixed offer price $\$B_1$ per share, while if the deal fails its price reverts to a fallback price $\$B_2$. We assume that B_2 has a log-normal distribution. We also assume that the success probability of the deal changes over time, i.e., it follows a stochastic process. 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.[4] When the success probability q and the fallback price B_2 are uncorrelated our formula takes a particularly simple form.

Our option formula can then be used to estimate the latent (unobserved) variables q and B_2 from the observables, in particular the price of the target company B and the prices of the various existing options on B. We estimate the times series of latent variables along with the structural parameters of the model using a Markov Chain Monte Carlo (MCMC) algorithm.[5] Though this algorithm is flexible enough to allow us to use any (or all) options traded on the stock, for simplicity we choose only one option each day, e.g., the call option with the maximum trading volume on that day. As we discuss later, this choice also allows us to perform specification checks on our model, including out of sample pricing of options with different strike prices.

We apply the option formula to all cash mergers during the 1996–2008 period with sufficiently liquid options traded on the target company. After removing companies with illiquid options, we obtain a final sample of 282 cash mergers. We test our model in three different ways. First, we compare the model-implied option prices to those coming from the Black-

[4]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 \$1 if the deal is successful and \$0 otherwise.

[5]For a discussion of Markov Chain Monte Carlo methods in finance, see the survey article of Johannes and Polson (2003). MCMC methods allow us to conduct inference by sampling from the joint distribution of model parameters and unobserved state variables (in this case q and B_2) given the observed data. MCMC methods for state space models are well established in Bayesian statistics and econometrics (see, e.g., Jacquier, Johannes, and Polson (2007)). MCMC is particularly convenient here because some parameters enter the model non-linearly, meaning standard techniques for latent variable models, such as the Kalman filter, would need to be modified. We emphasize that this choice is based on convenience and that many other estimation approaches (frequentist and Bayesian) are possible.

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—for 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. Finally, 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: the median percentage error is 26.06% for our model compared to an error of 33.02% in the case of the Black–Scholes model.[6] Our formula also does well compared to a modified Black–Scholes formula in which the volatility equals the previous-day implied volatility at the same strike price. This modified Black–Scholes formula is very difficult to surpass, as it already incorporates the observed volatility smile from the previous day. Even though we use only one option each day in our estimation process, our out-of-sample option pricing estimates are very close to the observed prices and therefore produce a volatility smile close to the observed one.

We test the implications of the model regarding the kink in the volatility smile. If instead of looking at the volatility plot we consider plotting the call option price against the strike price, then theoretically the magnitude of the kink normalized by the time discount coefficient should be precisely equal to the risk-neutral probability. A regression of the normalized kink on the estimated risk-neutral probability strongly supports the prediction that the intercept equals to 0 and the slope equals to 1. Our estimation procedure uses only one option each day, yet it matches well the whole cross section of options for that day, including the magnitude of the kink.

We show that the probabilities estimated using our formula predict the outcomes of deals in the data. In particular, this method does significantly better than the "naive" method widely used in the mergers and acquisitions literature, 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.

We also investigate how the fallback price compares to the price before the announcement. One might expect that the fallback price should be on average higher than the pre-

[6]The error is smaller than the average bid-ask spread for options in our sample.

announcement price. This may be due to the fact that a merger is usually a good signal about the quality of the target company, and indicate that other takeover attempts are now more likely. We find that indeed the fallback price is on average 27% higher than the pre-announcement price.

Another implication of our model is that the merger risk premium may be estimated as the drift coefficient in the diffusion process for the success probability. This is individually very noisy, but over the whole sample the estimated merger risk premium is significantly positive, at an 158% annual rate (with an error of $\pm 29\%$). This figure is comparable to the one obtained by Dukes, Frolich and Ma (1992), which examine arbitrage activity around 761 cash mergers between 1971 and 1985 and report returns to merger arbitrage of approximately 0.47% daily. See also Mitchell and Pulvino (2001) and Jindra and Walkling (2004) for a more detailed discussion of the risks and the transaction costs involved in merger arbitrage.

Background Literature

The literature on estimating the sources of uncertainty related to mergers is relatively scarce. Brown and Raymond (1986) reflect the widely spread practice in the industry of measuring the probability of the success of a merger by taking the fallback (failure) price to be the price before the deal was announced (an average over the past few weeks).

Samuelson and Rosenthal (1986) is the closest in spirit to our paper. They start with an empirical formula similar to our Equation (10), 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.[7] 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.

Barone-Adesi et al. (1994) point out that option prices should also be used in order to extract information about mergers. Subramanian (2004) discusses an arbitrage-free method to price options on stocks involved in mergers. He focuses mainly on stock-for-stock deals, since they allow a theoretically perfect correlation between the price of the acquirer and that

[7] 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.

of the target during the announcement period. In order to solve the model, he assumes that the fallback price follows a given basket of securities, and that the risk-neutral probability is determined by the arrival in a Poisson process with constant intensity.[8]

Hietala, Kaplan, and Robinson (2003) discuss the difficulty of information extraction around takeover contests, and estimate synergies and overpayment in the case of the 1994 takeover contest for Paramount in which Viacom overpaid by more than $2 billion.

In general, the literature on mergers has mostly been on the empirical side.[9] Several articles focus on the information contained in asset prices prior to mergers. Cao, Chen, and Griffin (2005) observe that option trading volume imbalances are informative prior to merger announcements, but not in general. From this, along the lines of Ross (1976), they deduce that option markets are important, especially when extreme informational events are pending. McDonald (1996) analyzed option prices on RJR Nabisco, which was the subject of a hostile takeover between October, 1988 and February, 1989, and noticed that there was a significant failure of the put–call parity during that time.

Mitchell and Pulvino (2001) survey the risk arbitrage industry and show that risk arbitrage returns are correlated with market returns in severely depreciating markets, but uncorrelated with market returns in flat and appreciating markets. This correlation shows that there is a positive merger risk premium.

There is also a recent related literature on credit risk and default rates. The similarity with our framework lies in that the underlying default is modeled as a process, and its estimation is central in pricing credit risk securities (see for example Duffie and Singleton (1997), Pan and Singleton (2005), Berndt et al. (2004)). Similar ideas, but involving earning announcements can be found in Dubinsky and Johannes (2005), who use options to extract information regarding earnings announcements.

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 empirical tests and the simulations of our model, and Section 4 concludes.

[8]According to this assumption, if the Poisson process has not jumped until the effective date, the merger is successful. This has the counter-factual implication that even deals that eventually fail become more likely as they approach the effective date.

[9]For a theoretical discussion about preemptive bidding, and an explanation of the offer premium or the choice between cash deals and stock deals, see Fishman (1988, 1989).

2 Model

2.1 Theory

Consider a merger in which one company, the acquirer, makes a cash offer to a second company, the target, in the amount of B_1 dollars per share. We assume that on some fixed future date T_e, which is called the "effective date," the uncertainty about the merger is resolved, and further that this date is known to all market participants and to the econometrician. If the deal succeeds, the target firm's shareholders receive B_1 per share on the effective date. Fluctuations in the price of the target company after the offer has been announced, but has not yet been accepted, come from two different sources. First, there are fluctuations in the probability of the merger being successful. Second, the "fundamental value" of the target firm may change for reasons unrelated to the takeover attempt.

At each time t define by $p_m = p_m(t)$ the price that the market would assign to a contract that pays 1 if the deal goes through and 0 if the deal fails.[10] Also, define the *fallback price* $B_2(t)$ as the value of the target company estimated by the market at t, conditional on the merger not being successful. Both B_2 and p_m are latent variables, assumed to be known to the market participants but unobservable by the econometrician. To allow for possible generalizations we assume that the offer B_1 is also stochastic, although we later analyze the case when B_1 is constant.

Let $W(t)$ be a 3-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. We assume that $B_1(t), B_2(t)$ are exponential diffusion processes:

$$B_i(t) \;=\; \mathrm{e}^{X_i(t)} \quad \text{with} \quad \mathrm{d}X_i(t) \;=\; \mu_i \;\mathrm{d}t + \sigma_i \;\mathrm{d}W_i(t), \quad i = 1, 2, \tag{1}$$

with constant drift μ_i and volatility σ_i. Also, p_m is an Itô process given by

$$\mathrm{d}p_m = \mu\big(p_m(t), t\big)\;\mathrm{d}t + \sigma\big(p_m(t), t\big)\;\mathrm{d}W_3(t), \tag{2}$$

where μ and σ satisfy some regularity conditions (see Duffie (2001)) and are such that p_m is

[10] If such a futures contract contingent on the success of the merger were traded on the Iowa Electronic Markets or Intrade, $p_m(t)$ would be the market price of this contract.

constrained to be between 0 and 1.[11] We assume that p_m is independent of B_1 and B_2.[12]

Denote by $\beta(t) = \mathrm{e}^{rt}$ the price of the bond (money market) at t. Define by Q the equivalent martingale measure associated to B_1, B_2, p_m. This is done as in Duffie (2001, Chapter 6), except that we want B_1, B_2, p_m to be Q-martingales after discounting by β. At each time t denote by $q = q(t)$ by

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

Then $q(t)$ is the risk-neutral probability of the state in which the merger succeeds. Because $p_m(t)$ is a discounted martingale with respect to Q, we have

$$\mathsf{E}_t^Q\left\{\frac{p_m(T_e)}{\beta(T_e)}\right\} = \frac{p_m(t)}{\beta(t)} \qquad \text{or equivalently} \qquad \mathsf{E}_t^Q\left\{q(T_e)\right\} = q(t). \tag{4}$$

We extend the probability space Ω on which Q is defined by including the binomial jump of p_m to either 1 or 0 with probability $p_m(T_e)$. This defines a new equivalent martingale measure Q' and a new filtration $\mathcal{F}'$. Denote by T_e' the instant after T_e at which we know whether p_m jumped to 1 or 0. Extend p_m as a stochastic process on $[0, T_e] \cup \{T_e'\}$ in the obvious way. Notice that $p_m = q$ at both T_e and T_e', so the payoff of stock B at T_e' can be written as

$$q(T_e')B_1(T_e') + (1 - q(T_e'))B_2(T_e'), \tag{5}$$

since $q(T_e')$ is either 1 or 0 depending on whether the merger was successful or not.

We now apply Theorem 6J in Duffie (2001) for redundant securities. Markets are complete because there are three Brownian motions and three securities B_1, B_2, p_m (plus a deterministic bond). Also at T_e the market, which displays only a binary uncertainty between T_e and T_e', 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. By the assumptions made above,

[11] We could require that on the effective date $p_m(T_e)$ is either 0 or 1, but we prefer the more general case when there is no such restriction. The intuition for this is that the market might not know the merger outcome even on the effective date, and so on that last day it assigns the probability $p_m(T_e)$. We also assume that at the end of day T_e the process p_m jumps either to 1 with probability $p_m(T_e)$, or to 0 with probability $1 - p_m(T_e)$, and that this jump is independent from the all the other sources of uncertainty.

[12] The case when p_m is correlated with B_2 is discussed after Theorem 1. Since later we treat B_1 as deterministic, we do not explicitly discuss the case when q and B_1 are correlated, but this can be solved as well using similar methods.

notice that this has B has final payoff equal to $q(T'_e)B_1(T'_e) + (1 - q(T'_e))B_2(T'_e)$. This allows us to establish the formula for $B(t)$ in the Theorem below.

Consider also a European call option on B with strike price K and maturity $T \geq T_e$. Define by $C(t)$ its price. Now define by $X_+ = \max\{X, 0\}$. Define $C_2(t)$ the price of a European call option on B_2 with strike price K and maturity T. Also, when B_1 is stochastic, define $C_1(t)$ the price of a European call option on B_1 with strike price K and maturity T_e. Under the assumption that the diffusion and volatility parameters are constant, the option price $C_2(t)$ satisfies the Black–Scholes formula:

$$C_2(t) = C^{BS}(B_2(t), K, r, T - t, \sigma_2) = B_2(t)N(d_1) - K\mathrm{e}^{-r(T-t)}N(d_2), \tag{6}$$

$$d_{1,2} = \frac{\log(B_2(t)/K) + (r \pm \frac{1}{2}\sigma_2^2)(T - t)}{\sigma_2\sqrt{T - t}}. \tag{7}$$

Theorem 1. *Assume q, B_1, B_2 satisfy the assumptions made above, with q is independent from B_1 and B_2. Then, if B_1 is stochastic the target stock price and option price satisfy*

$$B(t) = q(t)B_1(t) + (1 - q(t))B_2(t). \tag{8}$$

$$C(t) = q(t)C_1(t) + (1 - q(t))C_2(t). \tag{9}$$

If B_1 is constant the formulas become

$$B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1 - q(t))B_2(t). \tag{10}$$

$$C(t) = q(t)(B_1 - K)_+\mathrm{e}^{-r(T_e-t)} + (1 - q(t))C_2(t). \tag{11}$$

Proof. See the Appendix. □

When q and B_2 are correlated, one can still obtain similar results, but the formulas are more complicated.[13]

Now we study the Black–Scholes implied volatility curve under the hypothesis that our model is the correct one. The volatility curve plots the Black–Scholes implied volatility of

[13] To see where the difficulty comes from, suppose B_1 is constant. Let us consider the derivation of the formula for $B(t)$ in the proof of the Theorem: $B(t) = \frac{\beta(t)}{\beta(T_e)}\mathsf{E}_t^Q\Big\{q(T_e)B_1 + (1 - q(T_e))B_2(T_e)\Big\}$. The problem arises when attempting to calculate the integral $\mathsf{E}_t^Q\Big\{q(T_e)B_2(T_e)\Big\}$. This is in general a stochastic integral, but in particular cases, it can be reduced to an indefinite integral in two real variables.

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 in "at the money"), and then going up. This phenomenon is called the volatility "smile" or, when the curve is asymmetric, the volatility "smirk."

The next result shows that, in the case of options on cash mergers, the volatility smile arises naturally if the merger success probability is sufficiently high. Our model implies that the volatility curve is convex, with a kink at $K = B_1$. The magnitude of the kink (the difference between the slope of the curve on the right and left of $K = B_1$) equals the time discounted risk-neutral probability divided by the option vega. Recall the formula for d_2 in Equation (7) $d_2 = d_2(S, K, r, \tau, \sigma) = \frac{\log(S/K)+(r-\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$, with $\tau = T - t$. Denote by $\nu = \nu(S, K, r, \tau, \sigma) = \frac{\partial C}{\partial \sigma}$ the option vega; and by $\chi(\cdot)$ the indicator function: $\chi(x) = 1$ if $x > 0$ and $\chi(x) = 0$ otherwise.

Proposition 1. *If the offer price B_1 is constant, the slope of the implied volatility plot equals*

$$\frac{\partial \sigma^{\text{imp}}}{\partial K} = \frac{\mathrm{e}^{-r\tau}}{\nu(B,K,r,\tau,\sigma^{\text{imp}})}\Big(-q(t)\chi(B_1-K)-(1-q(t))N\big(d_2(B_2,K,r,\tau,\sigma_2)\big)+N\big(d_2(B,K,r,\tau,\sigma^{\text{imp}})\big)\Big),$$

where $\nu = \frac{\partial C}{\partial \sigma}$ is the option vega. For $q(t)$ sufficiently close to 1 the slope $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K\uparrow B_1}$ is negative and the slope $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K\downarrow B_1}$ is positive. The magnitude of the kink, i.e., the slope difference equals

$$\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K\downarrow B_1} - \left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K\uparrow B_1} = \frac{\mathrm{e}^{-r\tau}q(t)}{\nu(B,K,r,\tau,\sigma^{\text{imp}})}. \tag{12}$$

Proof. The formula for $\frac{\partial \sigma^{\text{imp}}}{\partial K}$ comes from differentiating with respect to K our option pricing formula for cash mergers: $C(t) = q(t)\mathrm{e}^{-r\tau}(B_1 - K)_+ + (1 - q(t))C^{BS}(B_2(t), K, r, \tau, \sigma_2) = C^{BS}(B, K, r, \tau, \sigma^{\text{imp}})$. This also implies the formula for the magnitude of the kink. Moreover, we get the following formula:

$$\left(\frac{\partial C}{\partial K}\right)_{K\downarrow B_1} - \left(\frac{\partial C}{\partial K}\right)_{K\uparrow B_1} = \mathrm{e}^{-r\tau}q(t). \tag{13}$$

Note that $\left(\frac{\partial\sigma^{\text{imp}}}{\partial K}\right)_{K\uparrow B_1}$ is proportional to $-q-(1-q)N(d_{2,B_2})+N(d_{2,B})$, which is negative for q sufficiently close to 1. Also, $\left(\frac{\partial\sigma^{\text{imp}}}{\partial K}\right)_{K\downarrow B_1}$ is proportional to $-(1-q)N(d_{2,B_2})+N(d_{2,B})$, which is positive for q sufficiently close to 1. □

In fact, one can check numerically that $\left(\frac{\partial\sigma^{\text{imp}}}{\partial K}\right)_{K\uparrow B_1}$ is negative and $\left(\frac{\partial\sigma^{\text{imp}}}{\partial K}\right)_{K\downarrow B_1}$ is positive for most of the relevant values of the parameters. This implies the usual convex shape for the volatility smile.

Now we prove a result about the instantaneous volatility that will be useful later. Define the instantaneous volatility of a positive Itô process $B(t)$ as the number $\sigma_{B}(t)$ that satisfies $\frac{\mathrm{d}B}{B}(t)=\mu_{B}(t)\,\mathrm{d}t+\sigma_{B}(t)\,\mathrm{d}W(t)$, where $W(t)$ is a standard Brownian motion. We compute the instantaneous volatility $\sigma_{B}(t)$ when the company B is the target of a cash merger.

Proposition 2. *Assume that the risk-neutral probability process follows the Itô process* $\frac{\mathrm{d}q}{q(1-q)}=\mu_1\,\mathrm{d}t+\sigma_1\,\mathrm{d}W_1(t)$. *The fallback price satisfies* $B_2(t)=\mathrm{e}^{X_2(t)}$, *with* $\mathrm{d}X_2=\mu_2\,\mathrm{d}t+\sigma_2\,\mathrm{d}W_2(t)$. *Assume that q and B_2 are independent and that B_1 is constant. Then the instantaneous volatility of B satisfies*

$$\left(\sigma_{B}(t)\right)^2 = \left(\frac{B_1\mathrm{e}^{-(T_e-t)}-B_2(t)}{B(t)}q(t)(1-q(t))\sigma_1\right)^2+\left(\frac{B_2(t)}{B(t)}(1-q(t))\sigma_2\right)^2 \tag{14}$$

Proof. Use Itô calculus to differentiate Equation 10 in Theorem 1. □

3 Empirical Results

3.1 Data

We study cash merger deals that were announced between January 1996 and June 2008, and have options traded on the target company. Merger data, e.g., company names, offer prices and effective dates, are from SDC Platinum, Thomson Reuters. 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.

During this period there are 7600 merger deals reported by SDC where the form of payment is exclusively cash. We restrict our sample to deals for which OptionMatrix has option prices on the target company. Since at the time of the analysis OptionMatrix only displays prices up to September 2008, we limit our sample to deals for which there is a resolution of the merger (success or failure) by this date. In other words, pending deals are excluded. We also exclude partial acquisitions: if the acquirer is wants to purchase less than 80% of the outstanding shares of the target company, the deal is excluded.

The resulting sample consists of 586 deals. Although cash is the most common type of payment when public companies are acquired, the significant reduction in sample size shows that most of these companies are usually small and have no options traded on their stock. Out of these 586 deals, 465 successfully completed the merger, while 121 failed to reach an agreement by the effective date.[14]

Table 1 reports some summary statistics for our initial sample. For example, the median deal lasted 84 trading days (until either it succeeded or failed). The average deal duration is 100 days, with 581 days being the longest. Various percentiles for the offer premium are also reported in Table 1. The offer premium is the percentage difference between the (cash) offer price, and the target company stock price on the day before the initial merger announcement. The median offer premium in our sample is 25%, while the mean is 31% and the standard deviation is 30%. The table also reports how often options are traded on the target company. The median percentage of trading days where there exists at least an option with positive trading volume is 28.57%, indicating that options are quite illiquid.

Since in order to obtain the pricing equations (10) and (11) we assume that the option matures after the effective date ($T > T_e$), we restrict the sample to include only deals for which there exist traded options which mature after the effective date. Next, we perform various checks to spot various data problems: missing underlying prices; prices which are inconsistent between OptionMatrix and CRSP; misreported offer values that did not include additional payments like special dividends; and missing offer values. We also exclude deals for which the total duration is less than 6 days. The resulting sample has 422 deals.

As mentioned earlier, options on the target companies of the deals selected are usually thinly traded. In order for our estimation procedure to work, we need to impose the requirement that there are enough options traded daily on each stock. For each stock i and

[14]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. An effective date is filed with the SEC at the time when the initial cash offer is made, but this date may subsequently change and in fact it often does.

each day t consider the number of options traded on that day which with positive trading volume. (An option can have quotes—bid and ask prices posted by the market maker—but zero trading volume.) Denote that number by $N_{i,t}$. Then for each stock i we define the mean of $N_{i,t}$ over time to be μ_i^N and the standard deviation over time σ_i^N. The intuition is that we want a high average number of options traded per day, but we do not want the number of options to vary too much, so we put a penalty if it varies. We then select the deals for which $\mu_i^N - 0.5\,\sigma_i^N > 0.9$.[15] Our final sample consists of 282 deals, out of which 246 succeeded and 36 failed. Table 2 reports the most liquid 5 successful deals and 5 failed deals in our sample, ranked by the liquidity measure mentioned above, the adjusted average number of traded options per day $\mu_i^N - 0.5\,\sigma_i^N$.

We use closing daily prices for the target stocks, and the closing bid and ask prices for the option prices. We only consider call options with maturities longer than the effective date of the deal. For deals that are successful by the effective date, 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.[16]

3.2 Methodology

Consider our sample of 282 cash merger deals with options traded on the target company. Start with the observed variables: (i) T_e, the effective date of the deal (measured as the number of trading days from the announcement $t = 0$); (ii) r, the risk-free interest rate, assumed constant throughout the deal; (iii) B_1, the cash offer price; (iv) $B(t)$, the stock price of the target company on day t; (v) $C(t)$, the price of a call option traded on B with a strike price of K; this is selected to have the maximum trading volume on that day.[17]

The latent variables in this model are $q(t)$, the risk-neutral probability that the merger is successful; and $B_2(t)$, the fallback price, i.e., the price of the target company if the deal fails.

[15] We choose the penalty slope 0.5 so that it is not too restrictive and we get enough deals. We choose the cutoff 0.9 so that the number of failed deals in our final sample is large enough (36). If we require that $\mu_i^N - 0.5\,\sigma_i^N > 1$ instead, the number of failed deals decreases from 36 to 20.

[16] This procedure is stipulated in the Options Clearing Corporation (OCC) By-Laws and Rules (Article VI, Section 11).

[17] If all option trading volumes are zero on that day, use the option with the strike K closest to the strike price for the most currently traded option with maximum volume.

The variables q and B_2 satisfy:

$$q \;=\; X_1(t) \quad \text{with} \quad \frac{\mathrm{d}X_1}{X_1(1-X_1)} \;=\; \mu_1\,\mathrm{d}t + \sigma_1\,\mathrm{d}W_1(t); \tag{15}$$

$$B_2(t) \;=\; \mathrm{e}^{X_2(t)} \quad \text{with} \quad \mathrm{d}X_2 \;=\; \mu_2\;\mathrm{d}t + \sigma_2\;\mathrm{d}W_2(t). \tag{16}$$

We assume that $\mathrm{d}W_1(t)$ and $\mathrm{d}W_2(t)$ are independent, which implies that q and B_2 are independent.

The parametrization of q is very similar to the Black–Scholes specification for the underlying price ($\mathrm{d}S/S = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$), except that the equation for q also has a term $1-q$ in the denominator, which ensures that q stays lower than 1. With this parametrization the drift μ_1 has a particularly useful interpretation in relation to the merger risk premium. 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:

$$\mathsf{E}_t\left(\frac{\mathrm{d}p_m}{p_m}\right) - r\,\mathrm{d}t \;=\; \mathsf{E}_t\left(\frac{\mathrm{d}q}{q}\right) \;=\; (1-q)\mu_1\,\mathrm{d}t. \tag{17}$$

Assume that equations (10) and (11) from Theorem 1 hold only approximately, with errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$:

$$B(t) \;=\; q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t) + \varepsilon_B(t), \tag{18}$$

$$C(t) \;=\; q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C^{BS}(B_2(t),K,r,T-t,\sigma_2) + \varepsilon_C(t). \tag{19}$$

The errors are IID bivariate normal:

$$\begin{bmatrix} \varepsilon_B(t) \\ \varepsilon_C(t) \end{bmatrix} \sim N(0,\Sigma_\varepsilon), \quad \text{where} \quad \Sigma_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}. \tag{20}$$

Equations (15), (16), (18) and (19) define a state space model with observables $B(t)$ and $C(t)$, latent (state) variables $q(t)$ and $B_2(t)$, and model parameters $\mu_1, \sigma_1\mu_2, \sigma_2, \sigma_{\varepsilon,B}, \sigma_{\varepsilon,C}$. We adopt a Bayesian approach and conduct inference by sampling from the joint posterior density of state variables and model parameters given the observables. We do this using a Markov Chain Monte Carlo (MCMC) method based on a state space representation of our model. In this framework, the state equations (15) and (16) specify the dynamics of latent variables, while the pricing equations (18) and (19) specify the relationship between the latent variables and the observables. The addition of errors ε_B and ε_C 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 discussion see, e.g., Johannes and Polson (2003), Koop (2003). The resulting estimation procedure is described in detail in Appendix B.[18] The priors used in our estimation are all flat, except for the case of σ_2, for which the prior has a very diffuse inverse gamma distribution.

To illustrate our methodology, we select a specific deal corresponding to the most liquid company in our sample, AWE. (See Table 2.) Then Figure 4 displays the histograms of the MCMC posterior draws for: the latent variables at half the effective date $(X_1\left(\frac{T_e}{2}\right), X_2\left(\frac{T_e}{2}\right))$, and the model parameters (μ_1, σ_1, μ_2, σ_2, $\sigma_{\varepsilon,B}$, and $\sigma_{\varepsilon,C}$).[19]

3.3 Results

As described in the data section, our sample contains 282 cash mergers during 1996–2008 with sufficiently liquid options traded on the target company. Recall that our estimation method assumes that the pricing formulas (18) and (19) 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 3 reports percentiles (computed over the cross section of firms) of the time series average pricing error $\frac{1}{T_e}\sum_{t=1}^{T_e}\left|\frac{\hat{B}(t)-B(t)}{B(t)}\right|$ for the target company stock price. The errors are very small, with a median error of only 5 basis points.

According to our liquidity measure, the adjusted average number of traded options per day, the most liquid company in our sample is AT&T Wireless (AWE) (see Table 2 for more

[18]As noted by Johannes and Polson (2003), equations of the type (18) or (19) are a non-linear filter. The problem is that it is quite hard to do the estimation using the actual filter. MCMC is a much cleaner estimation technique, but it does smoothing, because it uses all the data at once.

[19]The draws are considered only after the initial "burn-out" period, which in this case occurs after approximately 200,000 iterations.

details). Figure 5 shows that in the case of company AWE our model fits the call option prices, including the kink in the implied volatility curve. This is remarkable: 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.

Table 4 reports option pricing errors from four models. The first model is the one described in this paper, denoted "MRB" for short, and the other three models are versions of the Black–Scholes formula, with the volatility parameter estimated in there different ways. The first version ("BS1") uses an average of the Black–Scholes implied volatilities for the ATM call options over the duration of the deal. The second version ("BS2") uses the implied volatility for the previous-day ATM call option. The third version ("BS3") uses the implied volatility for the previous-day call option with the closest strike price to the option being priced. Note that BS3 sets a relatively high bar, as it uses the previous day's realization of the volatility smile to predict current option prices.

The table reports three types of errors: Panel A the percentage errors, Panel B the absolute errors, and Panel C the absolute errors divided by the bid-ask spread of the call option. (Panels D and E report the percentage bid-ask spread and absolute bid-ask spread of the call options, respectively.) Each type of error is computed by restricting the sample of call options based on the moneyness of the option, i.e., the ratio of the strike price K to the underlying stock price $B(t)$: (1) all call options; (2) near-in-the-money (Near ITM) calls, with $K/B \in [0.95, 1.00]$; (3) near-out-of-the-money (Near OTM) calls, with $K/B \in [1.00, 1.05]$; (4) in-the-money (ITM) calls, with $K/B \in [0.90, 0.95]$; (5) out-of-the-money (OTM) calls, with $K/B \in [1.05, 1.1]$; (6) deep-ITM calls, with $K/B < 0.90$; and (7) deep-OTM calls, with $K/B > 1.10$. 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 of the individual stocks in our sample.

To understand how the pricing errors are computed, consider, e.g., the results of Table 4, Panel A, fifth group. These are OTM calls. From the Table, we see that there are only 231 stocks for which the set of such options is non-empty. Then, for one of these stocks and for a call option $C(t)$ traded on day t on a stock $B(t)$ and with strike K, compute the 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 MRB, BS1, BS2, or BS3. Next, take the average error over this particular group of options (using equal weights). The Table then reports the 5-th, 25-th, 50-th, 75-th, and 95-th percentiles over the 231 corresponding stocks.

Overall, our model (MRB) does significantly better than both BS1 and BS2, where we use at-the-money implied volatilities. For example, in Panel A we see that, for all call options, the median percentage pricing error is 26.06% for the MRB model, with 33.02% for BS1 and 34.22% for BS2. As mentioned above, model BS3 is hard to surpass, and indeed it does better than our model: the median error is 18.19%. However, the MRB model does better in terms of the *absolute* pricing error (see Panel B): the median absolute error is 9.99% for the MRB model, compared with 12.10% for BS1, 12.04% for BS2, and 11.11% for BS3. The exception is for OTM calls, where BS3 does better than our model.

Panel C of Table 4 reports the ratio between the absolute pricing error and the bid-ask spread, which for the median stock in our sample is less than 0.5 for each moneyness. This indicates that for the median stock the profit is smaller than the bid-ask spread. However, some particular deep-OTM call options have a ratio larger than one, indicating that in that case one could devise a profitable trading strategy. Even in that case, the bid-ask spread represents only a part of the costs. The depth at the bid and ask for the deep-OTM options is very small, so price impact would prevent an arbitrageur from correcting the mispricing.

We test the implications of the model regarding the kink in the volatility smile, which can be observed in a particular case in Figure 5. Proposition 1 shows that this kink corresponds to a kink in the plot of the call option price against the strike price. Moreover, it shows that this kink, normalized by the time discount coefficient, should be equal to the risk-neutral probability q. Empirically, if we do an OLS regression of the normalized kink on the estimated risk-neutral probability, we should find that the intercept equals 0 and the slope equals 1. Table 5 shows that this is indeed the case.

In addition to pricing options, we also check whether the estimates of state variables recovered using our model are economically meaningful. We begin by asking whether estimated success probabilities, $\hat{q}$, predict the outcomes of deals in the sample. Figures 1 illustrates the results for the ten most liquid deals from Table 2, five of which succeeded, and five of which failed. Figures 1 and 2 display the time series of posterior means 90% credibility intervals (i.e., the 5% and 95% quantiles of the posterior) for the time series of the state variables $q(t)$

and $B_2(t)$. 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.[20]

Table 6 shows that in general $\hat{q}$ predicts well the outcome of the deal. We choose 10 evenly spaced days during the period of the merger deal: for $n = 1, \ldots, 10$, choose t_n as the closest integer strictly smaller than $n\frac{T_e}{10}$. The Table reports the pseudo-R^2 for 10 cross-sectional probit regressions of the deal outcome (1 if successful, 0 if it failed) on $\hat{q}(t_n)$. Notice that R^2 increases approximately from 10% to about 47%, which indicates that the success probability better predicts the outcome of the merger the closer one comes to the effective date. Note that we do not impose the success probability to be 0 or 1 at the effective date T_e. This would likely lead to an even better fit.

We contrast our model-implied risk-neutral probability to the "naive" method of Brown and Raymond (1986), which is used widely in the merger literature. This is defined by considering the current price $B(t)$ of the target company. If this is close to the offer price B_1, the naive probability is high. If instead $B(t)$ is close to the pre-announcement stock price $B_0(t)$, then the naive probability is low. Specifically define $q_{\text{naive}}(t) = \frac{B(t)-B_0}{B_1-B_0}$ if $B_0 < B(t) < B_1$. If $B(t) < B_0$ (or $> B_1$), q_{naive} is set equal to zero (one). Table 6 reports the results from a cross-sectional probit regression of the deal outcome on $q_{\text{naive}}(t_n)$. R^2 increases from 0% to 28%, indicating that our model does a better job at predicting the deal outcome than the "naive" one.

We also investigate how the fallback price $B_2(t)$ compares to the price B_0 before the announcement. One might expect that B_2 should be on average higher than B_0. This might be true because a merger is usually a good signal about the target company, e.g., it might indicate that other tender offers have become more likely. Table 7 reports the results from regressing $\ln(B_2(t_n))$ on $\ln(B_0(t_n)$. The slope is very close to one, as expected, and the intercept indicates that the fallback price is on average 20–30% higher than the pre-announcement price.

Our model has implications for measurement of the volatility of merger target firms. Proposition 2 shows that the volatility of target company B is given by $\sigma_B^2(t) = \left(\frac{B_2}{B}(1-q)\sigma_2\right)^2 + \left(\frac{B_1 e^{-(T_e-t)} - B_2}{B} q(1-q)\sigma_1\right)^2$. Notice that when the deal is close to completion, the success probability q is close to 1 and so the model-implied probability σ_B is close to 0. This explains the

[20]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 option 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.

empirical fact when a merger is close to completion, the Black–Scholes implied volatility of the target company converges to 0.

Finally, we explore the possibility to estimate the merger risk premium using the drift coefficient in the diffusion process for the success probability (15). According to Equation (17), 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 158%, with a standard deviation of 29%. The mean seems very high, but comparable figures for cash mergers have been reported in the literature.[21]

4 Conclusions

We propose an arbitrage-free option pricing formula on companies that are subject to takeover attempts. We use the formula to estimate several variables of interest in a cash merger: the success probability and the fallback price. The option formula does significantly better than the standard Black–Scholes formula, and produces results comparable to a modified Black–Scholes formula which estimates the volatility using the previous-day implied volatility for the same strike price. As a consequence, our model produces a volatility smile close to the one observed in practice, and goes some distance towards explaining the volatility smile when the underlying stock price is exposed to a significant binary event.

One implication of our theoretical model is the existence of a kink in the implied volatility curve near the money for mergers which are close to being successful. It can be shown that the magnitude of the kink equals the time discounted risk-neutral version of the success probability divided by the option vega. Empirically, we show that indeed a larger estimated risk-neutral probability is correlated with a bigger kink in the implied volatility curve.

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

[21] See, e.g., Dukes, Frolich and Ma (1992), who report an average daily premium of 0.47%, 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.

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 estimated average merger risk premium over our sample is 158% annually, which is a large figure, although consistent with the cash mergers literature.

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

A Proofs

Proof of Theorem 1:

By the independence of q and B_1, B_2, we have: $\frac{B(t)}{\beta(t)} = \mathsf{E}_t^{Q'}\left\{\frac{q(T_e')B_1(T_e')+(1-q(T_e'))B_2(T_e')}{\beta(T_e)}\right\} = \mathsf{E}_t^{Q}\left\{\frac{q(T_e)B_1(T_e)+(1-q(T_e))B_2(T_e)}{\beta(T_e)}\right\} = \mathsf{E}_t^{Q}\left\{q(T_e)\frac{B_1(T_e)}{\beta(T_e)} + (1-q(T_e))\frac{B_2(T_e)}{\beta(T_e)}\right\} = q(t)\frac{B_1(t)}{\beta(t)} + (1-q(t))\frac{B_2(t)}{\beta(t)}$. This implies, when B_1 is stochastic, that $B(t) = q(t)B_1(t) + (1-q(t))B_2(t)$.

When B_1 is constant, the formula is different: $\frac{B(t)}{\beta(t)} = \mathsf{E}_t^{Q}\left\{q(T_e)\frac{B_1}{\beta(T_e)} + (1-q(T_e))\frac{B_2(T_e)}{\beta(T_e)}\right\} = q(t)\frac{B_1}{\beta(T_e)} + (1-q(t))\frac{B_2(t)}{\beta(t)}$. This implies $B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t)$.

Recall that $C(t)$ is the price of a European call option on B with strike price K and maturity $T \geq T_e$. When B_1 is stochastic, it satisfies

$$\begin{aligned}\frac{C(t)}{\beta(t)} &= \mathsf{E}_t^{Q}\left\{q(T_e)\frac{(B_1(T_e)-K)_+}{\beta(T_e)} + (1-q(T_e))\mathsf{E}_{T_e}^{Q}\left\{\frac{(B_2(T)-K)_+}{\beta(T)}\right\}\right\}\\ &= q(t)\frac{C_1(t)}{\beta(t)} + (1-q(t))\frac{C_2(t)}{\beta(t)}.\end{aligned}$$

This implies $C(t) = q(t)\,C_1(t) + (1-q(t))\,C_2(t)$. Notice that C_1 and C_2 expire at different maturities (T_e and T, respectively).

When B_1 is constant, the formula is:

$$
\begin{aligned}
\frac{C_(t)}{\beta(t)} &= \mathsf{E}_t^Q\left\{q(T_e)\frac{(B_1-K)_+}{\beta(T_e)} + (1-q(T_e))\mathsf{E}_{T_e}^Q\left\{\frac{(B_2(T)-K)_+}{\beta(T)}\right\}\right\} \\
&= q(t)\frac{(B_1-K)_+}{\beta(T_e)} + (1-q(t))\frac{C_2(t)}{\beta(t)}.
\end{aligned}
$$

This implies $C(t) \;=\; q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C_2(t)$. □

B An MCMC Procedure for Mergers

Recall that for target companies in cash mergers the latent variables are $q(t)$ and $B_2(t)$, and the observed variables are $B(t)$ and $C(t)$. These variables are connected by Equations (18) and (19): $B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t) + \varepsilon_B(t)$, $C(t) = q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C^{BS}_{K,\sigma_2,r,T}(B_2(t),t) + \varepsilon_C(t)$. The errors ε_B and ε_C are IID bivariate normal and independent. Define the state variables X_1 and X_2 as Itô processes with constant drift and volatility

$$
\mathrm{d}X_{i,t} \;=\; \mu_i\,\mathrm{d}t + \sigma_i\,\mathrm{d}W_t^{(i)}, \quad \text{with} \quad i = 1,2. \tag{21}
$$

The risk-neutral probability q and the fallback price B_2 are defined by $q(t) = \frac{\mathrm{e}^{X_1(t)}}{1+\mathrm{e}^{X_1(t)}}$.[22]

Define the observed variables Y_B and Y_C by $Y_{B,t} = B(t)$, $Y_{C,t} = C(t)$. Also define the deterministic functions of t: $\bar{B}_t = B_1\mathrm{e}^{-r(T_e-t)}$, $\bar{C}_t = (B_1-K)_+\mathrm{e}^{-r(T_e-t)}$. With this notation, define the functions $f_{B,t}(x_1,x_2) = \frac{\exp(x_1)}{1+\exp(x_1)}\bar{B}_t + \frac{1}{1+\exp(x_1)}\exp(x_2)$, $f_{C,t}(x_1,x_2) = \frac{\exp(x_1)}{1+\exp(x_1)}\bar{C}_t + \frac{1}{1+\exp(x_1)}C^{BS}_{K,\sigma_2,r,T}(\exp(x_2),t)$. To make the dependence of $f_{C,t}$ on σ_2 explicit, sometimes we write $f_{C,t}(x_1,x_2) = f_{C,t}(x_1,x_2|\sigma_2)$. Notice that with the new notation Equations (18) and (19) become

$$
\begin{cases}
Y_{B,t} = f_{B,t}(X_{1,t},X_{2,t}) + \varepsilon_{B,t}, \\
Y_{C,t} = f_{C,t}(X_{1,t},X_{2,t}) + \varepsilon_{C,t}.
\end{cases} \tag{22}
$$

Define also $Z_{1,t} = X_{1,t} - X_{1,t-1}$, $Z_{2,t} = X_{2,t} - X_{2,t-1}$. The vector of parameters is

$$
\theta = \begin{bmatrix} \mu_1 & \mu_2 & \sigma_1 & \sigma_2 \end{bmatrix}^\top. \tag{23}
$$

[22]Note that the specification we choose here is not the same as in Equation (15), which is used in our empirical study. This is done in order to simplify the presentation.

The MCMC strategy is to sample from the posterior distribution with density $p(\theta, X, \Sigma_\varepsilon \mid Y)$, 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, one gets: $p(\theta, X, \Sigma_\varepsilon \mid Y) \ \propto\ p(Y \mid X, \Sigma_\varepsilon, \theta) \cdot p(X \mid \theta) \cdot p(\Sigma_\varepsilon) \cdot p(\theta)$. On the right hand side, the first term in the product is the likelihood for the observation equation (22); the second term is the likelihood for the state equation (21); and the third term is the prior density of the parameter θ. Define by $\phi(x \mid \mu, \Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$ the density of the n-dimensional multivariate normal density with mean μ and covariance matrix Σ. Then we have the following formulas:

$$p(Y \mid X, \Sigma_\varepsilon, \theta) = \prod_{t=1}^{T} \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t}(X_{1,t}, X_{2,t}) \\ f_{C,t}(X_{1,t}, X_{2,t}) \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}\right); \tag{24}$$

$$p(X \mid \theta) \ = p(X_{1,1} \mid \theta) \cdot p(X_{2,1} \mid \theta) \cdot \prod_{t=2}^{T} \phi\left(\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma^2_1 & 0 \\ 0 & \sigma^2_2 \end{bmatrix}\right). \tag{25}$$

Recall that $Z_{k,t} = X_{k,t} - X_{k,t-1}$ for $k = 1, 2$.

We now start 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 $p(\Sigma_\varepsilon \mid \theta^{(i)}, X^{(i)}, Y)$. Notice that with a flat prior for Σ_ε, one has

$$p(\Sigma_\varepsilon \mid \theta^{(i)}, X^{(i)}, Y) \ \ \propto \ \prod_{t=1}^{T} \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t} \\ f_{C,t} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}\right),$$

where $f_{j,t} = f_{j,t}\big(X^{(i)}_{1,t}, X^{(i)}_{2,t}\big)$, with $j = B, C$. This implies that $\left(\sigma^{(i+1)}_{\varepsilon,j}\right)^2, j = B, C$ is sampled from an inverted gamma-2 distribution: $\big(\sigma^{(i+1)}_{\varepsilon,j}\big)^2 \ \sim\ IG_2\Big(\sum_{t=1}^{T}(Y_{j,t} - f_{j,t})^2, T-1\Big)$, $j = B, C$. The inverted gamma-2 distribution $IG_2(s, \nu)$ has log-density $\log p_{IG_2}(x) \ = \ -\frac{\nu+1}{2}\log(x) - \frac{s}{2x}$. 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 $p(X \mid \theta^{(i)}, \Sigma_\varepsilon^{(i+1)}, Y)$. To simplify notation, denote by $\theta = \theta^{(i)}$, and $\Sigma_\varepsilon = \Sigma_\varepsilon^{(i+1)}$. Notice that $p(X \mid \theta, \Sigma_\varepsilon, Y) \ \propto\ p(Y \mid \theta, \Sigma_\varepsilon, X) \cdot p(X \mid \theta)$, assuming flat priors for X. Then, if $t = 2, \ldots, T-1$,

$$
\begin{aligned}
p(X_t \mid \theta, \Sigma_\varepsilon, Y) \;\propto\; & \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t}(X_{1,t}, X_{2,t}) \\ f_{C,t}(X_{1,t}, X_{2,t}) \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}\right) \\
& \cdot \phi\left(\begin{bmatrix} X_{1,t} - X^{(i+1)}_{1,t-1} \\ X_{2,t} - X^{(i+1)}_{2,t-1} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma^2_1 & 0 \\ 0 & \sigma^2_2 \end{bmatrix}\right) \qquad (26) \\
& \cdot \phi\left(\begin{bmatrix} X^{(i)}_{1,t+1} - X_{1,t} \\ X^{(i)}_{2,t+1} - X_{2,t} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma^2_1 & 0 \\ 0 & \sigma^2_2 \end{bmatrix}\right).
\end{aligned}
$$

If $t = 1$, replace the second term in the product with $p(X_1 \mid \theta)$; and if $t = T$, drop the third term out of the product. This is a non-standard density, so we have to perform the Metropolis–Hastings algorithm to sample from this distribution. This algorithm will be described later.

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

$$
\begin{aligned}
p(\theta \mid X, \Sigma_\varepsilon, Y) \;\propto\; & \prod_{t=1}^{T} \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t}(X_{1,t}, X_{2,t}) \\ f_{C,t}(X_{1,t}, X_{2,t} \mid \sigma_2) \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}\right) \\
& \cdot \prod_{t=2}^{T} \phi\left(\begin{bmatrix} X_{1,t} - X_{1,t-1} \\ X_{2,t} - X_{2,t-1} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma^2_1 & 0 \\ 0 & \sigma^2_2 \end{bmatrix}\right). \qquad (27)
\end{aligned}
$$

Notice that the first term appears in the product only because $f_{C,t}(X_{1,t}, X_{2,t}|\sigma_2)$ depends on σ_2. For the other parameters (μ_1, μ_2, and σ_1) we can drop this term. Since we denoted by $Z_{k,t} = X_{k,t} - X_{k,t-1}$ for $k = 1, 2$, we have the following updates: $\mu^{(i+1)}_k \sim N\Big(\frac{1}{T-1} \sum_{t=2}^{T} Z_{k,t}, \frac{\left(\sigma^{(i)}_k\right)^2}{T-1}\Big)$, $k = 1, 2$; $\left(\sigma^{(i+1)}_1\right)^2 \sim IG_2\Big(\sum_{t=2}^{T}\Big(Z_{1,t} - \mu^{(i+1)}_1\Big)^2, T-2\Big)$. For σ_2 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 $X_{PROP} \sim q(x \mid X_{CURR})$ (this is the "proposed" X).

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

(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 $q(y \mid x) = \phi(x \mid 0, a^2)$, for some positive value of a. Equivalently, $X_{PROP} = X_{CURR} + e$, where $e \sim N(0, a^2)$.

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

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

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

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

[7] Cao, Charles, Zhiwu Chen, and John M. Griffin (2005): Informational Content of Option Volume Prior to Takeovers, *Journal of Business*, forthcoming.

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

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

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

[11] 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.

[12] 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.

[13] 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.

[14] Fishman, Michael J. (1988): A Theory of Preemptive Takeover Bidding, *RAND Journal of Economics*, **19**, 88–101.

[15] Fishman, Michael J. (1989): Preemptive Bidding and the Role of the Medium of Exchange in Acquisitions, *Journal of Finance*, **44**, 41–57.

[16] 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.

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

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

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

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

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

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

[23] McDonald, Robert L. (1996): *Speculating on an Acquisition with Options: RJR Nabisco*, case study, South-Western College Publishing.

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

[25] Pan, Jun, and Kenneth J. Singleton (2005): Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads; working paper.

[26] Ross, Stephen A. (1976): Options and Efficiency, *Quarterly Journal of Economics*, **90**, 75–89.

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

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

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

Figure 1: Estimates of the risk-neutral success probability $q(t)$ for a subsample of ten cash mergers described in Table 2. The deals corresponding to target tickers AWE, DSP, GP, MLNM, PLAT succeeded, while those for CSC, GMSTF, MCIC, TTWO, UCL failed. The dash-dotted lines represent the 5% and 95% error bands around the estimated median values. These estimates in the model are obtained using all the options offered with positive trading volume.

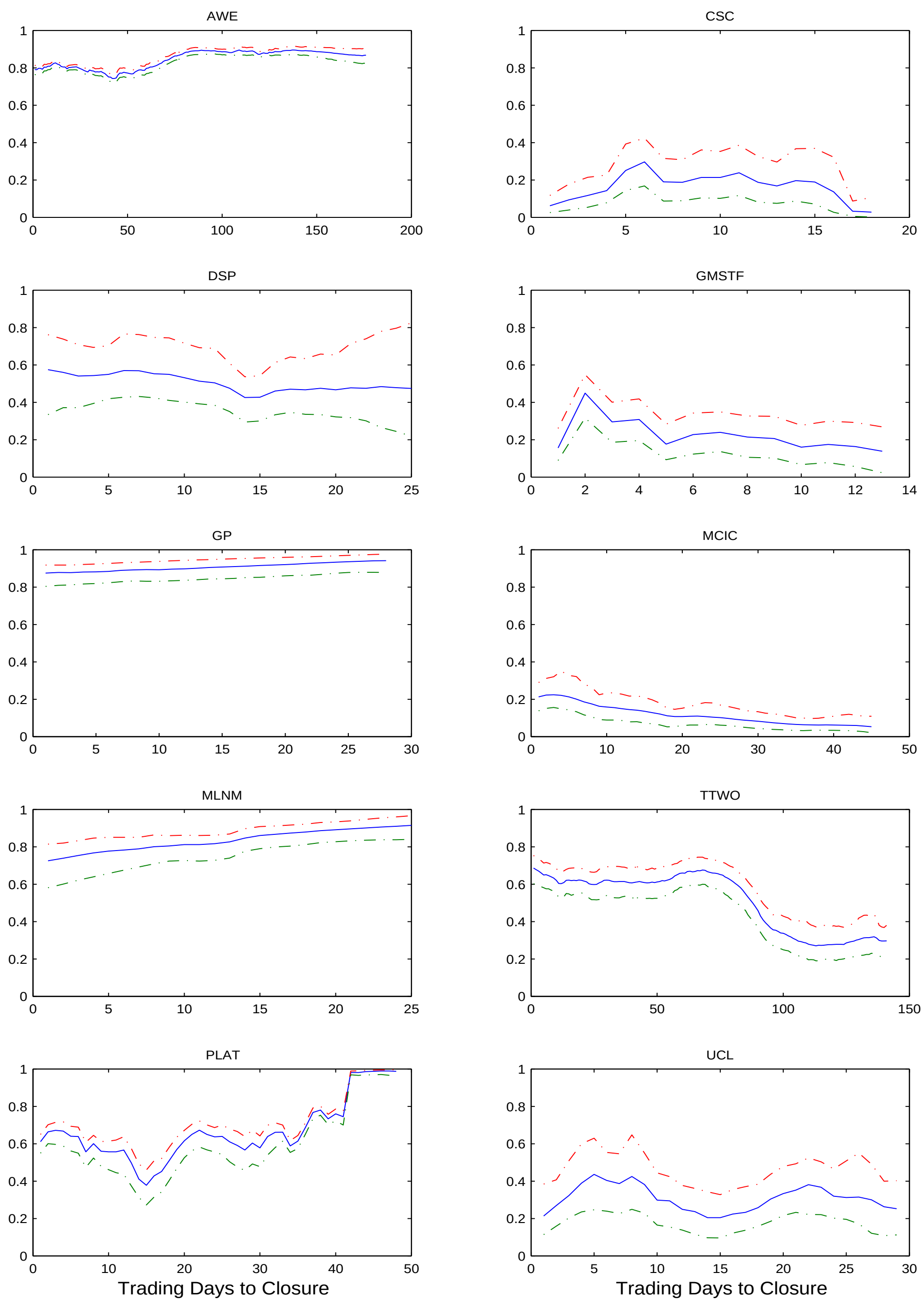


Figure 2: Estimates of the fallback prices of the target stock $B_2(t)$ for a subsample of ten cash mergers described in Table 2. The deals corresponding to target tickers AWE, DSP, GP, MLNM, PLAT succeeded, while those for CSC, GMSTF, MCIC, TTWO, UCL failed. The dash-dotted lines represent the 5% and 95% error bands around the estimated median values. These estimates in the model are obtained using all the options offered with positive trading volume.

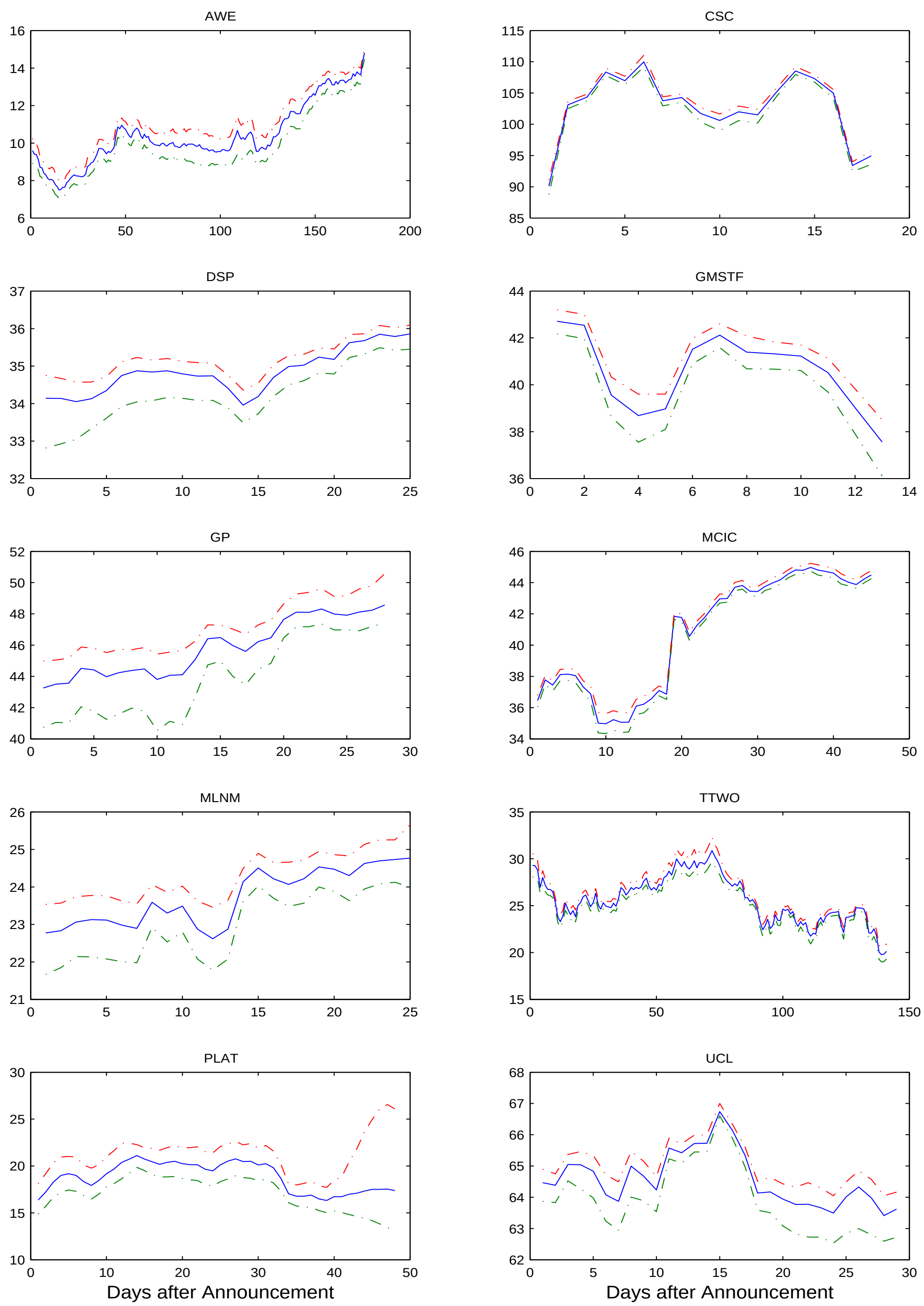


Figure 3: Consider the deal described in Table 2 corresponding to the target company AWE. This figure plots: (i) the offer price, discounted at the current interest rate, using a dashed-dotted line; (ii) the stock price, using a continuous line; and (iii) the estimated fallback price (the price of the target company if the deal fails), using a dashed line.

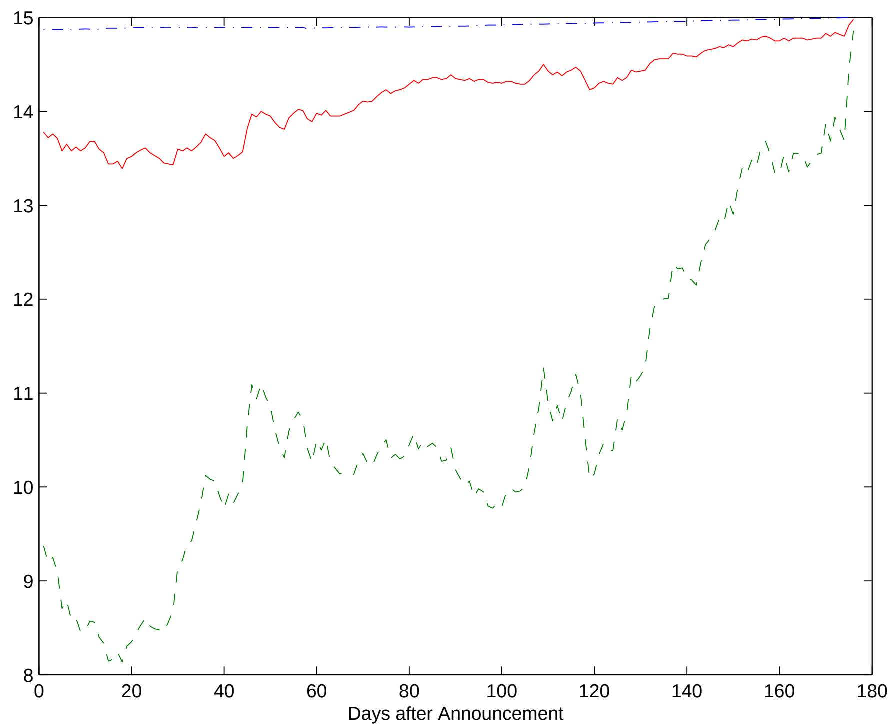


Figure 4: Consider the most liquid deal described in Table 2 for which the merger succeeded (where the target company is AWE). This figure plots the MCMC draws for a few latent variables, parameters, and model errors. Recall the chosen parametrization for the risk-neutral probability $q(t) = X_1(t)$: $\frac{\mathrm{d}q}{q(1-q)} = \mu_1 \,\mathrm{d}t + \sigma_1 \,\mathrm{d}W_1$, and for the fallback price $B_2(t) = \mathrm{e}^{X_2(t)}$: $\mathrm{d}X_2 = \mu_2 \,\mathrm{d}t + \sigma_2 \,\mathrm{d}W_2$. Recall also the model errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$ are assumed to have constant standard deviations $\sigma_{\varepsilon,B}$ and $\sigma_{\varepsilon,C}$, respectively. The figure plots the histogram of the 200,000 to 400,000 draws for: (i) X_1 at $t = \frac{T_e}{2}$, where $T_e = 176$ is the number of trading days for which the deal is ongoing; (ii) X_2 at $t = \frac{T_e}{2}$; (iii–iv) the drift parameters μ_1 and μ_2; (v–vi) the volatility parameters σ_1 and σ_2; (vii–viii) the model error standard deviations $\sigma_{\varepsilon,B}$ and $\sigma_{\varepsilon,C}$. All reported parameter values are annualized.

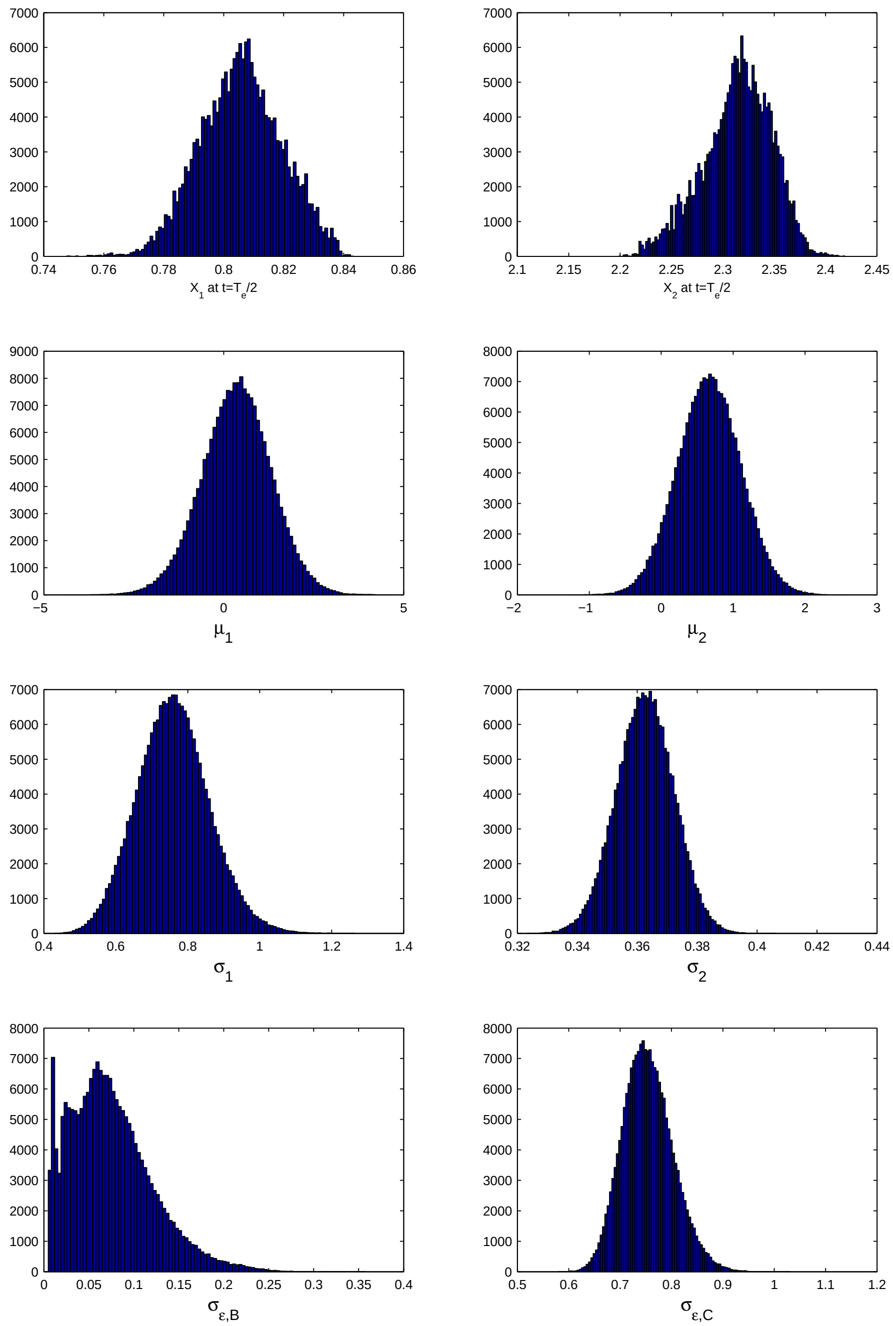
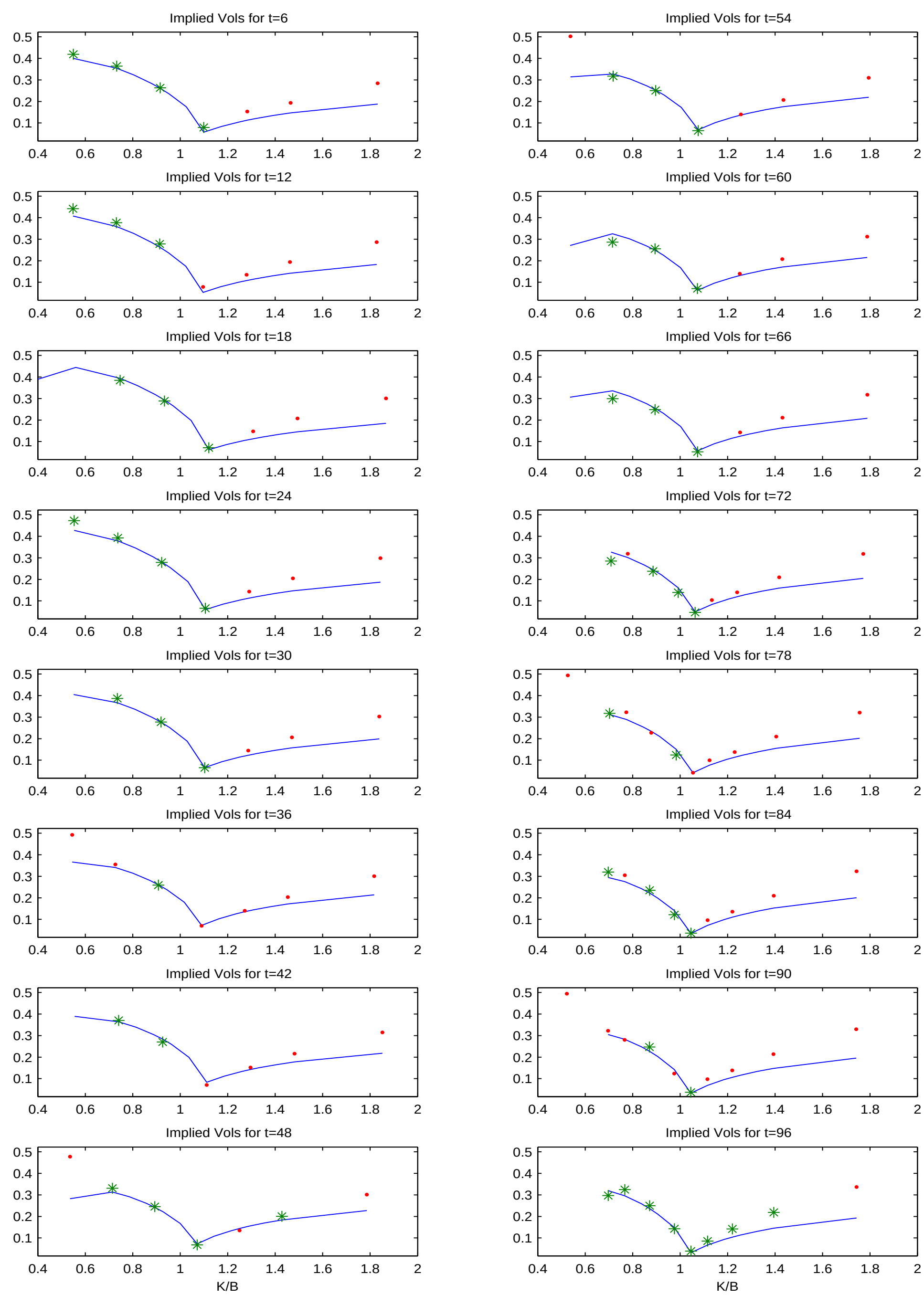


Figure 5: This compares the observed volatility smile with the theoretical volatility smile in the case of AWE. For every other sixth day while the deal is ongoing and there are more than six options offered, plot the option's Black–Scholes implied volatility against its moneyness (the ratio of strike price K to the underlying price $B(t)$). The Black–Scholes implied volatility for the observed option price is plotted 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 (for which the price is taken as the mid-point between the bid and ask). The Black–Scholes implied volatility for our theoretical option price is plotted using a continuous solid line. The parameters in the theoretical option price are estimated using only the option with the highest volume traded in each day.



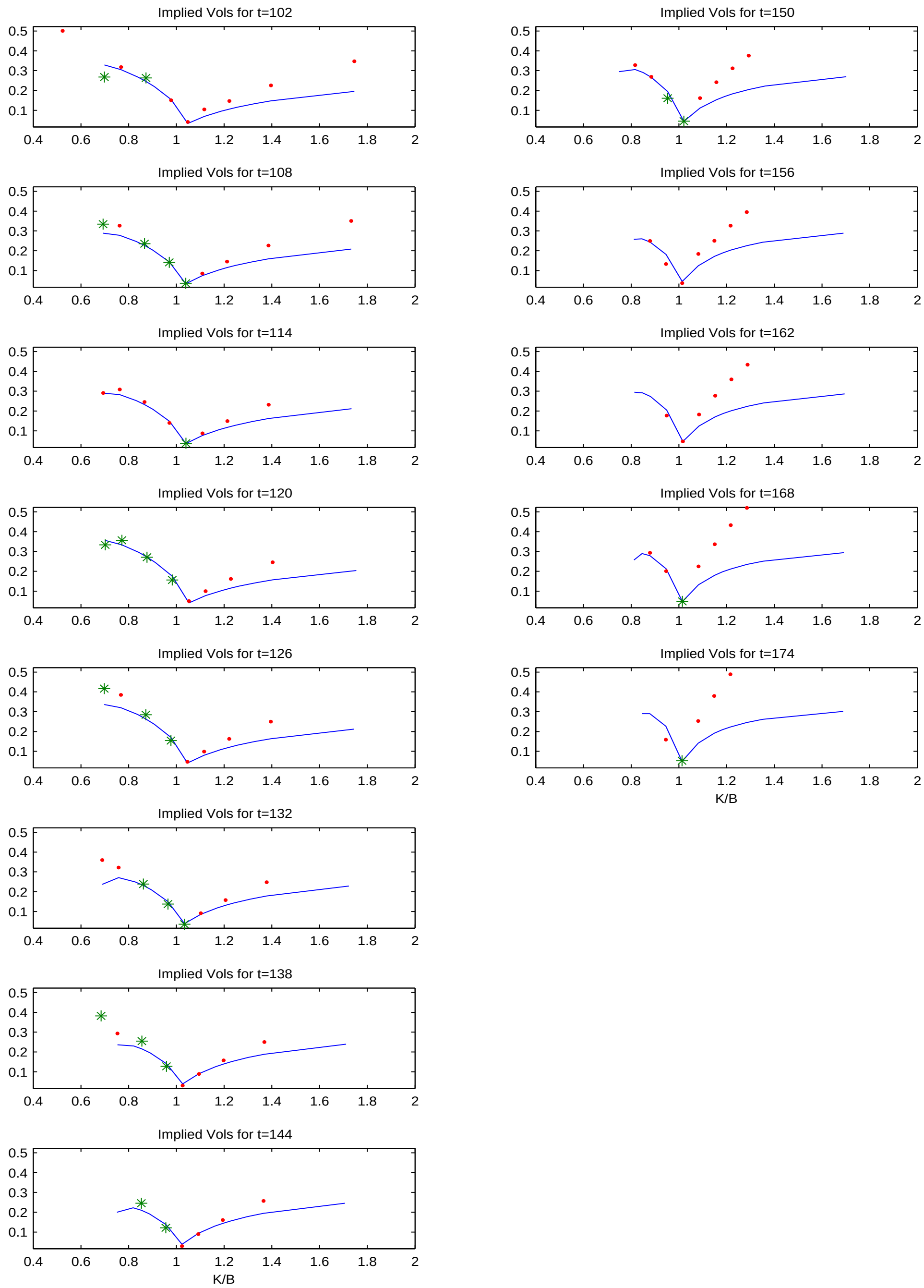

Implied Vols for t=102
Implied Vols for t=150
Implied Vols for t=108
Implied Vols for t=156
Implied Vols for t=114
Implied Vols for t=162
Implied Vols for t=120
Implied Vols for t=168
Implied Vols for t=126
Implied Vols for t=174
K/B
Implied Vols for t=132
Implied Vols for t=138
Implied Vols for t=144
K/B

Table 1: This reports summary statistics for the initial sample of 582 cash mergers from January 1996 to June 2008, for which the target company has sufficiently liquid options traded on its stock. The liquidity criterion is derived by looking at the average number of options traded per day (with positive trading volume). We report the 5-th, 25-th, 50-th, 75-th, and 95-th percentiles for: (i) the duration of the deal, i.e., the number of trading days after the deal was announced, but before either it was completed, or it failed; (ii) the offer premium, i.e., the percentage difference between the offer price per share, and the share price for the target company one day before the deal was announced; (iii) the percentage of trading days for which there is at least one traded option; and (iv) the average number of options traded on each day.

Data Description					
Percentile	5%	25%	50%	75%	95%
Deal Duration	32	52	84	137	223
Offer Premium	2.66%	12.20%	25.00%	41.85%	83.22%
% of Days with Options Traded	1.75%	14.68%	28.57%	50.89%	91.97%
Ave. No. of Options Traded Per Day	1.0%	1.09%	1.33%	1.79%	3.68%

Table 2: This reports summary data for ten cash mergers in our sample, five of which were successful acquisitions, and five of which failed. These are the deals with the largest $\eta_i - \frac{1}{2}\zeta_i$ as described in table 1. Panel A reports the names of the acquirer and target company, together with the ticker of the target company. For the ten selected deals, Panel B reports: the ticker, the announcement date, the date when the deal succeeded or failed, the offer price, and the target price one day before the announcement.

Panel A: List of Deals

Target Company	Target Ticker	Acquirer Company
Computer Science Corp.	CSC	Computer Assoc. Intl. Inc.
Gemstar International Group	GMSTF	United Video Satellite Group
MCI Communications Corp.	MCIC	GTE Corp.
Take-Two Interactive Software	TTWO	Electronic Arts Inc.
Unocal Corp.	UCL	CNOOC
AT&T Wireless Services Inc.	AWE	Cingular Wireless LLC.
DSP Communications Inc.	DSP	Intel Corp.
Georgia Pacific Corp.	GP	Koch Forest Products Inc.
Millennium Pharmaceuticals Inc.	MLNM	Mohagany Acquisition Corp.
Platinum Tech. Inc.	PLAT	Computer Assoc. Intl. Inc.

Panel B: Deal Description

	Announcement	Closure Date			Target Price
Ticker	Date	Succeeded	Failed	Offer Price	Before Announcement
CSC	10-Feb-1998		10-Mar-1998	$108.00	$88.50
GMSTF	06-Jul-1998		22-Jul-1998	$45.00	$38.875
MCIC	15-Oct-1997		17-Dec-1997	$40.00	$25.125
TTWO	24-Feb-2008		14-Sept-2008	$26.00	$20.85
UCL	22-Jun-2005		02-Aug-2005	$67.00	$44.34
AWE	17-Feb-2004	26-Oct-2004		$15.00	$8.55
DSP	14-Oct-1999	11-Nov-1999		$36.00	$28.00
GP	13-Nov-2005	23-Dec-2005		$48.00	$34.65
MLNM	10-Apr-2008	14-May-2008		$25.00	$16.35
PLAT	29-Mar-1999	06-Jun-1999		$29.25	$9.875

Table 3: Percentage Pricing Errors for the Target Price

This table uses a sample of 282 cash mergers from 1996–2008, which have sufficiently liquid options on the target company. The percentage stock price error is computed as follows: For each company i and on each day t, compute the stock pricing error $\left|\frac{\hat{B}^i(t)-B^i(t)}{B^i(t)}\right|$, where $\hat{B}^i(t)$ is the fitted price of company i according to our model: $\hat{B}^i(t) = q^i(t)B_1^i \mathrm{e}^{-r(T_e^i-t)} + (1-q^i(t))B_2^i(t)$. In this formula, $q^i(t)$ is the estimated risk-neutral probability that the deal is successful; B_1^i is the cash offer price; T_e^i is the effective date of the deal; and $B_2^i(t)$ is the fallback price, i.e., the price of company i if the deal fails. Next, for each stock i compute the mean μ_i over time of the stock pricing error, and the standard deviation σ_i. The table reports the 5-th, 25-th, 50-th, 75-th, and 95-th percentile of μ_i and σ_i over the 282 stocks in our sample. The estimates in the model are calculated using only the option with the highest volume each day. The errors shown are for all the options in the sample.

Percentiles of Percentage Pricing Errors for B

5%	25%	50%	75%	95%
0.00004	0.00011	0.00038	0.00178	0.00845

Table 4: Percentiles of Pricing Errors for Call Options on the Target Company

This table uses a sample of 282 cash mergers from 1996–2008, which have sufficiently liquid options on the target company. The table compares the pricing errors for four models: our model, denoted by "MRB"; and three versions of the Black–Scholes formula, which differ only in the way the volatility is computed. Panel A reports the percentage error: for a given call option on company i on day t, with strike price K and maturity T, the percentage error is defined as $\left|(C^i_{\text{model},K,T,t} - C^i_{K,T,t})/C^i_{K,T,t}\right|$, depending on the model used. The MRB model defines the call option price by $\hat{C}^i_{K,T,t} = q^i(t)\max\{B^i_1 - K, 0\}\mathrm{e}^{-r(T^i_e - t)} + (1 - q^i(t))C_{BS}(B^i_2(t), K, r, T-t, \sigma^i_2)$, where: $q^i(t)$ is the estimated risk-neutral probability that the deal is successful; B^i_1 is the cash offer price; T^i_e is the effective date of the deal; $B^i_2(t)$ is the fallback price, i.e., the price of company i if the deal fails; σ^i_2 is the estimated volatility of the fallback price $B^i_2(t)$; and $C_{BS}(S, K, r, T-t, \sigma)$ is the Black–Scholes formula for the European call option price over a stock with price S at time t and volatility σ. Under the three versions of the Black–Scholes formula, we define the option price by $C_{BS}(B^i(t), K, r, T-t, \sigma)$, where $B^i(t)$ is the stock price at t, and the volatility σ is defined as: (1) the average implied volatility for at-the-money (ATM) call options quoted on company i during the time of the deal; (2) the implied volatility for an ATM call option quoted on the previous day $(t-1)$; (3) the implied volatility for a call option quoted on the previous day with the strike price closest to K. Panel B uses the absolute error: $\left|C^i_{\text{model},K,T,t} - C^i_{K,T,t}\right|$. Panel C uses the absolute error divided by the bid-ask spread of the corresponding option. Once the error is computed for each stock i, time t, and strike K, fix the stock i and compute the mean μ_i of the stock pricing error over time and strike, equally weighted. The table reports various percentiles (5, 25, 50, 75, 95) for μ_i over the 282 stocks in our sample. Each panel reports the results for: all call options; near-in-the-money (Near-ITM) calls, i.e., call options with the strike price K over the target stock price B_t are in the range $K/B \in [0.95, 1.0]$; near-out-the-money (Near-OTM) calls with strike K so that $K/B \in [1.0, 1.05]$; in-the-money (NTM) with $K/B \in [0.90, 0.95]$; out-the-money (ITM) with $K/B \in [1.05, 1.10]$; deep-in-the-money (Deep-ITM) with $K/B < 0.90$; and deep-out-of-the-money (Deep-OTM) with $K/B > 1.10$. Panels D and E report percentiles over the absolute and the percentage bid-ask spread, respectively. In all panels, N represents the number of cross-sectional observations. The estimates in the model are calculated using only the option with the highest volume each day.

Panel A: Percentage Errors for Four Option Pricing Models

Selection	Model	Percentile 5%	25%	50%	75%	95%	N
All Call Options	MRB	0.08591	0.18146	0.26059	0.36489	0.56335	282
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.10578	0.19007	0.33019	0.48892	1.11483	282
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.09937	0.20711	0.34223	0.51132	1.17115	282
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05724	0.11697	0.18189	0.28775	0.51272	282
Deep In-The-Money Calls ($K/B < 0.9$)	MRB	0.00441	0.00866	0.01504	0.02186	0.04860	281
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.00490	0.01061	0.01647	0.02609	0.05725	281
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.00512	0.01191	0.01799	0.02664	0.05705	281
	BS: $\sigma = \sigma^i_{K,t-1}$	0.00580	0.01190	0.01820	0.02707	0.05023	281
In-The-Money Calls ($K/B \in [0.9, 0.95]$)	MRB	0.01033	0.02273	0.04254	0.07307	0.16680	213
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.01410	0.03369	0.05908	0.11064	0.26527	213
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.01711	0.03569	0.05514	0.08447	0.22226	213
	BS: $\sigma = \sigma^i_{K,t-1}$	0.01666	0.03244	0.05306	0.08650	0.19545	213
Near In-The-Money Calls ($K/B \in [0.95, 1.0]$)	MRB	0.01883	0.04720	0.08036	0.14519	0.47372	205
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.03429	0.07514	0.13361	0.24715	0.81787	205
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.03371	0.06185	0.10413	0.20543	0.84077	205
	BS: $\sigma = \sigma^i_{K,t-1}$	0.02609	0.05932	0.11563	0.25139	1.02767	205
Near Out-The-Money Calls ($K/B \in [1.0, 1.05]$)	MRB	0.06724	0.19704	0.42291	0.87243	2.18229	204
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.11584	0.33458	0.62771	1.19446	5.14470	204
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.06646	0.20413	0.48669	0.85884	2.09441	204
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05744	0.19738	0.38476	0.68261	1.59153	204
Out-The-Money Calls ($K/B \in [1.05, 1.10]$)	MRB	0.12166	0.41035	0.67209	0.98636	1.41742	206
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.20219	0.53450	0.81586	1.21704	3.91596	206
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.10950	0.47117	0.85423	1.58244	6.21376	206
	BS: $\sigma = \sigma^i_{K,t-1}$	0.06403	0.23956	0.41111	0.67903	1.78833	206
Deep Out-The-Money Calls ($K/B > 1.1$)	MRB	0.37606	0.71212	0.94370	1.00000	1.44190	231
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.36172	0.80532	0.97943	1.00000	2.22103	231
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.35484	0.86320	1.00000	1.30257	4.27091	231
	BS: $\sigma = \sigma^i_{K,t-1}$	0.09207	0.23402	0.46254	0.68593	1.18797	231

Panel B: Absolute Errors for Four Option Pricing Models

Selection	Model	Percentile 5%	25%	50%	75%	95%	N
All Call Options	MRB	0.03885	0.07081	0.09991	0.15929	0.39036	282
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.04276	0.07867	0.12097	0.22181	0.57306	282
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.04640	0.08307	0.12038	0.20428	0.52179	282
	BS: $\sigma = \sigma^i_{K,t-1}$	0.03718	0.07181	0.11111	0.17533	0.42242	282
Deep In-The-Money Calls ($K/B < 0.9$)	MRB	0.02796	0.05638	0.09579	0.15974	0.43023	281
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.03071	0.06251	0.10534	0.17794	0.53655	281
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.03299	0.06943	0.10793	0.18433	0.51245	281
	BS: $\sigma = \sigma^i_{K,t-1}$	0.03605	0.07058	0.11032	0.18822	0.51498	281
In-The-Money Calls ($K/B \in [0.9, 0.95]$)	MRB	0.02460	0.06100	0.11416	0.21899	0.48129	213
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.02794	0.09433	0.16740	0.33065	0.93220	213
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.03144	0.08530	0.14849	0.28918	0.69731	213
	BS: $\sigma = \sigma^i_{K,t-1}$	0.04069	0.09344	0.14196	0.25583	0.52557	213
Near In-The-Money Calls ($K/B \in [0.95, 1.0]$)	MRB	0.02482	0.06293	0.10605	0.21825	0.59304	205
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.03966	0.08944	0.18748	0.40083	1.27648	205
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.04046	0.08099	0.14916	0.29604	0.61143	205
	BS: $\sigma = \sigma^i_{K,t-1}$	0.04053	0.09083	0.17234	0.34007	0.62459	205
Near Out-The-Money Calls ($K/B \in [1.0, 1.05]$)	MRB	0.02072	0.05308	0.10883	0.21882	0.51736	204
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.02674	0.07389	0.14563	0.38565	1.21321	204
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.02687	0.06036	0.11406	0.23950	0.70509	204
	BS: $\sigma = \sigma^i_{K,t-1}$	0.02070	0.05155	0.09327	0.20685	0.56115	204
Out-The-Money Calls ($K/B \in [1.05, 1.10]$)	MRB	0.02978	0.06248	0.10612	0.19029	0.45875	206
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.03606	0.07438	0.13524	0.31414	1.04700	206
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.03733	0.07782	0.12686	0.28159	1.03029	206
	BS: $\sigma = \sigma^i_{K,t-1}$	0.01630	0.04183	0.06996	0.12167	0.39998	206
Deep Out-The-Money Calls ($K/B > 1.1$)	MRB	0.03974	0.08053	0.11923	0.19246	0.51778	231
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.04280	0.08340	0.12500	0.24525	0.68421	231
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.03864	0.08744	0.13679	0.26473	0.77783	231
	BS: $\sigma = \sigma^i_{K,t-1}$	0.01156	0.02901	0.05474	0.10208	0.28423	231

Panel C: Absolute Errors Divided by Observed Bid-Ask Spread for Four Option Pricing Models

Selection	Model	Percentile					
		5%	25%	50%	75%	95%	N
All Call Options	MRB	0.16514	0.24056	0.32222	0.42886	1.03515	282
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.18701	0.27737	0.40442	0.66493	1.65719	282
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.20566	0.32979	0.42908	0.63422	1.63183	282
	BS: $\sigma = \sigma^i_{K,t-1}$	0.17545	0.27418	0.35168	0.52383	1.03465	282
Deep In-The-Money Calls ($K/B < 0.9$)	MRB	0.06643	0.11734	0.17349	0.29305	0.85159	281
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.07847	0.12118	0.20511	0.35230	0.79279	281
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.08840	0.14413	0.22661	0.37496	0.88582	281
	BS: $\sigma = \sigma^i_{K,t-1}$	0.09922	0.16045	0.22724	0.36447	0.70894	281
In-The-Money Calls ($K/B \in [0.9, 0.95]$)	MRB	0.07869	0.19038	0.32855	0.57705	1.38377	213
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.09610	0.25648	0.52789	1.11088	2.66648	213
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.12387	0.26284	0.44932	0.91271	2.23472	213
	BS: $\sigma = \sigma^i_{K,t-1}$	0.14098	0.28764	0.48538	0.79276	1.70762	213
Near In-The-Money Calls ($K/B \in [0.95, 1.0]$)	MRB	0.10368	0.25221	0.41819	0.71268	1.96357	205
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.17202	0.38034	0.74865	1.49529	4.87438	205
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.16331	0.38419	0.58142	1.06230	2.47430	205
	BS: $\sigma = \sigma^i_{K,t-1}$	0.15370	0.38680	0.65649	1.34888	3.44949	205
Near Out-The-Money Calls ($K/B \in [1.0, 1.05]$)	MRB	0.15958	0.33090	0.49984	0.88331	2.31108	204
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.21536	0.44726	0.80527	1.82159	5.58248	204
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.16095	0.36714	0.56192	1.07262	3.61276	204
	BS: $\sigma = \sigma^i_{K,t-1}$	0.13582	0.32219	0.49326	0.85702	2.61802	204
Out-The-Money Calls ($K/B \in [1.05, 1.10]$)	MRB	0.19864	0.36258	0.50000	0.76543	2.31380	206
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.23385	0.44178	0.64746	1.54473	5.90195	206
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.22807	0.44571	0.71525	1.47772	6.87030	206
	BS: $\sigma = \sigma^i_{K,t-1}$	0.08484	0.22005	0.38292	0.69955	1.78870	206
Deep Out-The-Money Calls ($K/B > 1.1$)	MRB	0.28251	0.45203	0.50000	0.58485	1.84053	231
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.29794	0.47649	0.50000	0.76486	2.46184	231
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.32090	0.49447	0.57293	0.90661	3.19430	231
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05344	0.16482	0.29548	0.43982	0.96805	231

Panel D: Absolute Bid-Ask Spread for Call Options

	Percentile					
Selection	5%	25%	50%	75%	95%	N
All Call Options	0.18591	0.28996	0.42734	0.62779	1.49941	282
Deep-ITM Calls	0.24463	0.39450	0.56797	0.84693	2.08308	281
ITM Calls	0.15391	0.23968	0.34233	0.54270	1.28734	213
Near-ITM Calls	0.11199	0.19545	0.26270	0.41990	1.33650	205
Near-OTM Calls	0.06964	0.15052	0.21726	0.29989	0.84065	204
OTM Calls	0.08600	0.15000	0.21875	0.27500	0.62845	206
Deep-OTM Calls	0.08450	0.16392	0.22069	0.30525	1.07353	231

Panel E: Percentage Bid-Ask Spread for Call Options

	Percentile					
Selection	5%	25%	50%	75%	95%	N
All Call Options	0.13644	0.38293	0.58819	0.80644	1.15064	282
Deep-ITM Calls	0.03206	0.05535	0.07919	0.11602	0.23678	281
ITM Calls	0.05457	0.08642	0.13668	0.20160	0.40486	213
Near-ITM Calls	0.06632	0.12912	0.19560	0.33721	1.02209	205
Near-OTM Calls	0.12398	0.42639	1.16594	1.72551	2.00000	204
OTM Calls	0.16644	1.07135	1.63823	1.94286	2.00000	206
Deep-OTM Calls	0.55483	1.74929	1.96550	2.00000	2.00000	231

Table 5: The Call Price Kink and the Risk-Neutral Probability

This table uses a sample of 282 cash mergers from 1996–2008 with sufficiently liquid options on the target company. Consider the call price kink multiplied by the time discont factor, $C_{\text{kink}} = e^{r(T-t)}\left(\left(\frac{\partial C}{\partial K}\right)_{K\downarrow B_1} - \left(\frac{\partial C}{\partial K}\right)_{K\uparrow B_1}\right)$, which according to Equation (13) satisfies $C_{\text{kink}} = q(t)$. The table reports a pooled panel regression of the (natural) logarithm of the call price kink on the logarithm of the estimated risk-neutral probability. T-statistics are reported in parentheses, and are obtained from standard errors clustered by firm.

OLS Regression of $\ln(C_{\text{kink}})$ on $\ln(q)$

	const	$\ln(q)$	Adj. R^2	No. of Obs.
$\ln(C_{\text{kink}})$	-0.006	0.922^{***}	21.56%	13,181
	(-0.05)	(4.19)		

Table 6: The Success Probability as a Predictor of Deal Outcome

This table uses a sample of 282 cash mergers from 1996–2008, which have sufficiently liquid options on the target company. It compares the predictive power of the risk-neutral probability $\hat{q}$ estimated using our model with that estimated using a naive method. For each company i, consider 10 equally spaced days t_n throughout the deal: for each $n = 1, \ldots, 10$, choose t_n the closest integer strictly smaller than $n\frac{T_e}{10}$. Use the model to compute $\hat{q}^i(t_n)$, the risk-neutral probability that the deal is successful. Define also $q^i(t_n)$ using a "naive" method: $q^i_{\text{naive}}(t_n) = \frac{B^i(t_n) - B^i_0}{B^i_1 - B^i_0}$, where $B^i(t_n)$ is the stock price at t_n, B^i_1 is the cash offer price, and B^i_0 is the stock price before the deal was announced. At each t_n, compute perform a probit regression of outcome of deal i (1 if it succeeds, 0 if it fails) on the success probability $q^i(t_n)$. The figures reported in the table are the pseudo-R^2. The estimates in the model are calculated using only the option with the highest volume each day.

Pseudo-R^2 of Probit Regression of Outcome on the Success Probability Estimated at $n\frac{T_e}{10}$

n	R^2 for $\hat{q}$	R^2 for q_{naive}	N
1	0.11678	0.00116	279
2	0.11508	0.00311	279
3	0.13040	0.01684	279
4	0.20788	0.03757	279
5	0.22979	0.03091	279
6	0.29651	0.03250	279
7	0.38594	0.04887	279
8	0.43682	0.08845	279
9	0.42681	0.14203	279
10	0.47306	0.28219	279

Table 7: The Behavior of the Fallback Price after a Takeover Announcement

This table uses a sample of 282 cash mergers from 1996–2008, which have sufficiently liquid options on the target company. For a company i subject to a takeover deal in our sample, it compares the fallback price $B_2^i(t_n)$, estimated using our model, with the stock price B_0^i before the deal announcement. The fallback price is estimated at 10 equally spaced days t_n throughout the deal: for each $n = 1, \ldots, 10$, choose t_n the closest integer strictly smaller than $n\frac{T_e}{10}$. The regression model is $\ln(B_2^i(t_n)) = a + b\ln(B_0^i) + \varepsilon$. t-statistics are reported in parentheses. The estimates in the model are calculated using only the option with the highest volume each day.

Regression of Log-Fallback Price on Log-Price before Announcement

n	a	b	R^2	N
1	0.20846 (4.658)	0.97585 (69.365)	0.94556	279
2	0.24639 (5.509)	0.96766 (68.818)	0.94474	279
3	0.26334 (5.946)	0.96537 (69.338)	0.94552	279
4	0.25530 (5.458)	0.96815 (65.834)	0.93993	279
5	0.23855 (4.525)	0.97203 (58.647)	0.92547	279
6	0.21793 (3.917)	0.97876 (55.954)	0.91872	279
7	0.21081 (3.842)	0.98186 (56.919)	0.92124	279
8	0.22492 (4.218)	0.98073 (58.500)	0.92512	279
9	0.24132 (4.217)	0.97882 (54.409)	0.91444	279
10	0.27167 (4.619)	0.97273 (52.613)	0.90903	279