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. The formula can be easily generalized to other types of binary events and in particular we show it can be extended to stock-for-stock mergers. 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 show that the formula produces a volatility smile that is close to the one observed in practice. Moreover, our fomula predicts that the kink in the volatility smile is closely related to the success probability: the magnitude of the kink (the slope difference) should be equal to the time discounted risk-neutral probability of deal success, divided by the option vega.

[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.

To set up the model, consider a company B which is the target of a cash merger. If the deal is successful it receives a fixed offer price B_1, while if the deal fails its price reverts to the 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 observed variables: the price of the target company B and the prices of the various existing options on B. We estimate the whole times series of the two latent variables along with the structural parameters of the model by using a Markov Chain Monte Carlo (MCMC) algorithm.[5] Rather than using all options traded on the stock, it is enough to choose only one option each day, e.g., the call option with the maximum trading volume on that day.

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-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. (The error is large because

[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). The MCMC method is similar to the maximum likelihood estimation, except that one cannot directly sample from the joint distribution (conditional on the observables). Instead one can use the Gibbs sampling to form Markov chains by using the known conditional distributions, and show that these chains converge to a sample from the joint distribution.

option prices on individual stocks and especially on target companies in cash mergers have large bid-ask spreads and therefore are very noisy.) 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 obsered 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. In fact, when comparing the estimates for the options which have positive trading volume, our formula does better even than the modified Black–Scholes formula: the median percentage error is 17.49% for our model, and 23.57% for the modified Black–Scholes model.

We test the implications of the model regarding the kink in the volatility smile. We find that indeed a higher success probability is correlated with a bigger kink, both in the time series (for the same stock across time) and in the cross section (comparing the averages across stocks).

We show that the estimated success probability predicts well the outcome of the deal. This method does significantly better than the "naive" method widely used in the mergers and acquisitions literature, which estimates the success probability by looking at the current stock price and seeing how close it is to the offer price in comparison to the pre-announcement price. Our method is similar to the naive method except the success probability is estimated by using the fallback price instead of the pre-announcement price. (The fallback price has to be estimated as well in our model, as it is a latent variable.)

We also give a formula for the instantaneous stock volatility based on the probability of success and its volatility, as well as on the fallback price and its volatility. This model-implied volatility is relatively closer to the at-the-money Black–Scholes implied volatility than the volatility fitted from a GARCH(1,1)-model. Our model can explain the fact that 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.

An interesting question is 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-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 may indicate that other takeover attempts are now more likely. We find that indeed the fallback price is on average 30–40% higher than the pre-announcement price.

Another implication of our model is the possibility to estimate the merger risk premium, 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 167% annual rate (with an error of ±43%). 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.

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.[6] 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 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.[7]

[6] 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.

[7] According to this assumption, if the Poisson process has not jumped until the effective date, the merger

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.[8] 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. Within our framework, this indicates that there is a difference between the real probability distribution of the outcome of a merger and its risk-neutral counterpart, which corresponds to the 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.

---

is successful. This has the counter-factual implication that even deals that eventually fail become more likely as they approach the effective date.

[8]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 for which company A makes a cash offer to company B in the amount of B_1 dollars per share. The deal must be decided by the time T_e, which is called the "effective date," when the uncertainty about the merger is resolved. (In the event of success, the shareholders of B are assumed to get B_1 on the effective date.) When one considers the price of the target company B after the offer has been announced but has not yet been accepted, its fluctuations come mainly from two different sources. First, there are fluctuations in the probability of the merger being successful. Second, there are fluctuations in the component of the price of B which is independent of the takeover attempt.

To make these risks more precise, define the following notions. 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.[9] Also, define the *fallback price* $B_2(t)$ as the market-estimated value at t of company B 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.

We setup the model in continuous time. (See for example Duffie (2001), Chapters 5 and 6.) Let $W(t)$ be a 3-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F})$ with the standard associated Brownian filtration. 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

[9] 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.

is constrained to be between 0 and 1. We assume that p_m is independent of B_1 and B_2.[10] 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.

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 goes through. Because $p_m(t)$ is a discounted martingale with respect to Q, we have

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

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',

[10] 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.

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, 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 prove 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. (Options are forced to expire at the effective date if the deal is successful.) 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. 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.

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 the call option price against the strike price. If the Black–Scholes model were the correct one the plot 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 volatiltiy plot 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 sometimes the volatility "smirk" if the plot does not go up as much as it came down.

The next result shows that in the case of options on cash mergers the volatility smile arises naturally if the merger is the success probability is high enough. Moreover, when the strike price K equals the offer price B_1 there is a kink in the (Black–Scholes) implied volatility plot, and we show that the magnitude of the kink (the slope difference) 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*

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

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

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

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.

Note that $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K<B_1}$ 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, $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K>B_1}$ 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 $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K<B_1}$ is negative and $\left(\frac{\partial \sigma^{\text{imp}}}{\partial K}\right)_{K>B_1}$ 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*

$$\left(\sigma_B(t)\right)^2 = \left(\frac{B_1\mathrm{e}^{-(T_e-t)} - B_2(t)}{B(t)}q(t)(1-q(t))\sigma_1\right)^2 + \left(\frac{B_2(t)}{B(t)}(1-q(t))\sigma_2\right)^2 \tag{12}$$

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

2.2 Intuition

Start again with the formulas we obtained in the case of cash mergers: $B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t)$, $C(t) = q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C_2(t)$. Denote by $B_1(t) = B_1\mathrm{e}^{-r(T_e-t)}$, and $C_1(t) = (B_1-K)_+\mathrm{e}^{-r(T_e-t)}$. Rewrite the equations for B and C by removing the dependence on t: $B = qB_1 + (1-q)B_2$, $C = qC_1 + (1-q)C_2$. The transformation

$(p, B_2) \mapsto (B, C)$ has the following linearization

$$\begin{cases} \Delta B = (B_1 - B_2)\Delta q + (1-q)\Delta B_2, \\ \Delta C = (C_1 - C_2)\Delta q + (1-q)\delta_2 \Delta B_2, \end{cases} \tag{13}$$

where $\delta_2 = N(d_1)$ is the stock option delta, as in Equation (7). Define $r_{CB} = \frac{C_1 - C_2}{B_1 - B_2}$. The inverse has the following linearization

$$\begin{cases} \Delta q = \frac{1}{(1-q)(r_{CB} - \delta_2)}(r_{CB}\Delta B - \Delta C), \\ \Delta B_2 = \frac{-1}{(B_1 - B_2)(r_{CB} - \delta_2)}(\delta_2 \Delta B - \Delta C). \end{cases} \tag{14}$$

Note that if $r_{CB} \neq \delta_2$, one can attribute changes in the risk-neutral probability q and the fallback price B_2 to changes in the stock price B and the option price C. The reason why this method works is that options are non-linear in the price of the underlying. So for maximum power this method should be used where the curvature of the option price is at a maximum, i.e., when the option is at the money.

3 Empirical Results

3.1 Data

We study cash merger deals that were announced between January 1996 and June 2007, 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.

Although cash deals are the most common when public companies are acquired, these companies are usually small and have no options traded on their stock. Therefore, our initial sample contains all the cash deals for which the target company has a market capitalization greater than USD 500 million, and has options traded with data available in OptionMetrics. This gives us a total sample of 350 deals. Out of these 350 deals, 280 successfully completed the merger while 62 failed to reach an agreement by the time of the effective date. For the rest of the 8 deals, the transaction was still pending as of the end of June 2007.

Table 1 reports some summary statistics for our initial sample. For example, the median deal lasted 87 trading days (until either it succeeded or failed). The average deal duration is 111 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 26.58%, while the mean is 31% and the standard deviation is 30%. The table further 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 30.46%, indicating that options are quite illiquid.

In order to use our statistical procedure, we narrow the sample to include only merger deals for which the offer price was not renegotiated after the announcement day, and for which the target company has at least 30% of trading days when at least one option has positive trading volume. Since we use options with relatively short maturities, we further restrict the sample to deals where the duration of the deal was not greater than 120 trading days. Finally, we also exclude deals where the acquirer would end up with less than 80% of the outstanding shares of the target company. These restrictions give us a final sample of 282 deals.

For the duration of each deal, we collect the closing stock prices, and the closing bid and offer prices for 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. This procedure is stipulated in the Options Clearing Corporation (OCC) By-Laws and Rules (Article VI, Section 11).

Table 2 reports ten deals in our sample, five of which were successful acquisitions, and five of which failed. These are the deals with the most liquid options, i.e., with the largest percentage of days when at least one option has positive trading volume.

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 $t = 0$); if the deal succeeds or fails before the effective date, we redefine T_e to denote the success or failure date; (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 at time 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.[11]

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. The variables q and B_2 satisfy:

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

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

Since q and B_2 are independent, $\mathrm{d}W_1(t)$ and $\mathrm{d}W_2(t)$ are independent. Assume that equations (10) and (11) from Theorem 1 hold only approximately, with errors $\varepsilon_B(t)$ and $\varepsilon_C(t)$:

$$B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t) + \varepsilon_B(t), \tag{17}$$

$$C(t) = q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C^{BS}(B_2(t), K, r, T-t, \sigma_2) + \varepsilon_C(t). \tag{18}$$

The errors are IID bivariate normal:

$$\begin{bmatrix} \varepsilon_B(t) \\ \varepsilon_C(t) \end{bmatrix} \sim N(0, \Sigma_\varepsilon), \qquad \text{where} \qquad \Sigma_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}. \tag{19}$$

In conclusion, we have the observed variables $B(t)$, $C(t)$ and the observed parameters B_1, r, T_e, together with the latent (unobserved) variables $q(t) = X_1(t)$ and $B_2(t) = \mathrm{e}^{X_2(t)}$ and the latent parameters μ_1, σ_1, μ_2, σ_2, σ_{ε_B} and $\sigma_{\varepsilon,C}$. The idea is to estimate the latent variables and parameters by using the observed variables and parameters. We do this using a Markov

[11] 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.

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 (17) and (18) specify the relationship between the latent variables and the observables. We introduce the errors ε_B and ε_C in the pricing equations to allow for model misspecification; this also allows us to easily extend the estimation procedure to multiple options and missing data. This approach is not new to our paper and is discussed in general in Johannes and Polson (2003), Koop (2003), and many other sources. The resulting estimation procedure is a generalization of the Kalman filter, and is described in detail in Appendix B.[12]

To illustrate the results, we select a total of ten deals, reported in Table 2, five of which succeeded, and five of which failed. These are selected based on the target company having most liquid options, measured by the percentage of trading days for which there is at least one option with a positive trading volume. The estimation results for the risk-neutral probability $q(t)$ and the fallback price $B_2(t)$ are plotted in Figures 1 and 2. One sees that the q estimates for the five deals that succeeded—on the left column—are much higher than for the five deals that failed—on the right column. (The statistical results over the whole sample of 282 companies for a probit regression of the deal outcome on the q estimates are reported in Table 5.) We also select a specific deal, corresponding to the target company PLAT (see Table 2), for which we plot the MCMC draws for: the latent variables at half the effective date $\left(X_1\left(\frac{T_e}{2}\right), X_2\left(\frac{T_e}{2}\right)\right)$, and the latent parameters (μ_1, σ_1, μ_2, σ_2, $\sigma_{\varepsilon,B}$, and $\sigma_{\varepsilon,C}$).

Once we have the estimates for the latent variables and parameters, we can use them, e.g., to compute the prices of the other options. This is done assuming that equations (10) and (11) are true, i.e., equations (17) and (18) hold without error. This means that the model also implies a theoretical price for the observed option, since this is supposed to be priced with error in the estimation process. To see how well the model prices the cross section of options for each day, we report the results for companies PLAT and KMET in Figures 5 and 6. There the observed call option prices are compared with the model implied option prices. (The option prices are reported using the Black–Scholes implied volatility, so one can also study the volatility smile.)

[12] As noted by Johannes and Polson (2003), equations of the type (17) or (18) 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 rather than filtering, because it uses all the data at once.

An appealing feature of MCMC is that it can incorporate many types of information and priors. For example, one can make the returns corresponding to the fallback price B_2 correlated to some fitted Fama–French portfolio, or the DGTW characteristics portfolio. This corresponds to the systematic risk of the target company during the merger. One can also account for idiosyncratic risk, e.g., by examining the news related to the merger. One can rank each day as a positive, negative, or neutral day in terms of the underlying price. For example, if several analysts downgraded the company, it is a negative news day. Then each day one can add an extra term to the return of B_2, proportional to the sign of the news day. Therefore, in principle one can use the systematic and idiosyncratic information about the fallback price to estimate the success probability of the merger even when there are no options traded on the target company. Furthermore, one can impose priors on the parameters of the model in order to improve the precision of the estimation. For example, one may use the observed failed deals to extract information about the fallback price, and then use the results as priors for the estimation of each particular deal. Any information one has about the parameters of the model can be in principle used as a prior in the MCMC estimation.

Remark 1. One may wonder why we chose the particular specification (15) for q. One reason is that it 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 we also introduced the term $1-q$ in the denominator to ensure that q stays lower than 1.[13] But it also turns out that with this specification, the parameter μ_1 has a particularly useful interpretation. To see that, write

$$\mu_1 = \frac{\mathrm{d}\mathsf{E}_t}{\mathrm{d}t}\left\{\frac{\mathrm{d}q}{q(1-q)}\right\} = \frac{\mathrm{d}\mathsf{E}_t}{\mathrm{d}t}\left\{\frac{\mathrm{d}q}{q}\right\} - \frac{\mathrm{d}\mathsf{E}_t}{\mathrm{d}t}\left\{\frac{\mathrm{d}(1-q)}{1-q}\right\}. \tag{20}$$

But $q(t)$, time discounted, is the contingent price of the state in which the merger is successful, and $1-q(t)$ is the contingent price of the state when the merger fails. The difference in expected returns is then a measure of how much the risk-neutral probability deviates from the actual one, i.e., μ_1 has the same sign as $p(t)-q(t)$.[14] This implies that estimating μ_1 for many deals allows one to see whether there is a positive risk premium for merger risk, and

[13]Another reason is that for the natural logistic specification $q = \exp(X_1)/(1+\exp(X_1))$ the MCMC procedure runs into problems. This is because when X_1 becomes large, the posterior distribution in the MCMC algorithm is almost flat, so convergence of the algorithm becomes a problem.

[14]There is also the difference between p and q coming from the discrete jump on the very last day, but that is assumed to go in the same direction as in the continuous case.

whether this risk is correlated with particular macroeconomic variables.

3.3 Results

As described in the previous section, our sample contains 282 cash mergers during 1996–2007, which have sufficiently liquid options traded on the target company. Recall that our estimation method assumes that the pricing formulas (17) and (18) 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 (with zero errors) are our estimates for the stock price $\hat{B}(t)$ and the option price $\hat{C}(t)$. Table 3 reports cross-sectional percentiles over 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.

Figures 5 and 6 show that, in the case of companies PLAT and KMET (taken from Table 2 as the first company for which the merger succeeded and failed, respectively), our model fits well the time series and cross section of call options. A statistical analysis for all companies in our sample is done in Table 4, Panels A–C. This table reports pricing errors for four models of various types of call options on B. 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 we try to estimate. The model BS3 is in general the hardest one to beat, since it use the strike price closest to the current one, so it only needs to adjust for the change in the underlying price.

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 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) at-the-money (ATM) calls, with K closest to B; (3) near-the-money (NTM) calls, with $K/B \in [0.9, 1.1]$; (4) in-the-money (ITM) calls, with $K/B < 0.9$; (5) out-of-the-money (OTM) calls, with $K/B > 1.1$; (6) deep-ITM calls, with $K/B < 0.7$; and (7) deep-OTM calls, with $K/B > 1.3$.

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 47 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 47 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 16.15% for the BRM model, with 35.72% for BS1 and 35.08% for BS2. As mentioned above, model BS3 is hard to beat, and indeed it does better than our model: the median error is 11.26%. However, the MRB model does better in terms of the *absolute* pricing error (see Panel B): the median absolute error is 9.79% for the BRM model, compared with 14.76% for BS1, 14.24% for BS2, and 10.23% for BS3. The exception is for OTM calls, where BS3 does better than our model.[15] In conclusion, our model does relatively well compared with the BS3 model, which means that it goes some distance towards explaining the *volatility smile*, i.e., the tendency for at-the-money (ATM) options to have lower Black–Scholes implied volatilities than the other options.

Since our model is arbitrage-free, a natural question is whether one could use our model to make a profit. In principle, every deviation from the theoretical model should lead to an arbitrage opportunity. In reality, there are limits to arbitrage, including the most simple one, the bid-ask spread of the option. It turns out that the option bid-ask spreads are usually larger than the pricing error. Panel C 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

[15] A result not reported in Table 4 is that the performance of our model relative to the three versions of the Black–Scholes formula is very similar when we restrict our attention only to *out-of sample* call options, i.e., to the options that have not been used in our estimation method (i.e., they are not the options with the maximum trading volume on that day).

moneyness. So for the median stock the profit is dwarfed by 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 (with the benefit of hindsight). But the question is if one can devise an *ex-ante* profitable arbitrage strategy. Since this cannot be done for the median stock in our sample, the most likely answer is no.

Next, we consider the estimated time series of latent variables: the success probability $\hat{q}(t)$ and the fallback price $\hat{B}_2(t)$. Table 5 shows that $\hat{q}$ predicts well the outcome of the deal. First, 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 40% to about 75%, which indicates that the success probability predicts better and better the outcome of the merger. This is more remarkable since we do not impose the success probability to be 0 or 1 at the end. This presumably would lead to an even better fit.

One can 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. Mathematically, one defines $q_{naive}(t) = \frac{B(t)-B_0}{B_1-B_0}$, making sure also that if q is set to 1 if it goes above 1, and to 0 if it goes below 0. Table 5 reports the results from a cross-sectional probit regression of the deal outcome on $q_{naive}(t_n)$. Now, R^2 increases from 1% to 30%, indicating that our model does a better job at predicting the deal outcome than the "naive" one.

An interesting question is 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 6 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 is typically 30–40%, indicating that the fallback price is on average 30–40% higher than the pre-announcement price.

Another way in which we can test our model, is to use Proposition 2 to define a model-implied volatility $\sigma_{_B}$ of the target company B. The formula is $\sigma^2_{_B}(t) = \left(\frac{B_1 e^{-(T_e-t)} - B_2}{B} q(1-q)\sigma_1\right)^2 + \left(\frac{B_2}{B}(1-q)\sigma_2\right)^2$. Notice that when the deal is close to completion, the success probability q is close to 1 and the model-implied probability $\sigma_{_B}$ is close to 0. This corresponds to the known fact that the Black–Scholes implied probabilities when a merger is close to being completed are close to zero. Table 7 reports the differences in implied volatilities when one uses our model, the Black–Scholes model, and a GARCH(1,1)-model. Our model-implied volatility is indeed close to the Black–Scholes implied volatility, with a median difference of 4.89%, while the GARCH(1,1)-implied volatility has a median difference of 6.44% from the Black–Scholes implied volatility. See also Figure 7 for a graphic illustration, in the case of companies PLAT and KMET, of the volatility time series estimated using the three models. One can see that in the case of PLAT, the implied volatilities computed using our model and Black–Scholes beecome close to zero at the end, when the deal is close to completion, while this is not true when the GARCH(1,1) model is used.

Finally, we explore the possibility to estimate the merger risk premium, as the drift coefficient in the diffusion process for the success probability (15). According to Remark 1, $\frac{\mu_1}{2}$ can be considered as the merger risk premium. The individual estimates for μ_1 are very noisy, but over the whole sample the estimated merger risk premium is significantly positive, and the annual figure is 180%. This figure seems very high, but comparable figures for cash mergers have been reported in the literature. (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.

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.

An intersting 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 impllied 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 on how the current target stock price is situated between the offer price and the pre-merger announcement price. Besieds the success probability itself, we can also estimate its drift parameter. This number turns out to be related to the merger risk premium, which we estimate to be around 180%. This is a large figure, but is consistent with the cash mergers literature, e.g., see Dukes, Frolich and Ma (1992).

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 (17) and (18): $B(t) = q(t)B_1\mathrm{e}^{-r(T_e-t)} + (1-q(t))B_2(t) + \varepsilon_B(t)$, $C(t) = q(t)(B_1-K)_+\mathrm{e}^{-r(T_e-t)} + (1-q(t))C^{BS}_{K,\sigma_2,r,T}(B_2(t),t) + \varepsilon_C(t)$. 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 $q(t) = \frac{\mathrm{e}^{X_1(t)}}{1+\mathrm{e}^{X_1(t)}}$.[16]

Define the observed variables Y_B and Y_C by $Y_{B,t} = B(t)$, $Y_{C,t} = C(t)$. Also define the numbers $\bar{B}_t = B_1 \mathrm{e}^{-r(T_e - t)}$, $\bar{C}_t = (B_1 - K)_+ \mathrm{e}^{-r(T_e - t)}$, which vary deterministically with 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 $f_{C,t}(x_1, x_2) = f_{C,t}(x_1, x_2 \mid \sigma_2)$. Notice that with the new notation Equations (17) and (18) 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

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

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: $p(\theta, X, \Sigma_\varepsilon \mid Y) \propto p(Y \mid X, \Sigma_\varepsilon, \theta) \cdot p(X \mid \theta) \cdot p(\theta)$. 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:

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

$$p(X \mid \theta) = p(X_{1,1} \mid \theta) \cdot p(X_{2,1} \mid \theta) \cdot \prod_{t=2}^{T} \phi\left(\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right). \tag{25}$$

Recall that $Z_{k,t} = X_{k,t} - X_{k,t-1}$ for $k = 1, 2$, therefore it is normally distributed.

[16] 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.

We now start the MCMC algorithm.

STEP 0. Initialize $\theta^{(1)}$, $X^{(1)}$, $\Sigma_\varepsilon^{(1)}$. Fix a large number of iterations M ($M = 10,000$ is a good first choice). Then for each $i = 1, \dots M-1$ 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 $t = 2, \dots, T-1$,

$$\begin{aligned} p(X_t \mid \theta, \Sigma_\varepsilon, Y) \;\propto\; & \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t}(X_{1,t}, X_{2,t}) \\ f_{C,t}(X_{1,t}, X_{2,t}) \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix} \right) \\ & \cdot \phi\left(\begin{bmatrix} X_{1,t} - X^{(i+1)}_{1,t-1} \\ X_{2,t} - X^{(i+1)}_{2,t-1} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right) \\ & \cdot \phi\left(\begin{bmatrix} X^{(i)}_{1,t+1} - X_{1,t} \\ X^{(i)}_{2,t+1} - X_{2,t} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right). \end{aligned} \tag{26}$$

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 $p(\theta \mid X^{(i+1)}, \Sigma_\varepsilon^{(i+1)}, Y)$. To simplify notation, denote by $X = X^{(i+1)}$, and $\Sigma_\varepsilon = \Sigma_\varepsilon^{(i+1)}$. Assuming a flat prior for θ, $p(\theta \mid X, \Sigma_\varepsilon, Y) \propto p(Y \mid \theta, \Sigma_\varepsilon, X) \cdot p(X \mid \theta)$.

Then, if we assume that $X_{k,1}$ does not depend on θ,

$$p(\theta \mid X, \Sigma_\varepsilon, Y) \;\propto\; \prod_{t=1}^{T} \phi\left(\begin{bmatrix} Y_{B,t} \\ Y_{C,t} \end{bmatrix} \mid \begin{bmatrix} f_{B,t}(X_{1,t}, X_{2,t}) \\ f_{C,t}(X_{1,t}, X_{2,t} \mid \sigma_2) \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon,B} & 0 \\ 0 & \sigma^2_{\varepsilon,C} \end{bmatrix}\right)$$
$$\cdot \prod_{t=2}^{T} \phi\left(\begin{bmatrix} X_{1,t} - X_{1,t-1} \\ X_{2,t} - X_{2,t-1} \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right). \tag{27}$$

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: $\mu_k^{(i+1)} \sim N\Big(\frac{1}{T-1}\sum_{t=2}^{T} Z_{k,t}\,,\ \frac{\left(\sigma_k^{(i)}\right)^2}{T-1}\Big)$, $k = 1, 2$; $\left(\sigma_1^{(i+1)}\right)^2 \sim IG_2\Big(\sum_{t=2}^{T}\Big(Z_{1,t} - \mu_1^{(i+1)}\Big)^2, T-2\Big)$. For σ_2 the density is non-standard, so we need to perform the Metropolis–Hastings algorithm. Also, for the algorithm to converge it might be necessary to choose an appropriate prior distribution for σ_2.

METROPOLIS–HASTINGS. The goal of this algorithm is to draw a random element X out of 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 $X_{CURR} = X_0$. Now repeat the following steps as many times as necessary (usually once it is enough):

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

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

(3) Draw $u \sim U[0,1]$ (the uniform distribution on $[0,1]$). Then define $X^{(i+1)}$ by: if $u < \alpha$, $X^{(i+1)} = X_{PROP}$; if $u \geq \alpha$, $X^{(i+1)} = X_{CURR}$.

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] 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.

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

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

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

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

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

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

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

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

[10] 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.

[11] 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.

[12] 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.

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

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

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

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

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

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

[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 (2005): Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads; working paper.

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

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

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

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

Table 1: This reports summary statistics for the initial sample of 350 cash mergers from January 1996 to June 2007, for which the target company has options traded on its stock, and has a market capitalization of at least USD 500 million. 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; and (iii) the percentage of trading days for which there is at least one option with a positive trading volume.

Data Description					
Percentile	5%	25%	50%	75%	95%
Deal Duration	34	57	87	151	273
Offer Premium	3.27%	15.78%	26.58%	41.00%	78.78%
% of Days with Options Traded	3.29%	15.79%	30.46%	55.17%	92.71%

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 with the largest percentage of days when at least one option has positive trading volume. 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
Kemet Corp.	KMET	Vishay Intertechnology Inc.
MCI Communications Corp.	MCIC	GTE Corp.
Computer Science Corp.	CSC	Computer Assoc. Intl. Inc.
Gemstar Int. Group	GMST	United Video Satellite Group Inc.
Maytag Corp.	MYG	Investor Group
Platinum Tech. Inc.	PLAT	Computer Assoc. Intl. Inc.
FORE Systems	GMST	General Electric Co. PLC
Verio Inc.	VRIO	NTT Communications Corp.
Adv. Neuromodulations Sys. Inc.	ANSI	St. Jude Medical Inc.
Medlmmune Inc.	MEDI	AstraZeneca PLC.

Panel B: Deal Description

	Announcement	Closure Date			Target Price
Ticker	Date	Succeeded	Failed	Offer Price	Before Announcement
KMET	26-Jun-1996		26-Aug-1996	$22.00	$16.25
MCIC	15-Oct-1997		17-Dec-1997	$40.00	$25.125
CSC	10-Feb-1998		10-Mar-1998	$108.00	$88.5
GMST	26-Jun-1996		23-Aug-1996	$45.00	$38.875
MYG	20-Jun-2005		19-Jul-2005	$16.00	$11.47
PLAT	29-Mar-1999	06-Jun-1999		$29.25	$9.875
FORE	26-Apr-1999	30-Jun-1999		$35.00	$24.50
VRIO	05-May-2000	13-Sep-2000		$60.00	$36.875
ANSI	16-Oct-2005	30-Nov-2005		$61.25	$46.98
MEDI	23-Apr-2007	16-Jun-2007		$58.00	$48.01

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 PLAT, FORE, VRIO, ANSI, MEDI succeeded, while those for KMET, MCIC, CSC, GMST, MYG failed. The dash-dotted lines represent the 5% and 95% error bands around the estimated median values.

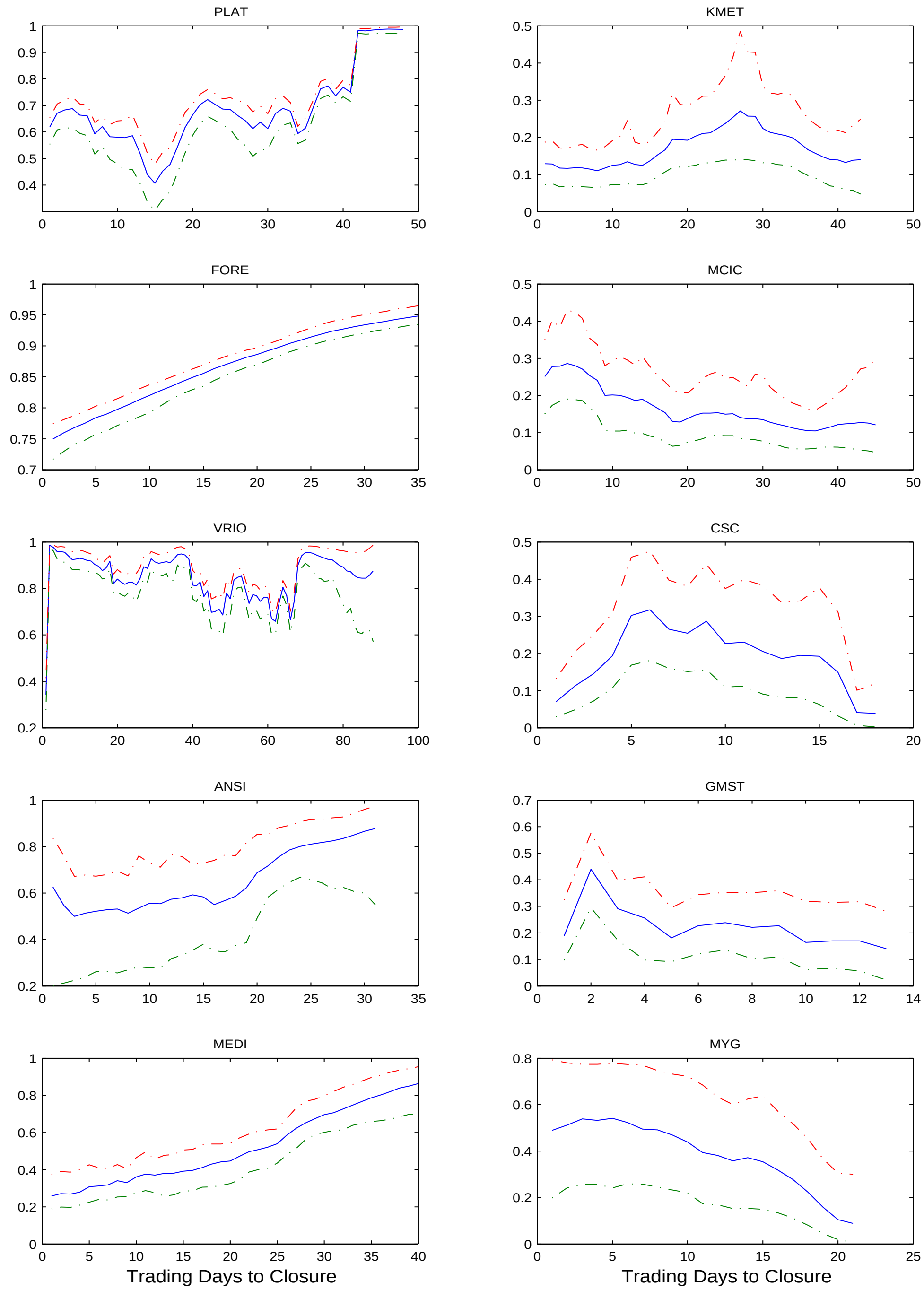


Figure 2: Estimates of the fallback prices of the target stock $B_2(t)$ for a subsample of ten cash mergers described in Table 2. The deals corresponding to target tickers PLAT, FORE, VRIO, ANSI, MEDI succeeded, while those for KMET, MCIC, CSC, GMST, MYG failed. The dash-dotted lines represent the 5% and 95% error bands around the estimated median values.

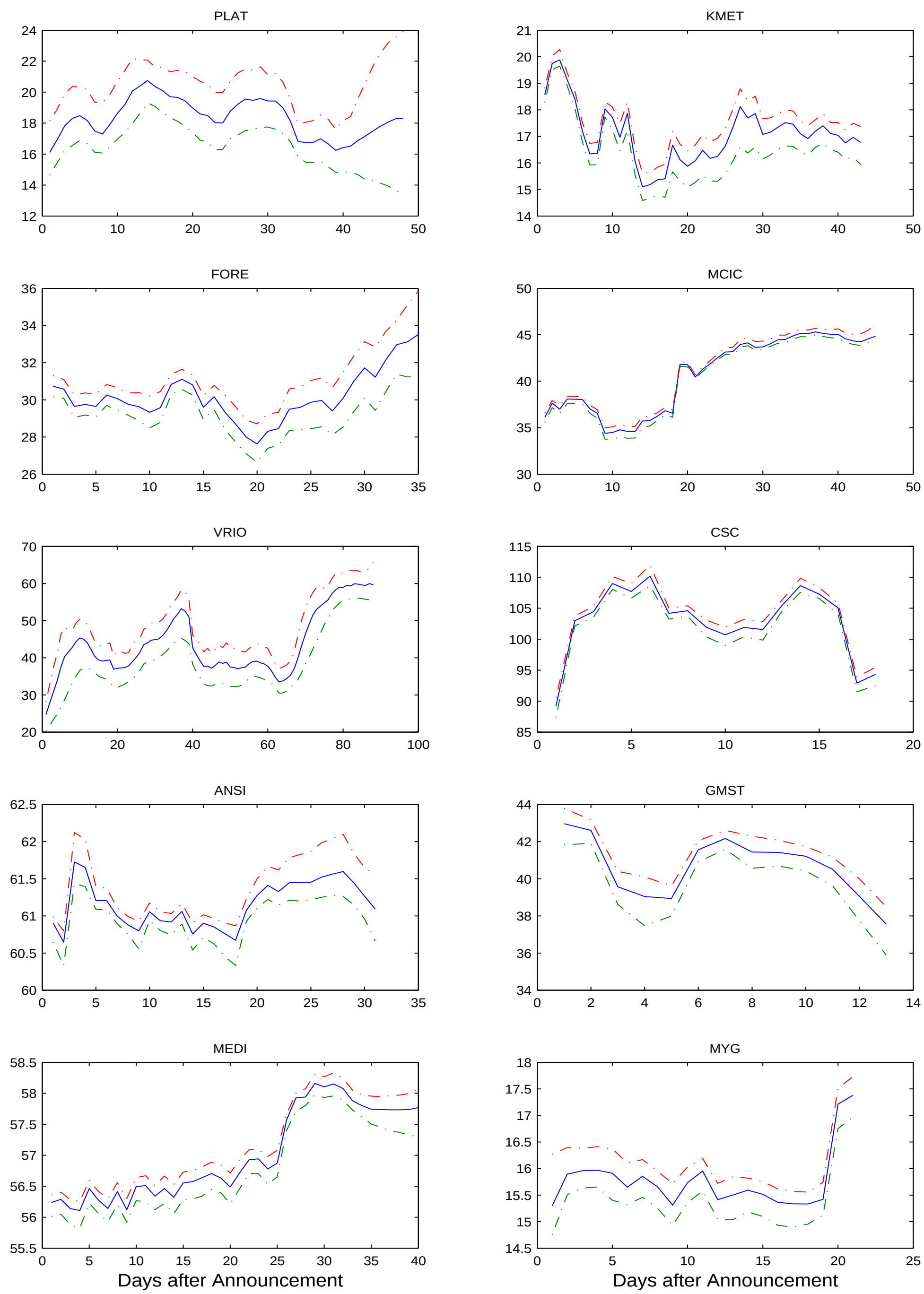


Figure 3: Consider the deal described in Table 2 corresponding to the target company PLAT. This figure plots: (i) the offer price, discounted at the current interest rate, using a dashed-dotted line; (ii) the stock price, using a continuous line; and (iii) the estimated fallback price (the price of the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

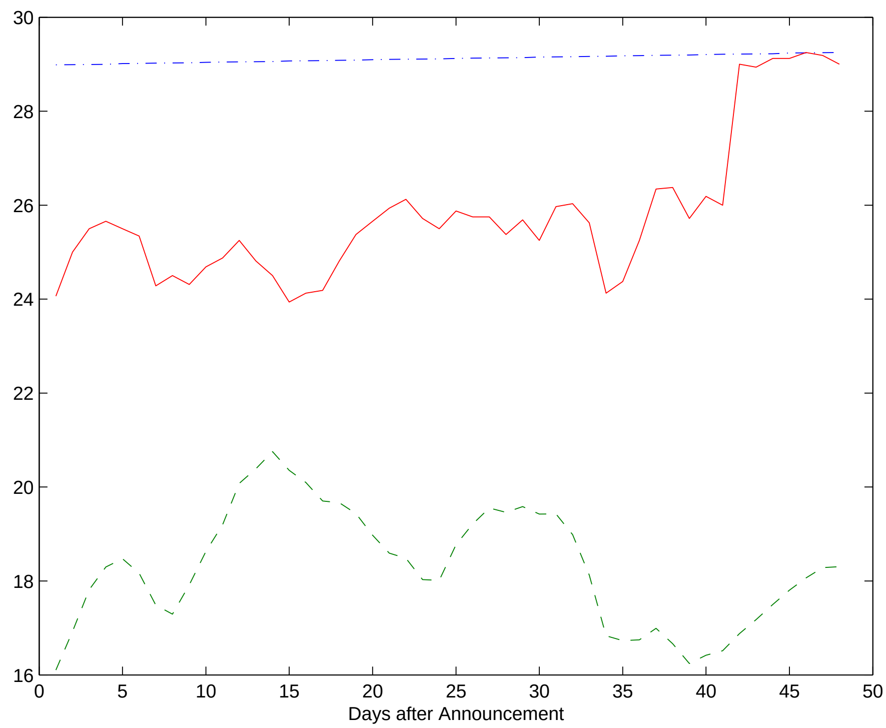


Figure 4: Consider the first deal described in Table 2 for which the merger succeeded (where the target company is PLAT). This figure plots the MCMC draws for a few latent variables, parameters, and model errors. Recall the chosen parametrization for the risk-neutral probability $q(t) = X_1(t)$: $\frac{\mathrm{d}q}{q(1-q)} = \mu_1 \, \mathrm{d}t + \sigma_1 \, \mathrm{d}W_1$, 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 400,000 draws for: (i) X_1 at $t = \frac{T_e}{2}$, where $T_e = 48$ 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.

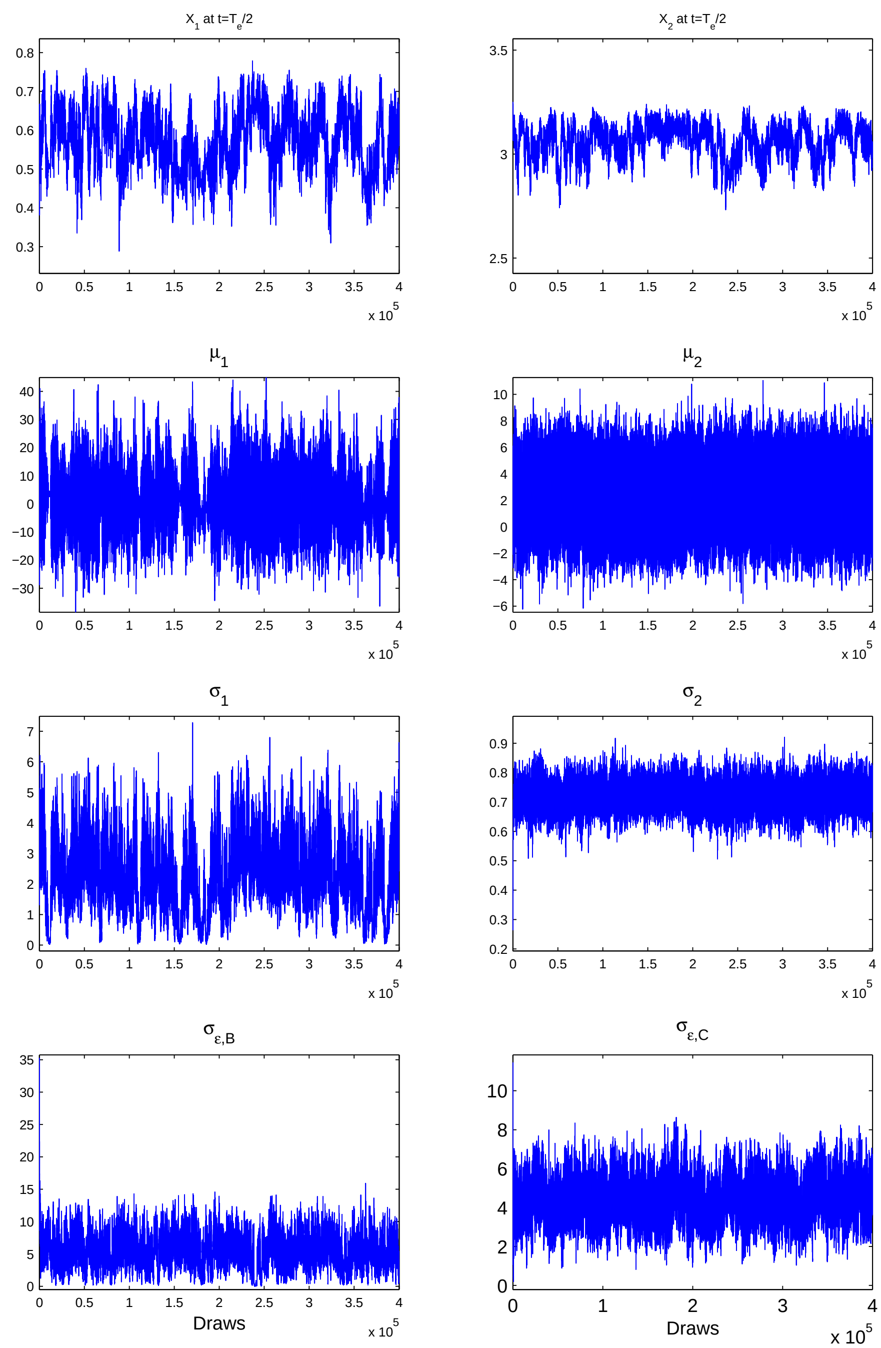
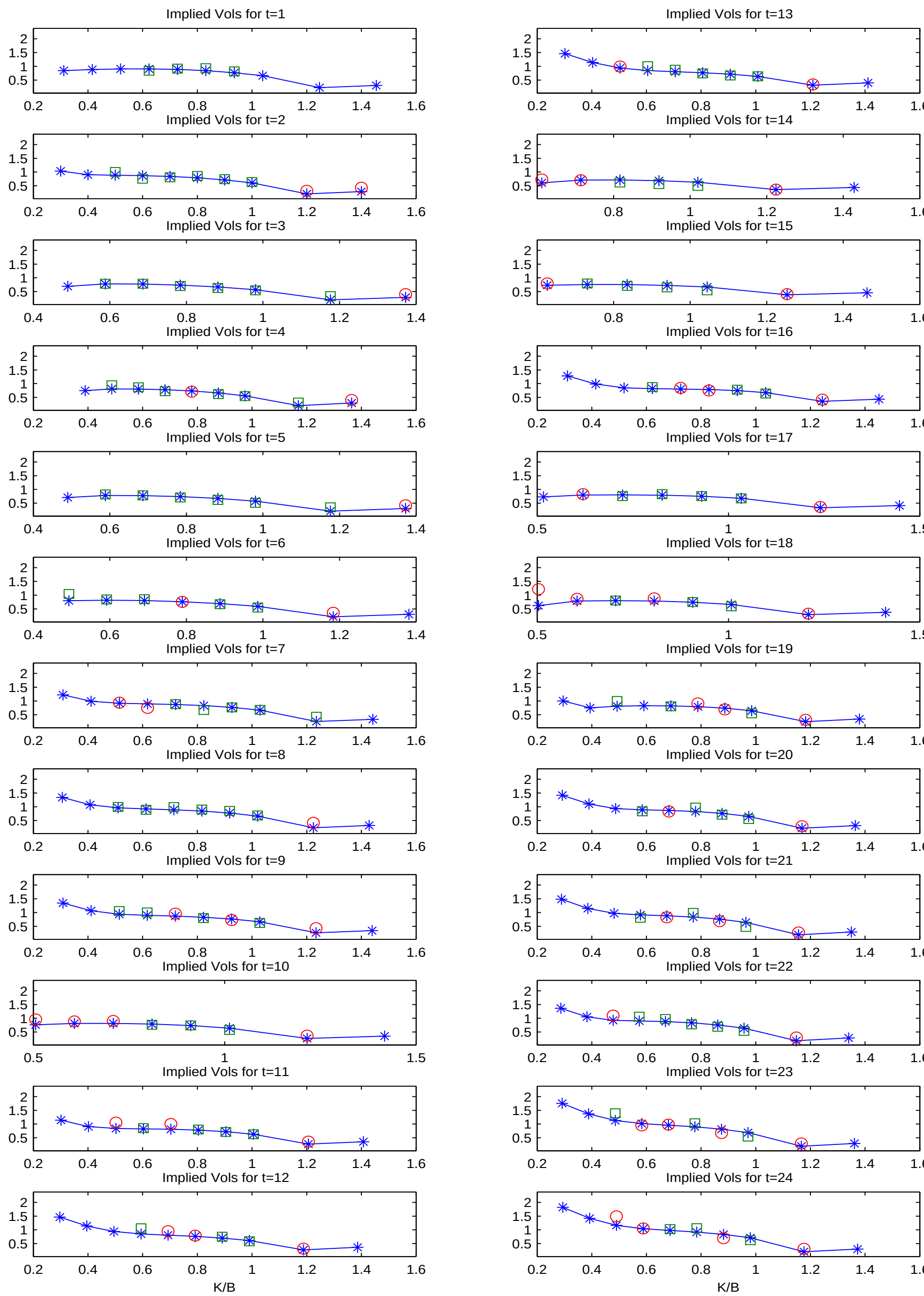


Figure 5: This compares the observed volatility smile with the theoretical volatility smile in the case of PLAT. For each day t while the deal is ongoing, plot the option's Black–Scholes implied volatility against its moneyness (the ratio of strike price K to the underlying price $B(t)$). The Black–Scholes implied volatility for the observed option price is plotted using either a square or a circle: a square for an option with positive trading volume, or a circle 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 the theoretical option price is plotted using a star, and is connected with a continuous line.



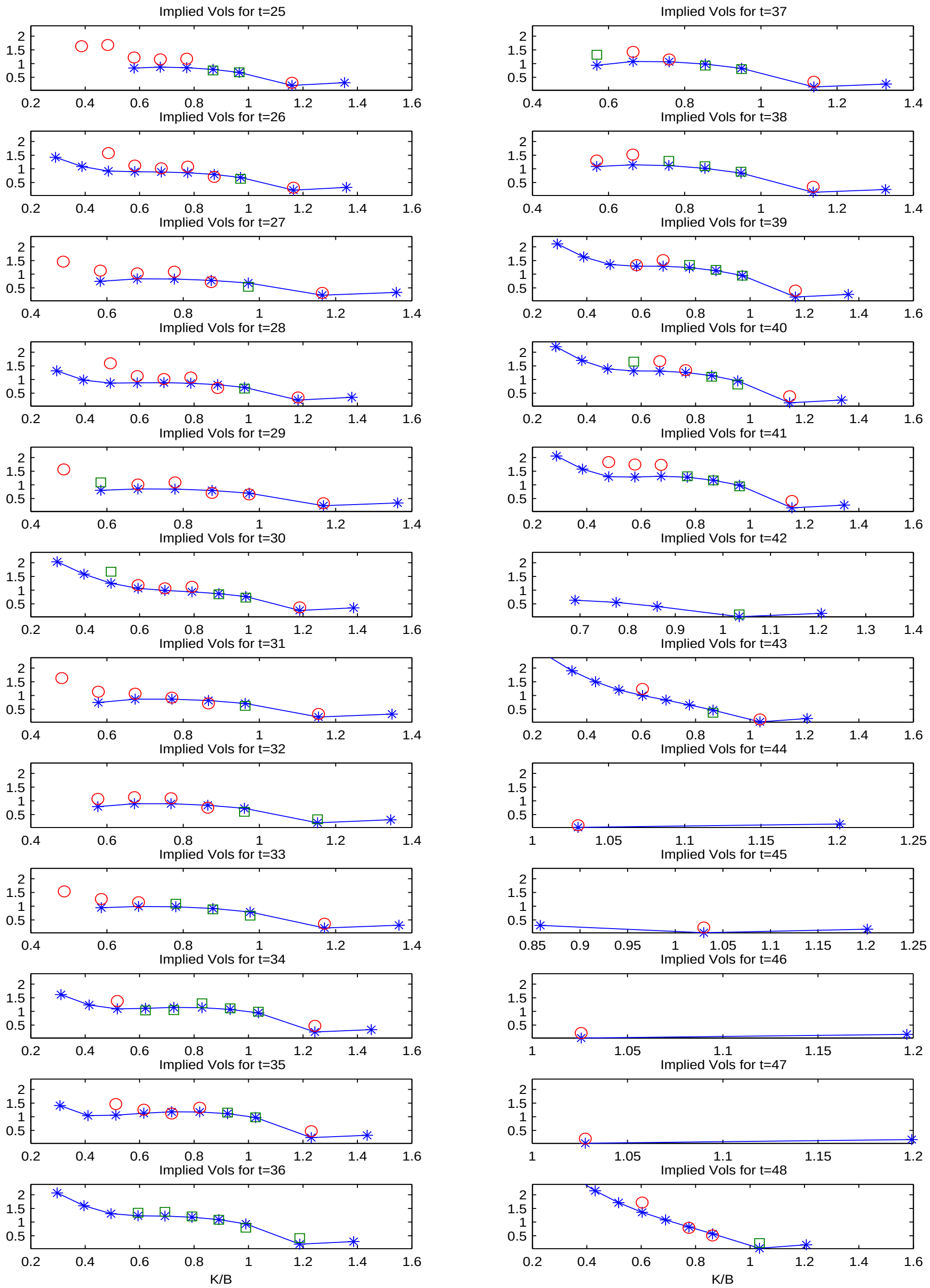
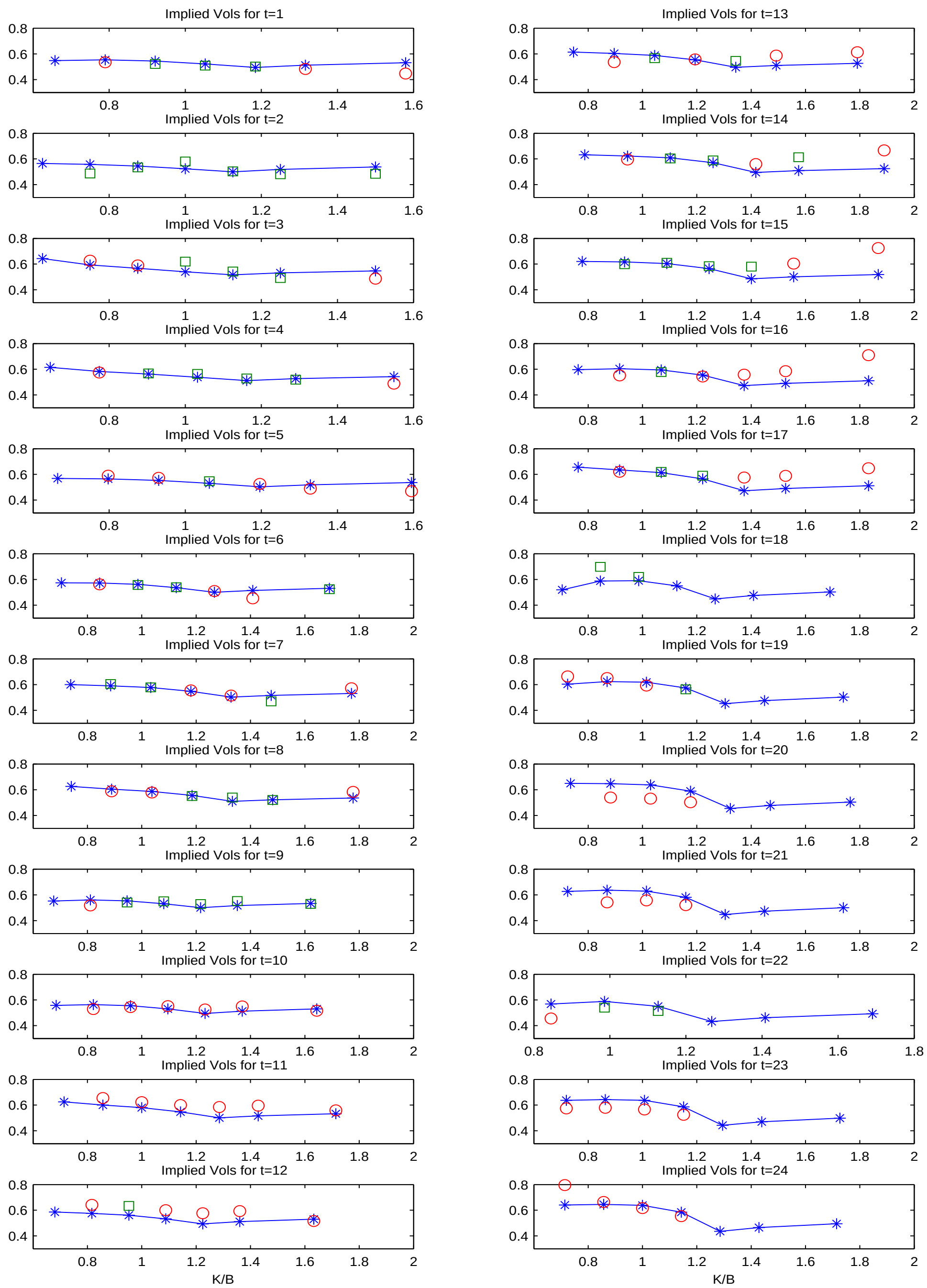
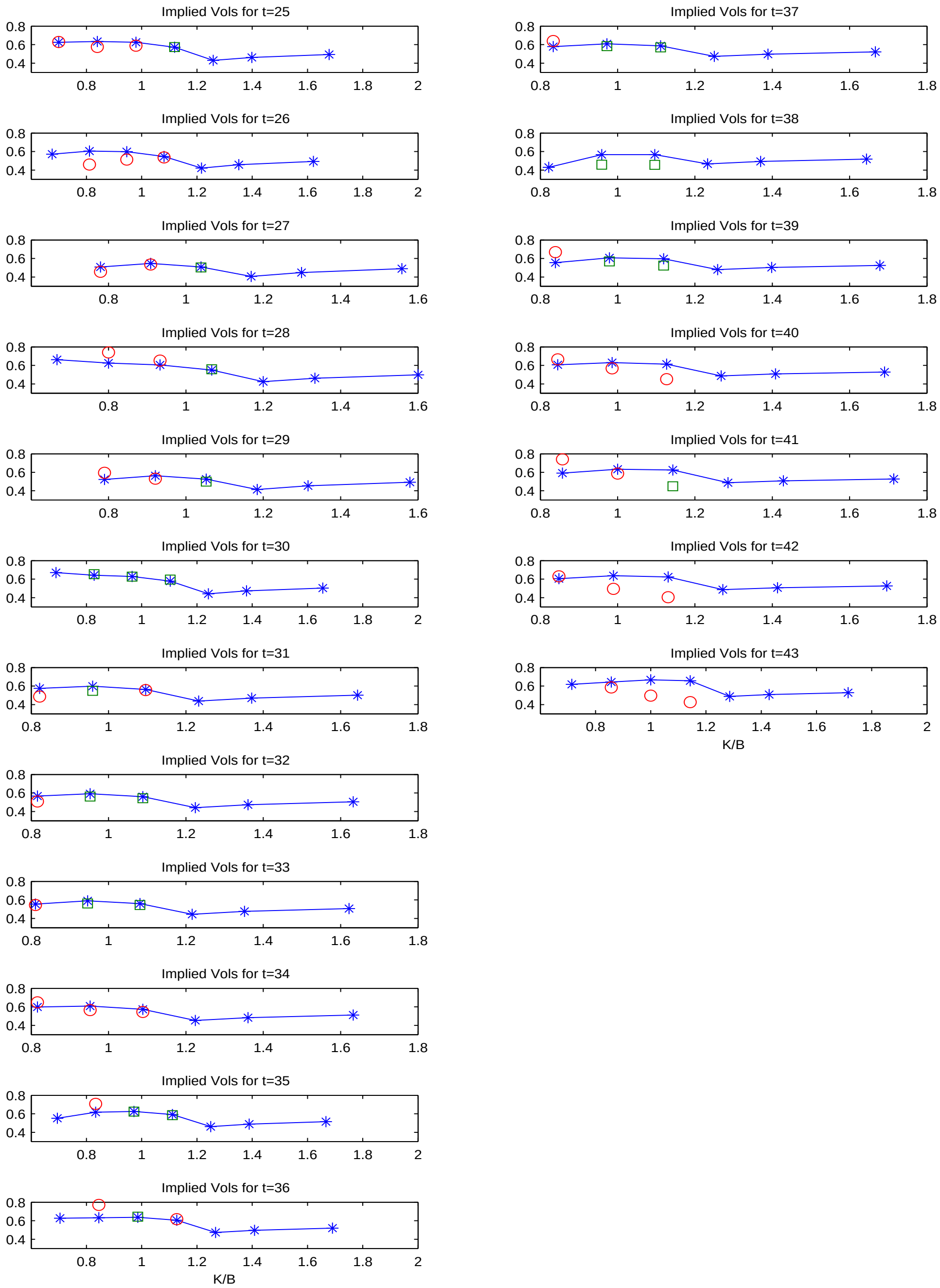

Implied Vols for t=25
Implied Vols for t=26
Implied Vols for t=27
Implied Vols for t=28
Implied Vols for t=29
Implied Vols for t=30
Implied Vols for t=31
Implied Vols for t=32
Implied Vols for t=33
Implied Vols for t=34
Implied Vols for t=35
Implied Vols for t=36
K/B
Implied Vols for t=37
Implied Vols for t=38
Implied Vols for t=39
Implied Vols for t=40
Implied Vols for t=41
Implied Vols for t=42
Implied Vols for t=43
Implied Vols for t=44
Implied Vols for t=45
Implied Vols for t=46
Implied Vols for t=47
Implied Vols for t=48
K/B

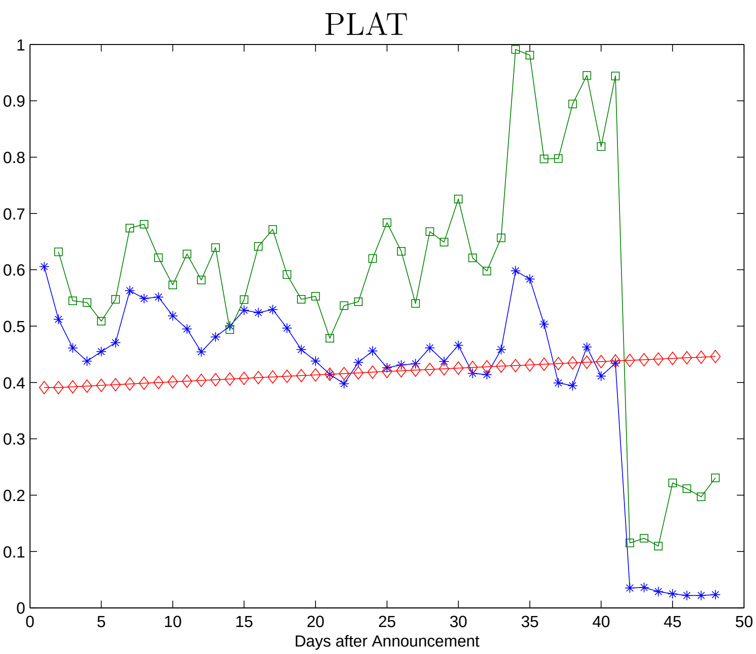
Figure 6: This compares the observed volatility smile with the theoretical volatility smile in the case of KMET. For each day t while the deal is ongoing, plot the option's Black–Scholes implied volatility against its moneyness (the ratio of strike price K to the underlying price $B(t)$). The Black–Scholes implied volatility for the observed option price is plotted using either a square or a circle: a square for an option with positive trading volume, or a circle 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 the theoretical option price is plotted using a star, and is connected with a continuous line.





Implied Vols for t=25
Implied Vols for t=26
Implied Vols for t=27
Implied Vols for t=28
Implied Vols for t=29
Implied Vols for t=30
Implied Vols for t=31
Implied Vols for t=32
Implied Vols for t=33
Implied Vols for t=34
Implied Vols for t=35
Implied Vols for t=36
K/B
Implied Vols for t=37
Implied Vols for t=38
Implied Vols for t=39
Implied Vols for t=40
Implied Vols for t=41
Implied Vols for t=42
Implied Vols for t=43
K/B

Figure 7: This plots the time series of implied volatilities for PLAT and KMET, using: (i) our theoretical model, where the volatility is computed as in Equation (12) (marked on the plot with a star); (ii) the Black–Scholes at-the-money implied volatility (marked with a square); and (iii) the volatility estimated a GARCH(1,1) model (marked with a diamond). All reported volatilities are annualized.



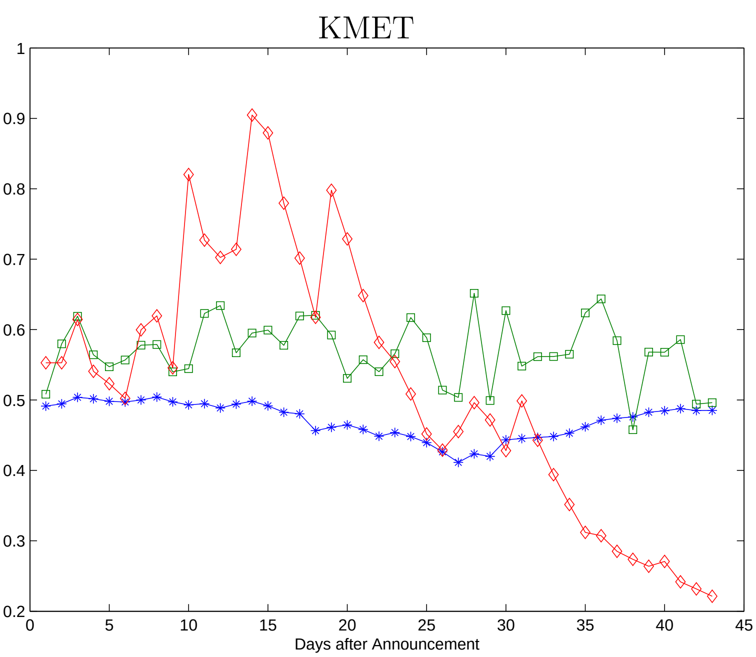


Table 3: Percentage Pricing Errors for the Target Price

This table uses a sample of 65 cash mergers from 1996–2007, 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 65 stocks in our sample.

Percentiles of Percentage Pricing Errors for B

5%	25%	50%	75%	95%
0.00004	0.00009	0.00050	0.00138	0.00559

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

This table uses a sample of 65 cash mergers from 1996–2007, 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 $\left|\left(C^i_{\text{model},K,T,t} - C^i_{K,T,t}\right)/C^i_{K,T,t}\right|$, 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 65 stocks in our sample. Each panel reports the results for: all call options; at-the-money (ATM) calls, i.e., call options with the strike price K closest to the target stock price B_t; near-the-money (NTM) calls with strike K so that $K/B \in [0.9, 1.1]$; in-the-money (ITM) with $K/B < 0.9$; out-of-the-money (OTM) with $K/B > 1.1$; deep-in-the-money (Deep-ITM) with $K/B < 0.7$; and deep-out-of-the-money (Deep-OTM) with $K/B > 1.3$. 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.

Panel A: Percentage Errors for Four Option Pricing Models

		Percentile					
Selection	Model	5%	25%	50%	75%	95%	N
All Call Options	MRB	0.02912	0.10882	0.16154	0.24461	0.46742	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.05838	0.17602	0.35722	0.96760	2.09349	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.07669	0.19985	0.35080	0.88682	2.36116	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.04793	0.07695	0.11263	0.16863	0.28735	65
At-The-Money Calls ($K \approx B$)	MRB	0.01603	0.02915	0.05725	0.08544	0.15182	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.02857	0.05018	0.09491	0.17773	0.52310	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.02324	0.03832	0.06248	0.12889	0.47735	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.02277	0.03786	0.06477	0.12672	0.29553	65
Near-The-Money Calls ($K/B \in [0.9, 1.1]$)	MRB	0.03338	0.13303	0.25659	0.41045	0.67445	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.06132	0.27993	0.67916	1.84527	6.35943	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.04858	0.27483	0.61407	2.16868	9.73362	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05309	0.13740	0.19967	0.31982	0.53904	65
In-The-Money Calls ($K/B < 0.9$)	MRB	0.00297	0.00603	0.00997	0.01821	0.03304	64
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.00335	0.00725	0.01297	0.01910	0.03075	64
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.00330	0.00800	0.01356	0.01914	0.03057	64
	BS: $\sigma = \sigma^i_{K,t-1}$	0.00391	0.00863	0.01265	0.02006	0.03793	64
Out-of-The-Money Calls ($K/B > 1.1$)	MRB	0.28729	0.56222	0.84593	0.99697	1.00000	47
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.44748	0.68115	0.88279	1.00000	5.24072	47
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.41089	0.79160	0.99855	1.78398	10.76406	47
	BS: $\sigma = \sigma^i_{K,t-1}$	0.10257	0.21786	0.37162	0.54960	1.08068	47
Deep In-The-Money Calls ($K/B < 0.7$)	MRB	0.00194	0.00360	0.00529	0.01058	0.01573	55
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.00218	0.00438	0.00624	0.01128	0.02343	55
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.00217	0.00423	0.00635	0.01176	0.02402	55
	BS: $\sigma = \sigma^i_{K,t-1}$	0.00246	0.00424	0.00677	0.01139	0.02641	55
Deep Out-of-The-Money Calls ($K/B > 1.3$)	MRB	0.31266	0.57972	0.88894	0.98496	1.00000	12
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.45174	0.73920	0.97419	1.00000	5.15626	12
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.36695	0.84526	0.99118	1.00000	4.28096	12
	BS: $\sigma = \sigma^i_{K,t-1}$	0.16952	0.31665	0.83283	0.99731	2.20539	12

Panel B: Absolute Errors for Four Option Pricing Models

		Percentile					
Selection	Model	5%	25%	50%	75%	95%	N
All Call Options	MRB	0.04138	0.07188	0.09787	0.13006	0.30882	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.06702	0.10397	0.14760	0.25189	0.58200	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.05791	0.09114	0.14236	0.25460	0.60650	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.04779	0.07780	0.10225	0.14488	0.33302	65
At-The-Money Calls ($K \approx B$)	MRB	0.03919	0.06496	0.10337	0.16754	0.33395	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.06683	0.09862	0.15470	0.24782	0.56985	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.05181	0.07529	0.10255	0.17496	0.71944	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05334	0.07523	0.12872	0.21340	0.56088	65
Near-The-Money Calls ($K/B \in [0.9, 1.1]$)	MRB	0.03218	0.05875	0.09174	0.13225	0.24504	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.08101	0.13269	0.20882	0.42309	1.09082	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.05836	0.10258	0.20501	0.47198	1.18788	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05112	0.06607	0.10136	0.14715	0.41819	65
In-The-Money Calls ($K/B < 0.9$)	MRB	0.04150	0.06881	0.09784	0.14141	0.39043	64
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.04714	0.07392	0.10203	0.17532	0.47980	64
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.04657	0.07560	0.09517	0.16148	0.47650	64
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05130	0.07878	0.10798	0.16415	0.45205	64
Out-of-The-Money Calls ($K/B > 1.1$)	MRB	0.02677	0.05091	0.09833	0.12983	0.38856	47
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.02712	0.06021	0.12253	0.25766	1.36217	47
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.02876	0.06752	0.12500	0.34561	3.16174	47
	BS: $\sigma = \sigma^i_{K,t-1}$	0.01079	0.01753	0.04065	0.12314	0.51624	47
Deep In-The-Money Calls ($K/B < 0.7$)	MRB	0.03106	0.06026	0.08683	0.15052	0.32561	55
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.04128	0.06852	0.09522	0.16169	0.44577	55
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.04123	0.06766	0.09522	0.16207	0.44577	55
	BS: $\sigma = \sigma^i_{K,t-1}$	0.04150	0.06813	0.09940	0.16512	0.44577	55
Deep Out-of-The-Money Calls ($K/B > 1.3$)	MRB	0.02663	0.04750	0.10122	0.24947	3.56104	12
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.02645	0.04177	0.12506	0.62613	5.37421	12
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.02647	0.06257	0.11529	0.55198	5.37421	12
	BS: $\sigma = \sigma^i_{K,t-1}$	0.01743	0.02855	0.04256	0.42121	5.37421	12

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

		Percentile					
Selection	Model	5%	25%	50%	75%	95%	N
All Call Options	MRB	0.19423	0.25527	0.34475	0.45075	0.79646	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.25351	0.42241	0.66947	1.17981	2.38788	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.21405	0.39738	0.68102	1.52722	2.51563	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.20031	0.28160	0.39139	0.53851	1.27496	65
At-The-Money Calls ($K \approx B$)	MRB	0.17355	0.30280	0.44060	0.60543	1.39451	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.31529	0.46620	0.79286	1.36497	2.51171	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.26000	0.35674	0.54181	0.92342	2.66176	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.24809	0.35967	0.55379	1.09694	2.56146	65
Near-The-Money Calls ($K/B \in [0.9, 1.1]$)	MRB	0.22702	0.33859	0.43564	0.64210	1.20514	65
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.33948	0.65118	1.41491	2.67379	10.84737	65
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.27451	0.59057	1.57101	2.97128	9.01710	65
	BS: $\sigma = \sigma^i_{K,t-1}$	0.24326	0.42834	0.61321	0.85886	1.97500	65
In-The-Money Calls ($K/B < 0.9$)	MRB	0.08498	0.15688	0.21632	0.34987	0.55889	64
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.09564	0.16326	0.25078	0.39473	0.82930	64
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.09854	0.17197	0.24845	0.40039	0.74020	64
	BS: $\sigma = \sigma^i_{K,t-1}$	0.12752	0.17231	0.26771	0.42906	0.74958	64
Out-of-The-Money Calls ($K/B > 1.1$)	MRB	0.27862	0.40937	0.49993	0.55338	1.33429	47
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.26147	0.43858	0.50148	1.39626	4.32469	47
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.31300	0.49129	0.63736	1.90042	10.21699	47
	BS: $\sigma = \sigma^i_{K,t-1}$	0.05128	0.14541	0.27020	0.49194	1.73949	47
Deep In-The-Money Calls ($K/B < 0.7$)	MRB	0.07034	0.10956	0.15858	0.26791	0.56942	55
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.08092	0.11874	0.18569	0.32742	0.64397	55
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.07851	0.12013	0.18569	0.34408	0.64397	55
	BS: $\sigma = \sigma^i_{K,t-1}$	0.08274	0.11961	0.19318	0.34278	0.64397	55
Deep Out-of-The-Money Calls ($K/B > 1.3$)	MRB	0.43714	0.47758	0.50000	1.04208	18.59440	12
	BS: $\sigma = \bar{\sigma}^i_{ATM}$	0.30337	0.48671	0.66191	2.75452	25.57853	12
	BS: $\sigma = \sigma^i_{ATM,t-1}$	0.41876	0.49561	0.91420	2.45791	25.57853	12
	BS: $\sigma = \sigma^i_{K,t-1}$	0.08991	0.33690	0.41641	1.93483	25.57853	12

Panel D: Absolute Bid-Ask Spread for Call Options

	Percentile					
Selection	5%	25%	50%	75%	95%	N
All Call Options	0.20205	0.27644	0.40522	0.53865	0.71783	65
ATM Calls	0.13819	0.22438	0.26818	0.35074	0.46719	65
NTM Calls	0.10150	0.18000	0.22825	0.29409	0.43349	65
ITM Calls	0.26199	0.36760	0.54165	0.70033	0.95401	64
OTM Calls	0.06628	0.11977	0.17966	0.25000	0.53844	47
Deep-ITM Calls	0.31643	0.40286	0.65147	0.90737	1.04225	55
Deep-OTM Calls	0.05325	0.09594	0.15313	0.25000	0.46375	12

Panel E: Percentage Bid-Ask Spread for Call Options

	Percentile					
Selection	5%	25%	50%	75%	95%	N
All Call Options	0.08761	0.24089	0.36812	0.61028	0.90122	65
ATM Calls	0.06106	0.07955	0.13212	0.17526	0.26286	65
NTM Calls	0.09419	0.26244	0.66550	0.86436	1.38569	65
ITM Calls	0.02535	0.03366	0.04929	0.06075	0.10959	64
OTM Calls	0.44901	1.62059	1.97403	2.00000	2.00000	47
Deep-ITM Calls	0.02020	0.02667	0.03626	0.04513	0.08339	55
Deep-OTM Calls	0.05835	0.57016	2.00000	2.00000	2.00000	12

Table 5: The Success Probability as a Predictor of Deal Outcome

This table uses a sample of 65 cash mergers from 1996–2007, 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: $q^i_{naive}(t_n) = \frac{B^i(t_n) - B^i_0}{B^i_1 - B^i_0}$, 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.

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	0.40959	0.01122	65
2	0.42164	0.03749	65
3	0.45470	0.05817	65
4	0.58524	0.05891	65
5	0.61426	0.03097	65
6	0.86092	0.04317	65
7	0.75041	0.05056	65
8	0.83286	0.10042	65
9	0.77634	0.21125	65
10	0.74281	0.30956	65

Table 6: The Behavior of the Fallback Price after a Takeover Announcement

This table uses a sample of 65 cash mergers from 1996–2007, 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 $n = 1, \dots, 10$, 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.

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

n	a	b	R^2	N
1	0.44825 (2.384)	0.92538 (16.999)	0.82101	65
2	0.42885 (2.331)	0.93276 (17.508)	0.82951	65
3	0.41050 (2.195)	0.93900 (17.339)	0.82675	65
4	0.36591 (1.900)	0.95262 (17.086)	0.82250	65
5	0.33580 (1.729)	0.96013 (17.074)	0.82230	65
6	0.35364 (1.792)	0.95737 (16.753)	0.81668	65
7	0.34701 (1.825)	0.96182 (17.472)	0.82892	65
8	0.32476 (1.754)	0.97177 (18.133)	0.83920	65
9	0.31421 (1.681)	0.97789 (18.070)	0.83826	65
10	0.38483 (2.048)	0.95981 (17.640)	0.83162	65

Table 7: Comparison of Model-Implied Volatility with the Black–Scholes Implied Volatility and the Volatility estimated with a GARCH(1,1) Model

This table uses a sample of 65 cash mergers from 1996–2007, which have sufficiently liquid options on the target company. For a company i subject to a takeover deal in our sample, consider three ways of defining the instantaneous volatility of the stock price $B^i(t)$ on day t: (1) our model-implied volatility $\hat{\sigma}^i(t)$ given by $\left(\hat{\sigma}^i_B(t)\right)^2 = \left(\frac{B^i_1 \exp(-(T^i_e - t)) - B^i_2(t)}{B^i(t)} q^i(t)(1-q^i(t))\sigma^i_1\right)^2 + \left(\frac{B^i_2(t)}{B^i(t)}(1-q^i(t))\sigma^i_2\right)^2$, 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_1 is the estimated volatility of $q^i(t)$, i.e., $\frac{\mathrm{d}q^i}{q^i(1-q^i)} = \mu^i_1\,\mathrm{d}t + \sigma^i_1\,\mathrm{d}W^i_1(t)$; and σ^i_2 is the estimated volatility of the fallback price $B^i_2(t)$, i.e., $\frac{\mathrm{d}B^i_2}{B^i_2} = \mu^i_2\,\mathrm{d}t + \sigma^i_2\,\mathrm{d}W^i_2(t)$; (2) the Black–Scholes implied volatility $\sigma^i_{BS}(t)$ corresponding to an at-the-money call option at t; and (3) the estimated volatility $\sigma^i_{GARCH}(t)$ of B^i at t, according to a GARCH(1,1)-model. Compute the absolute value of the differences between the three estimated volatilities. Once each difference was computed for each stock i and time t, fix the stock i and compute the mean μ_i over time of the difference, 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 over the 65 stocks in our sample. All reported figures are annualized.

Difference in Estimated Volatilities

	5%	25%	50%	75%	95%	N
$\lvert\hat{\sigma} - \sigma_{BS}\rvert$	0.00983	0.02329	0.04887	0.07827	0.16393	65
$\lvert\sigma_{GARCH} - \sigma_{BS}\rvert$	0.01798	0.02767	0.06435	0.10727	0.25746	65
$\lvert\hat{\sigma} - \sigma_{GARCH}\rvert$	0.00587	0.01389	0.02153	0.05086	0.24578	65