Modeling Merger Arbitrage Situations Using Stochastic Processes

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Acknowledgements

I am sincerely grateful to my brilliant advisor, Chris Rycroft, without whom this thesis simply would not have been possible. His meaningful insights, deep curiosity, and encouragement to delve further truly shaped this document. I am incredibly honored by the amount of hands-on support I have received from him in this endeavor. I would further like to thank my colleague, Tiffany Lee, for her endless guidance and expertise at every stage of this project, from formulating my initial idea to proof-reading my final drafts. I appreciate the many endearing friends who gave their support and encouragement through this process.

I finally thank my family for their immeasurable contributions to my academic journey. It is on their shoulders that I have always stood, and the completion of this work is entirely their accomplishment.

Contents

Chapter 1

Introduction

1.1 Merger Arbitrage

Merger arbitrage (also known as risk arbitrage) is an event-driven investment strategy that exploits market inefficiencies during the course of a corporate merger or acquisition involving publicly-traded companies. After extensive due diligence and closed-door discussions, companies publicly announce their intention to acquire or consolidate by affirming a merger agreement (MGA) or a purchase & sale agreement (PSA). The contract lays out the material terms of the deal, including the type of transaction, purchase price, consideration paid, deal milestones, and expected closing date (the date upon which ownership and payment are officially transferred). The duration of the deal, representing the time period between the public announcement and the closing date, can vary greatly: while pre-approved transactions between small companies can close in a matter of weeks, large companies sometimes spend

years navigating litigious regulatory battles. Upon the announcement of a transaction, the stock of the target company generally begins to trade at some small discount to the offer price dictated in the MGA/PSA. This delta between the offer price and the target company's stock price, referred to as the arbitrage spread, is a factor of the expected duration of the deal and the market's perception that the deal will succeed or fail. Merger arbitrageurs attempt to profit from this spread. If a hedge fund expects a deal to complete, it may purchase shares of the target company hoping to lock in the fixed arbitrage spread, and if it expects a deal to fail, it may short sell shares of the target company stock, expecting a price drop to the pre-announcement level. As the closing date approaches, the target company stock generally approaches the offer price, and the arbitrage spread closes.

1.2 Related Literature

Over the past few decades, a number of studies have shown statistically significant excess returns to merger arbitrage strategies. Larcker and Lys (1987) identified a mean cumulative excess return of 5.32% over the duration of a transaction, Mitchell and Pulvino (2001) found excess annualized returns of 4% accounting for transaction costs, and Baker and Savasoglu (2002) reported average annualized risk-adjusted returns of 9.6%. Several academics have posited various theories to understand the excess returns related to the strategy. Tassel (2016) explained excess returns relative to standard risk models as compensation for liquidity provision and a nonlinear risk profile (arbitrageurs earn a small gain when deals succeed but

accept a large loss when deals fail). As a result of the excess returns, use of the strategy has become widespread, especially in the last twenty years. According to Hedge Fund Research (2008), assets under management of merger arbitrage hedge funds grew from \$233 million at the end of 1990 to \$28 billion by the end of 2007. Jetley and Ji (2010) showed that this uptick in popularity of the strategy has actually been matched by a trend of declining arbitrage spreads. Reduced transaction costs and increased trading in the target companies' stocks following announcement have led to an average arbitrage spread decline of 400 bps since 2002, thus reducing the potential upside for arbitrageurs.

To combat the declining profitability of this strategy, hedge funds and institutional investors have sought increasingly more innovative ways to predict deal outcomes, motivating a body of academic work. Brown and Raymond (1986) provided one of the first frameworks to assess, ex-ante, the probability that a merger will be completed. Though now a fundamental assumption in merger arbitrage, they were the first to show that the market prices of firms involved in a takeover attempt will necessarily reflect the attitudes of the investing public, and thus the post-announcement price of the target firm could be used to infer the probability that an acquisition will ultimately succeed. Barone-Adesi, Brown, and Harlow (1992) further showed that arbitrageurs in the market for takeover candidates behave rationally, though with less-than-perfect foresight. Hoffmeister and Dyl (1981) performed the first of many regression-based analyses to identify important variables in takeover success: their multivariate discriminant analysis found that the firm size and the management's reaction to the takeover are the two most important factors in determining the success of a

tender offer. Further studies by Jindra and Walking (2003) found arbitrage spreads to be greater for failed deals than successful deals, and identified that deal success is negatively correlated with price revisions and positively correlated with the deal duration. Branch and Wang (2003) built on Jindra and Walking's work by only considering ex-ante variables, and established a number of significant predictors of success: bidder's return volatility, bidder's systematic risk (beta), transaction cost, arbitrageurs' activity, relative size of target, and bid premium. Branch and Wang (2009) constructed a takeover success prediction model using a weighted logistic regression approach and found that the target's stock price run-up, resistance, arbitrage spread, and bidding competition also play significant roles in predicting a deals outcome. Lin, Lan, and Chuang (2012) proposed an option-based approach to improve prediction accuracy of deals in emerging markets, where data on takeover attempts are typically either unavailable or of poor quality, and achieved considerably higher accuracy than comparable qualitative regression models. Most recently, Tassel (2016) developed a dynamic asset pricing model that exploits the joint information in stock and option prices to forecast deal outcomes.

Due to the convenience afforded by the large population of historical M&A transactions, the majority of merger arbitrage literature has focused on the construction and analysis of logistic regression models. While there has been an overwhelming focus on the qualitative factors that may indicate deal success, few attempts have been made to quantitatively understand or model the exact dynamics of the target company stock price through the course of a transaction. Lu and Yau (2008) and Wilfling and Gelman (2009) have both explored the

use of stochastic processes to represent the target stock. These studies, however, use only geometric Brownian motion processes to derive risk-neutral distributions of the future stock price, in order to price European call options using Black-Scholes. Less consideration has been given to stochastic processes as a means of actually predicting the overall likelihood of deal success.

1.3   Stochastic Processes in Finance

A stochastic process is a sequence of random variables indexed by a set of numbers or points in time. Stochastic processes have long been used as mathematical models to represent real world phenomena that appear to follow random behavior, with applications in disciplines ranging from ecology to cryptography. The application of stochastic processes to finance, and more specifically asset prices, hinges on the theory of the efficient market hypothesis (EMH). First developed by Fama (1970), the efficient market hypothesis assumes 1) that the past history of a stock is fully reflected in the present price, and 2) that markets respond immediately to new information. Thus, the task of modeling asset prices is fundamentally about modeling new information. Changes in asset price follow a Markovian process, and if we assume no bias in price movement, they can also be described as a Martingale process. Although stochastic processes in finance are built using a history of asset price movements, it is important to note that exact price paths cannot be forecast. Instead, these models are used to examine the probable distribution of future asset prices. Stochastic processes

in finance gained widespread notoriety when Black and Scholes (1973) introduced a new methodology for the valuation of European call and put options, which can be generalized to other financial instruments (also referred to as contingent claims). The Black-Scholes model was built on the assumption that stock prices follow a geometric Brownian motion process with constant drift and volatility.

Beyond just Black-Scholes, the majority of stochastic processes in finance are built upon the assumption that short-term stock returns fit normal distributions: the Gaussian is easy to work with, has defined higher moments, and is well studied. Investors are often wary of fat-tailed distributions because of the additional risk that they suggest. Under a normal distribution, tail risk (the likelihood of an asset moving more than 3 standard deviations from its current price) is less than a thousandth of a percent. While the returns of well-behaved price data can be accurately represented by a normal distribution, stocks in the real-world often experience large, unexpected shifts in price. In the context of mergers and acquisitions, target stocks often skyrocket on the announcement of competing bidders and frequently tank on the announcement of DoJ or FTC regulatory intervention.

In this paper, making assumptions similar to EMH, we assume that the stock price of a target company in a transaction can be modeled by a stochastic process. We present two novel approaches to predict M&A deal outcomes. In the first approach, using a large dataset of historical transactions, we construct and test various stochastic models that accurately represent target stock price movement in completed transactions. We provide model calibration techniques, and, using Monte Carlo methods, show that the models can be fit to ongoing

deals to predict likelihood of success. In the second approach, we build a generalized random walk model of target stock prices during a merger and present computational approaches to determine success probability given deal-specific parameters. In both approaches, diverging from past merger arbitrage literature, we extend traditional stochastic models to allow returns from fat-tailed distributions, and compare their effectiveness to classical Gaussian models.

Chapter 2

Data Overview

2.1 Data Sources

We obtained data for our study from Thomson ONE, S&P Capital IQ, and Bloomberg's M&A database (accessed via Bloomberg Terminal). We compiled an initial dataset of transactions that were completed or terminated between 1/1/1998 and 12/31/2015. Our sample included only cash transactions, as the per share offer price can be difficult to discern in stock or cash & stock deals, where the offer price is a function of the acquiring company's stock price. Deals involving convertible debt or preferred stock were excluded because the market value of such considerations is generally difficult to obtain. We further limited our sample to only include deals in which the target company was publicly traded, as merger arbitrage is impossible in transactions involving only private companies. Additionally, we narrowed our dataset to only include deals in which the target company was publicly traded on a

US exchange. Deals involving foreign companies are likely to involve foreign currency risks and transaction costs that are difficult to estimate. Furthermore, since hedge funds that employ merger arbitrage strategies rarely take positions in micro-cap companies (generally due to liquidity constraints), we narrowed our dataset to only include deals in which the total transaction value exceeded \$250 mm. Our final sample of 2,787 transactions included acquirer ticker symbol, target ticker symbol, deal announcement date, deal completion date, cash terms (offer price per share), and final deal status (completed or terminated).

Fitting a model to the data required a standardized price history for each transaction. For each deal, we pulled the daily closing price of the target stock for each day between the announcement date and the completion/termination date. As the data for each transaction began the day after announcement, the price history did not include the initial surprise[1]. We then scaled all price histories to the same time series, $t \in [0, 1]$, such that the price of a target stock (defined by the variable S_t) at time $t = 0.5$ reflects the price of the stock exactly halfway through the deal. This normalization allowed us to easily work with data of inconsistent lengths (given each deal has a different duration). We also down-sampled our price data for each target stock to, at most, 101 data points over the range $[0, 1]$, such that the smallest Δt was 0.01. Lastly, we standardized all price levels by normalizing the daily prices of each deal to the offer price for the deal. For a deal that successfully completed, the price of the stock on the last day was typically at or near the offer price (i.e. $S_1 = 1$).

---

[1]Because acquirers typically offer target shareholders a premium to the existing stock price, the announcement of a deal is usually accompanied by a significant spike in the target share price (referred to as the 'initial surprise'). While other studies have investigated the impact of the magnitude of the surprise on deal outcome, we are concerned only with price movements during the actual transaction.

Figure 2.1 and Figure 2.2 demonstrate the effect of this standardization process on two target stocks. As both the expected completion date and the offer price are included in the initial deal announcement, we see no issue with this standardization process.

Figure 2.1: Price chart of two target stocks (EMC and PWRD) from announcement to completion. Both transactions are successful, yet the data is difficult to compare.
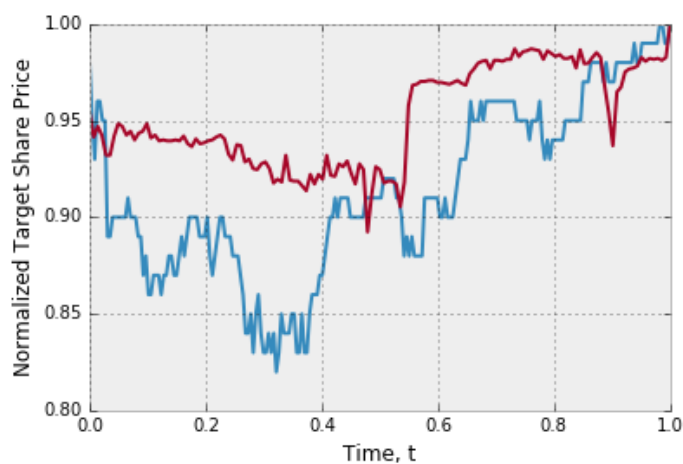

Figure 2.2: Price chart of two target stocks (EMC and PWRD) with normalized share prices and scaled time axis. By standardizing the axes, we can easily compare the arbitrage spreads at any portion of the price history.

2.2 Summary Statistics

Table 2.1 presents a summary of the 2,787 cash mergers included in our data sample. Of these mergers, 2,366 were completed transactions and 421 were terminated, representing a completion rate of 81%. This proportion is consistent with recent studies, including Christensen et al. (2011) which estimated that 70-90% of announced deals are successful. Our dataset had relatively fewer transactions announced before 2004, mostly due to the lack of accessible information.

We found that the average merger arbitrage return for a completed transaction was around 2.8%, where the return is defined as $R = \frac{S_1 - S_0}{S_0}$. We did not identify a trend of declining merger arbitrage spreads for cash deals in the past 15 years, as was suggested by Jetley and Ji (2010) (though their dataset covered both cash and stock deals from 1990 to 2007). For transactions that were terminated, we found that an investor who bought the target stock on the day of the announcement, and sold on the day of termination, would have experienced a substantial loss between 8% and 25%. Additionally, our dataset showed that completed deals have an average duration of 83 days, slightly larger than the average duration of 74 days for terminated deals. We also found that, in a given year, the total number of deals was roughly correlated with general macro-economic conditions in the US, as expected.

Table 2.1: Number of Deals, Average Arbitrage Return, and Average Deal Duration by Year

| Year | Completed | | | Terminated | | Total No. of Deals | % Completed |
	No. of Deals	Arb Return	Avg. Duration	No. of Deals	Avg. Duration		
1998	10	2.0%	77	5	55	15	66.7%
1999	8	1.3%	64	8	53	16	44.0%
2000	33	4.2%	118	9	71	42	75.0%
2001	68	2.3%	70	8	78	76	86.7%
2002	78	4.3%	87	3	40	81	88.1%
2003	79	2.7%	82	11	76	90	79.8%
2004	135	2.8%	90	19	74	154	84.6%
2005	210	2.0%	73	22	68	232	85.2%
2006	291	2.9%	86	55	62	346	79.0%
2007	329	2.6%	81	64	85	393	81.9%
2008	156	2.5%	79	38	65	194	77.0%
2009	88	4.9%	69	11	59	99	83.6%
2010	154	2.2%	84	25	78	179	83.4%
2011	158	2.3%	77	27	88	185	83.9%
2012	154	2.3%	86	29	84	183	82.1%
2013	126	2.3%	82	23	101	149	82.6%
2014	129	2.6%	92	33	90	162	79.1%
2015	160	4.5%	94	31	113	191	79.7%
Total	**2,366**	**2.8%**	**83**	**421**	**74**	**2,787**	**81.4%**

Notes: 'Arb. Return' reflects the return to an investor who purchased the target stock on the day of announcement and held to completion. 'Avg. Duration' is the average number of days between deal announcement and deal completion/termination.

2.3 Preliminary Data Analysis

Before constructing any stochastic model, we require a complete sense of how target stock prices move in the course of a transaction. In this section, we perform qualitative and quantitative analyses on completed and terminated deals in our sample to develop this understanding. Figure 2.3(a) shows a random sample of completed transactions. The price

movement shown here is highly representative of target stocks in completed deals: post-announcement trading generally begins in a tight band between 95% and 100% of the offer price (average of 97.41% across all completed deals). Though uncommon, if the market has conviction that other bidders may be involved in the process, or that the acquirer will offer a higher bid, the target stock may even trade above the offer. All target stocks eventually mean revert to the offer price, yet short term movements seem random and uncorrelated. Price movements appear especially erratic in the beginning quarter of the transaction; we understand this as the market reacting to large amounts of new information and attempting to estimate the deal outcome. Thereafter, average step sizes ($\frac{dS}{dt}$) tend to decrease. Similarly,

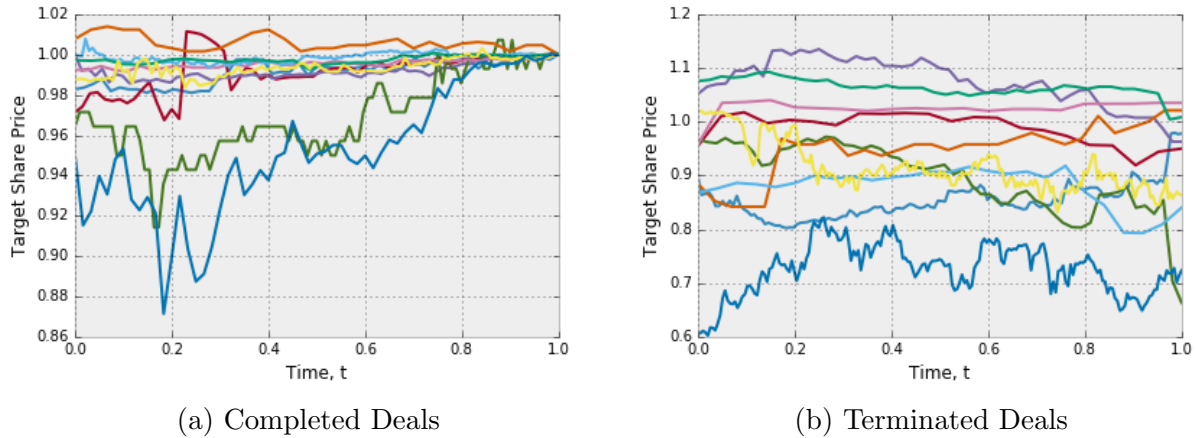(a) Completed Deals   (b) Terminated Deals

Figure 2.3: A random sample of 10 completed and 10 terminated transactions.

Figure 2.3(b) shows a random sample of terminated transactions. The initial distribution of prices is noticeably broader for these stocks, ranging from 60% to 120% of the offer price (average of 93.92% across all terminated deals). The fact that the arbitrage spread on day one tends to be substantially larger provides an initial indication that the market is uncertain

about the outcome. Unlike for completed deals, we see that the volatility of the processes does not necessarily decrease over time. Furthermore, there is no obvious long term trend in the price movements; the stocks tend to remain flat for the majority of the duration and may increase or decrease sporadically only at the very end. This trend is captured very well in Figure 2.4, which depicts a price chart for the average of all completed deals and the average of all terminated deals. Averaging over the dataset eliminates the stochastic element of the process, but reveals the positive bias for completed deals and the typical price drop toward the end of terminated deals. This suggests that the market receives some indication that a deal is likely to fail in the last quarter of the transaction. We found that, on average, target stocks are trading at 89% of the offer price on the termination date.

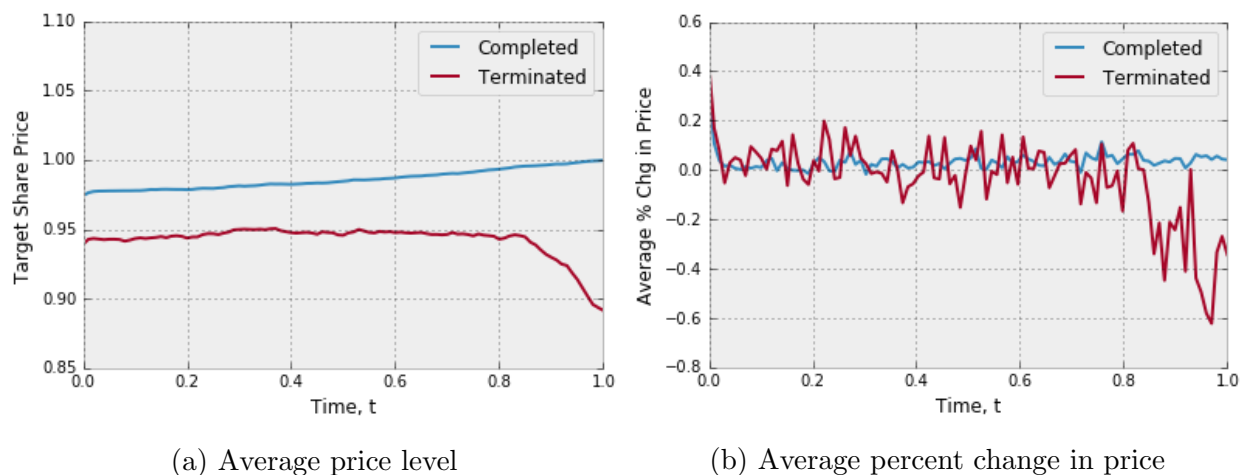(a) Average price level   (b) Average percent change in price

Figure 2.4: Average price level and percent change in price across all deals.

We next investigate the volatility of each process in further detail. Figure 2.5 shows how the variance across all deals develops with time. Fitting the variances for completed deals to a least squares regression line reveals that completed deals have a linear expected variance.

This means that the price charts of completed deals disperse at a constant rate over time. As we discuss later, linear expected variance is a feature of Brownian motion, in which the variance of a process X can be described as $\text{Var}(X) = \sigma^2 t$. It is worth noting, however, that the slope of the LSR line was very small, suggesting near zero increase in variance across deals over the duration. Terminated deals seem to follow a similar linear trend for the first 80% of the deal duration, only with a higher volatility σ. In the last portion, however, volatility spikes upwards in exponential fashion.

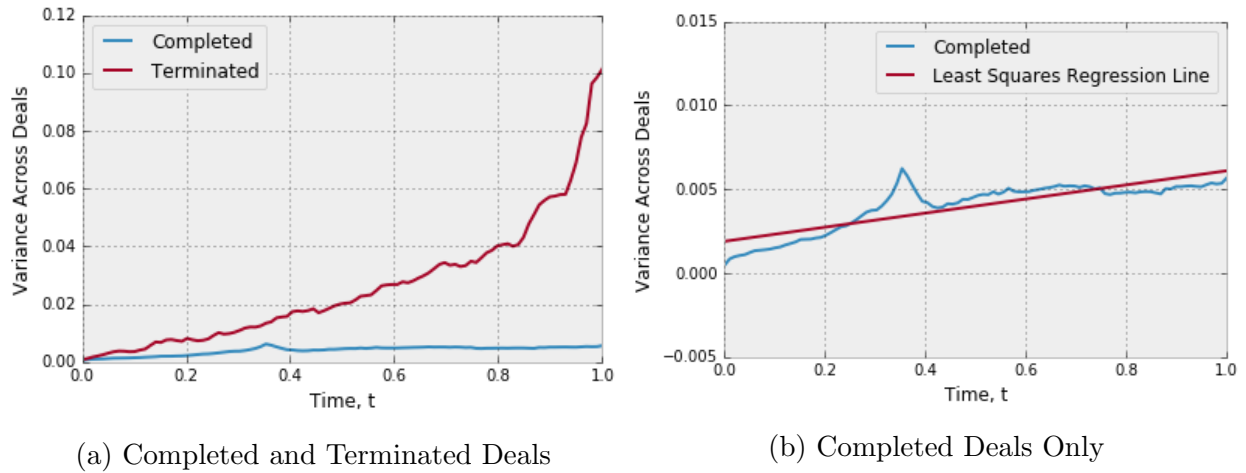(a) Completed and Terminated Deals   (b) Completed Deals Only

Figure 2.5: Variance across all deals at time = t.

We also wish to understand how the variance in price movement changes for different time-steps. Figure 2.6(a) shows the variance of changes in price for all time steps of a given size dt across the entire dataset. We find that target stocks in both completed and terminated deals exhibit a steep exponential decay in price variance as time-steps grow. This supports our initial assumption that, in the short-term ($dt < 0.1$), price movement is stochastic and highly variable, while over a longer horizon, dispersion in stock prices is

minimal. In Figure 2.6(b), we present the same data on a logarithmic scale and also find

that, 1) variance is consistently larger for stocks in terminated deals, and 2) variance actually

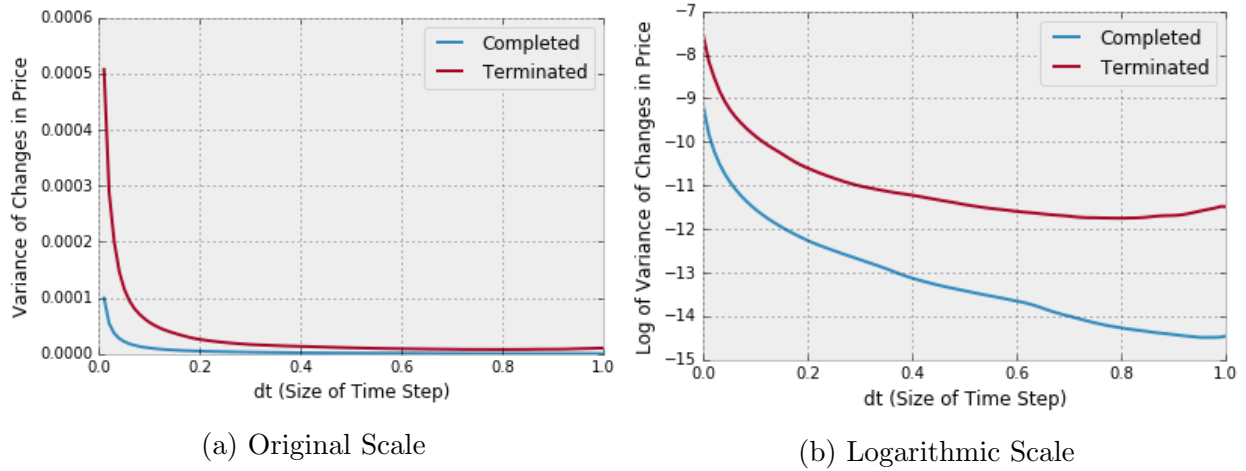increases slightly towards the end of a terminated deal.

(a) Original Scale   (b) Logarithmic Scale

Figure 2.6: Variance of change in price (Var(dS)) for all time step sizes dt.

Proceeding on the assumption that short-term price movements can be treated as ran-

dom, we attempt to identify a probability distribution that accurately represents the data.

Figure 2.7(a) presents a histogram of all instantaneous (i.e. $dt = 0.01$) price changes across

the dataset of completed deals. In Figure 2.7(b), we present the same data on a logarithmic

scale y-axis. The distribution of price changes for terminated deals was almost identical. As

we postulated in the introduction, the histogram appears to be fat-tailed and leptokurtic.

While the majority of values for dS are at or near zero, the range of values is quite large.

This makes sense given our understanding of M&A transactions. During the transaction, as

the market awaits information, stock price changes of the target company are generally very

small - often exhibiting even lower volatility than stocks outside of transactions, which have

(a) Original Frequency Scale
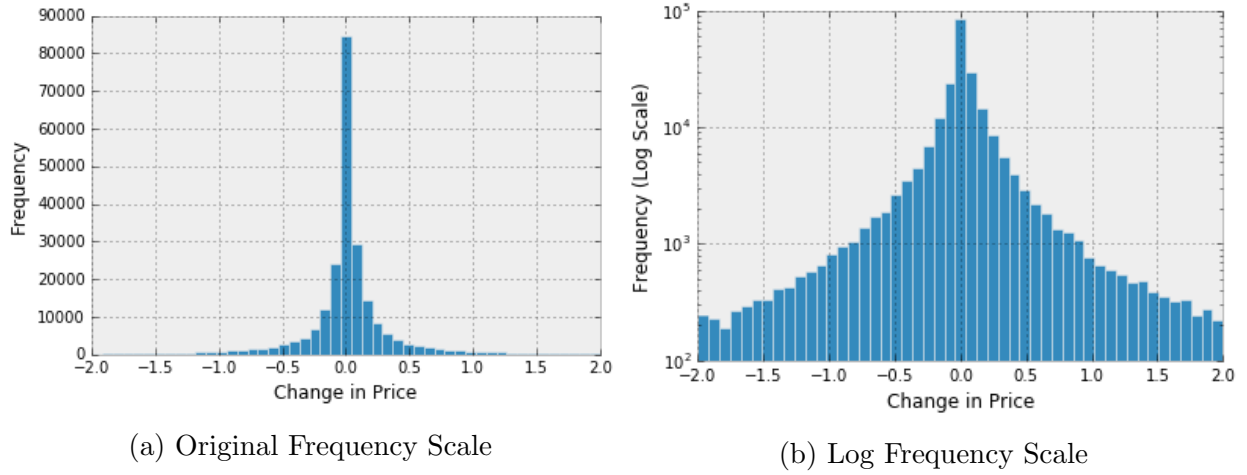
(b) Log Frequency Scale

Figure 2.7: Histogram of all instantaneous price changes (i.e. $dt = 0.01$) for completed deals.

greater exposure to broader market movements (Mitchell and Pulvino (2001)). However, when transaction-related news is released (e.g. regulatory status update, shareholders' vote, etc.), price movements are often large tail events. Figure 2.8(a) shows a histogram of a dataset of randomly sampled values from a Cauchy distribution, with logarithmic frequency scale (the sample is equal in size, center, and IQR to the data shown in Figure 2.7). The Cauchy is a common fat-tailed distribution that we discuss in-depth later on. For comparison, Figure 2.8(b) shows a dataset of randomly sampled values from a standard Gaussian distribution (also on a logarithmic frequency scale, with the same size, center and IQR as the previous histograms). It is immediately clear that the price changes in our dataset are more accurately represented by the Cauchy. The Gaussian gives higher probabilities to intermediate price movements than we see in in the actual data. Attempting to deal with this issue by reducing the variance of the distribution eliminates any potential for large price fluctuations or tail events.
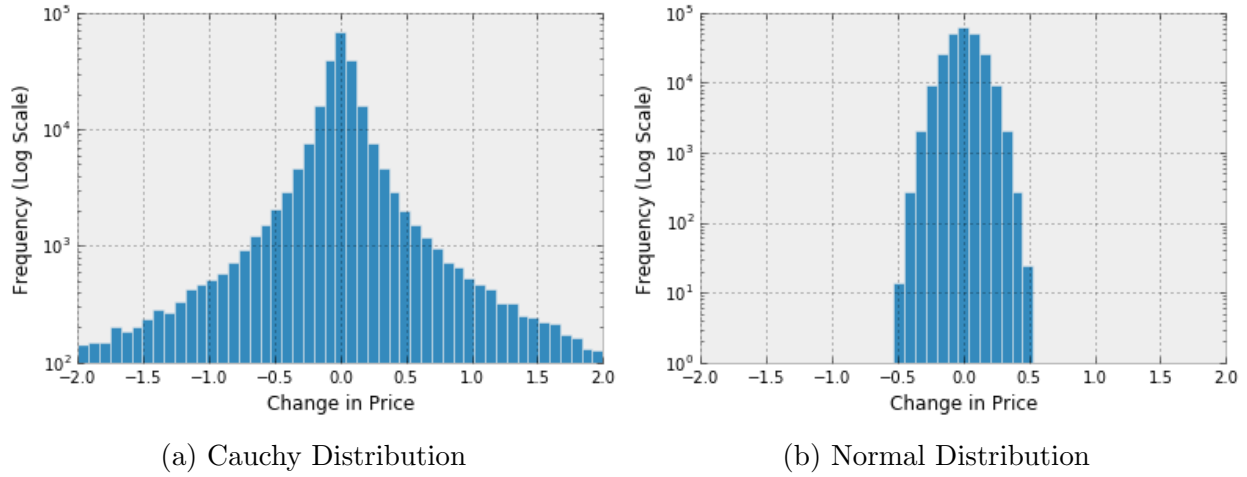
(a) Cauchy Distribution   (b) Normal Distribution

Figure 2.8: Log frequency scale histograms of randomly sampled values for dS/dt from a Cauchy and a Normal distribution (equal in size, center, and IQR to the sample shown in Fig. 2.7).

We next attempt to understand how the distribution of dS/dt (which we now know to be fat-tailed) changes with time. Figure 2.9 presents histograms of all instantaneous price changes at five different time points in the data. While the overall shape of the distribution stays the same, we do see that kurtosis increases over time for completed deals. Figure 2.10, which plots the interquartile range of the instantaneous price change (dS/dt) distributions over time, validates this finding[2]. This finding suggests that the volatility of short-term random movements in completed transactions actually declines over time; in other words, we are less likely to see large, one-off price fluctuations as the deal progresses. For terminated transactions, however, the distributions have noticeably wider spreads throughout the deal duration. Furthermore, roughly halfway through the transaction, we see that the distributions actually begin to widen, suggesting the probability of large price fluctuations

---

[2]We use interquartile range as our measure of spread, because variance is undefined for the Cauchy distribution and other stable fat-tailed distributions.

(a) $t = 10$ (b) $t = 30$ (c) $t = 50$

(d) $t = 70$ (e) $t = 90$
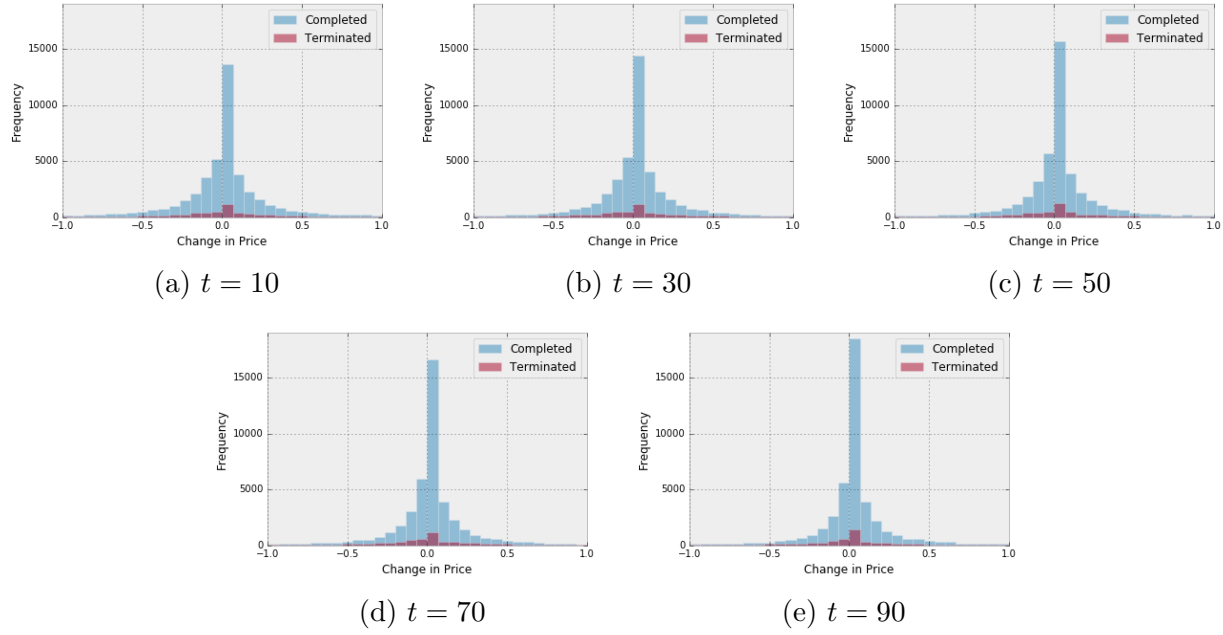
Figure 2.9: Histograms of instantaneous price changes (dS/dt) at $t = 10, 30, 50, 70, 90$.

increases. At the very end of the terminated transactions, we see that the spread narrows again, suggesting that the market has already priced in deal termination.
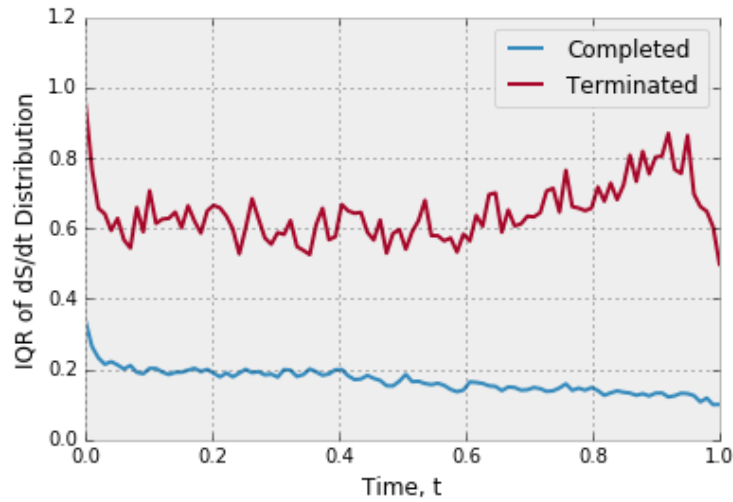

Figure 2.10: IQR of dS/dt distribution over time, representing the likelihood of tail-events or large price fluctuations.

2.4 Takeaways

Before constructing models, we performed a series of analyses on our sample of 2,787 transactions in order to better understand the data. We identified two principal forces acting on all target stocks: 1) a long-term mean-reversion effect, and 2) short-term stochastic price movements. We believe that a representative stochastic model must take both of these processes into account.

Across the duration of the deal, target stocks in completed transactions, on average, tended to move at a constant rate towards the deal offer price. Consequently, the variance in price level across all completed deals stayed essentially flat. For terminated deals, target stocks remained roughly constant until the last quarter of the deal duration, at which point prices tended to drop very quickly. This price drop likely reflects early information that the deal would fail. For terminated stocks, variance in price across all deals rose linearly for the majority of the transaction, at a rather steep rate. Towards the end of the transaction, variance grew exponentially.

By comparing $\text{Var}(dS)$ for different time step sizes, dt, we found that short-term target stock price movements can be treated as random. We were able to show that instantaneous price movements actually draw from a fat-tailed distribution, rather than the Gaussian distribution assumed by previous literature. This finding fits our understanding of M&A transactions: target stocks in transactions are not well-behaved and are subject to large fluctuations with the introduction of new information pertinent to the deal outcome. We

compared the distribution of short-term price changes (dS/dt) across the duration of the deals, and found that the distribution narrows at a constant rate for target stocks in completed deals. This suggests that the volatility of the random movement decreases over time and large shifts become less probable. On the other hand, for target stocks in terminated deals, the probability of tail events increases with time.

Chapter 3

A Model for Target Stocks in Completed Transactions

3.1 Overview

We now present our first approach to estimating the probability that a transaction will successfully complete. Across this chapter, we introduce five potential stochastic processes that can be used to represent the price of a target stock in a completed transaction. We introduce each model by assessing its various properties and considering its application to merger arbitrage. We then discuss calibration techniques and present results from model testing. In testing our models, we assumed the role of an arbitrageur, and used the following Monte Carlo method to estimate predictive power and robustness:

1. For each completed and terminated deal in our sample, we calibrated the model at 10 different uniformly distributed points in time $(t \in \{0.1, ..., 0.9, 1.0\})$[1]. A robust model would require less data to make an accurate determination deal outcome. As we showed in Section 2.3, arbitrage spreads decline over the course of the transaction, and thus, an arbitrageur maximizes her return by identifying the outcome as early as possible.

2. For each deal and time range, we then used the calibrated model to produce n simulated price path projections, which go from the last available date and price to the completion date. We set $n = 50^2$[2].

3. We established a test-statistic that represents the portion of simulated price paths that close at the offer price (thus representing a completed deal). The test-statistic reflects the simulated probability that the deal will complete, defined mathematically as $z \in [0, 1]$ such that

$$z = \frac{\sum\limits_{i=1}^{n} p_i}{n} \tag{3.1}$$

where p_i is given by

$$p_i = \begin{cases} 1 & \text{if } |S_{1,i} - K| < \epsilon, \\ 0 & \text{if } |S_{1,i} - K| \geq \epsilon, \end{cases} \tag{3.2}$$

---

[1]We did not calculate probability of completion at each point in time due to computational burden.
[2]We found from testing that increasing the number of simulations per deal does not significantly affect the test statistic.

in which K represents the offer price of the given deal and $S_{1,i}$ represents the final

price of a given simulation i. Though we were working in a discretized time field, the

random variable for price took on a continuous range of values. Thus, we considered

any simulations for which the final price satisfied $|S_{1,i} - K| < \epsilon$ to be completed deals.

We set $\epsilon = 0.01$. Our reliance on this threshold value to determine the deal outcome

is a limitation of this study, but we do not believe that it is unrealistic; analysis of the

transactions in our dataset found that less than 2% of terminated deals closed within

a 1% range of the offer price, while nearly all completed deals closed in the same

range. We thus believe that thresholds at $\pm 1\%$ of the offer price create a range that is

large enough to catch simulations of completed deals, yet narrow enough to limit false

prediction of terminated deals. It is worth nothing that we treated all simulations for

which the final price was $S_{1,i} > 1.01$ as terminated. In our dataset, the majority of

situations where the final price exceeded the offer price were in terminated transactions

involving competing bidders. In these situations, a merger may have occurred at a later

date with another acquirer, but the originally announced deal did not complete. Since

many arbitrageurs match long positions in the target stock with short positions in

the acquirer stock, it is important for our model to still consider these simulations as

terminated, even if there may have been a positive return on the target stock over

the deal duration. In sum, the test-statistic z for a given deal represents the portion

of simulations that can be considered completed deals (ex. if for a given deal and

time range, $z = 0.4$, we know that 20 out of 50 simulations closed in the price range

$S_{1,i} \in [0.99, 1.01])$.

4. We then aggregated and analyzed simulated probabilities for all completed deals and for all terminated deals, at each point in the time range. In order to determine whether a deal would complete, we set a cutoff for the test-statistic. We determined this by comparing distributions of test-statistics for completed and terminated deals, and identifying cutoffs that minimized type II error while maintaining predictive power.[3]. As an example, if our cutoff is 25%, we would predict a successful outcome for a transaction if more than 25% of its simulations reached completion.

5. Lastly, we tested for predictive power by calibrating the model to the stock price histories at the various points in time (e.g. target company stock prices from the first 40% of the transaction), simulating price paths, calculating test-statistics, and comparing to the cutoffs. We then compared the simulated outcomes to the actual outcomes to produce an estimate of predictive power.

3.2 Standard Brownian Motion

The standard Brownian motion process, also known as a Wiener process, is one of the simplest continuous-time stochastic processes. The process was first rigorously constructed by by Norbert Wiener in 1918, based upon observations of the botanist Robert Brown in

---

[3]Given the asymmetric return profile of merger arbitrage (i.e. an arbitrageur loses more from incorrectly assuming a deal will complete), we attempted to minimize false-negatives.

1827, to describe the random motion of particles suspended in a fluid. It is one of the best known Lévy processes (a Lévy process is essentially a random walk in continuous-time).

Let W_t represent the y-component of a particle in standard Brownian motion as a function of time, where x is the position of the particle at time $t = 0$. W_t is generally characterized by the following properties:

1. By convention, we generally set $W_0 = x = 0$.

2. W_t has stationary increments. This means that the distribution of displacement of a particle in a given time interval $[s, t]$ depends only on the length of the interval.

3. W_t has independent increments. Thus, for $u_1, u_2 \geq 0$, $W_{t+u_1} - W_t$ and $W_{t+u_2} - W_t$ are independent and normally distributed.

4. W_t is continuous on $[0, \infty)$.

We define $f_{W_t}(y; x, \sigma)$ to be the probability density function of the Wiener process, in y. f_{W_t} satisfies the partial differential equation

$$\frac{df}{dt} = \frac{1}{2}\sigma^2 \frac{d^2 f}{dx^2}. \tag{3.3}$$

This PDE is referred to as the diffusion equation and σ^2 is the diffusion coefficient. When $\sigma^2 = 1$, solving Eq. 3.3 yields the probability density function

$$f_{W_t}(y; x, \sigma) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2t}(y - x)^2\right). \tag{3.4}$$

Brownian motion was first considered for use in financial markets in 1900 by Louis Bache-

lier in "The Theory of Speculation," and has been widely used since. Despite a number of

recent studies that have shown that asset returns do not fit a Gaussian distribution, Brown-

ian motion has remained relevant because of its mathematical convenience and the 50-year

body of academic financial literature that hinges on it. Returning to merger arbitrage, we

treated S_t, the price of the target stock at time t, as a Wiener process. We calibrated the

process to a particular deal by adjusting the two parameters, x and σ. We set x to be

the initial value of the target stock ($S_{0.1}$, $S_{0.2}$, $S_{0.3}$, etc.), and we set σ to be the standard

deviation of the target stock's price returns. Figure 3.1 shows the path of an actual target
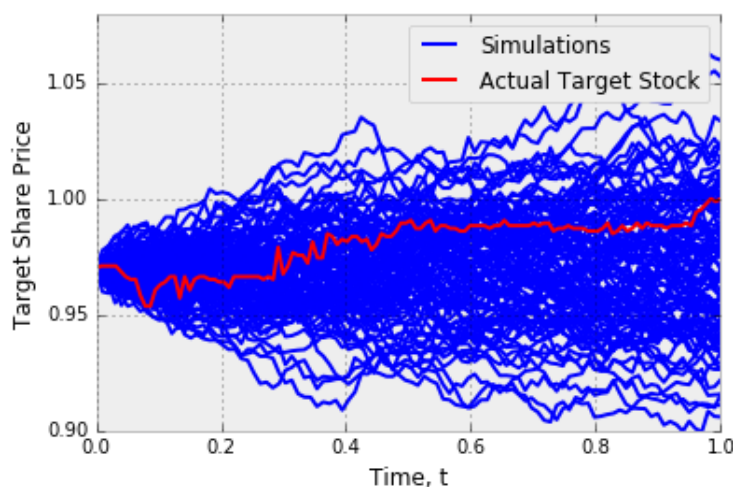

Figure 3.1: Brownian motion simulations of a target stock in transaction.

stock, along with 100 Brownian motion simulations run from time $t = 0$. We then computed

the simulated probability of completion for every transaction, at various points in the time-

series. In general, for a completed transaction, 38% of the simulated Brownian motion paths

successfully completed. For a given terminated transaction, however, only 6% the simulated paths would complete. In order to minimize the number false-negatives (predicting a deal completes when it actually fails), we set a cutoff value of around 14%. We then ran simulations on our data set from the various points in time. Figure 3.2 shows the percent of completed and terminated deals that were correctly predicted. We found that the model, halfway through the deal, predicted 70% of completed transactions correctly and 71% of terminated transactions correctly.
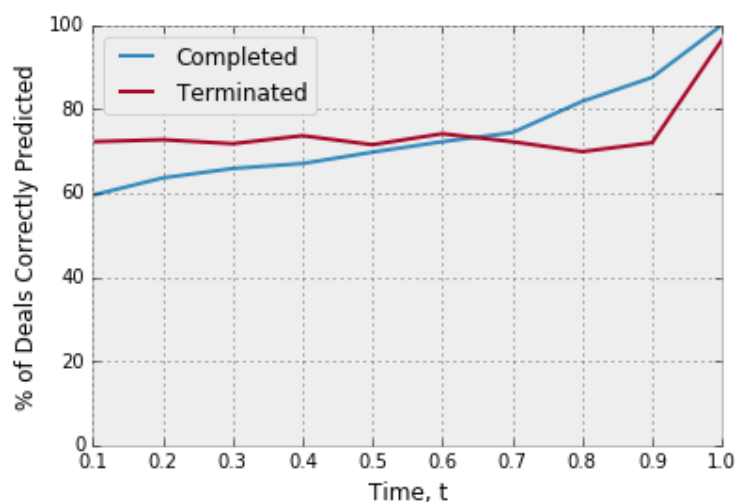
Figure 3.2: Standard Brownian Motion - Deal outcome prediction success rate.

We see a number of interesting observations here. Firstly, the prediction rates improve over time. This is expected: a set of simulations started from $t = 0$ will naturally have higher price variance at $t = 1$ than a set of simulations started from $t = 0.9$. Furthermore, we found that the model initially had a marginally higher prediction success rate for terminated deals than for completed deals. We believe that this is mostly due to the method we used:

because the model only considers a simulation to be a completed deal if $S_1 \in [0.99, 1.01]$, it is inherently more probable that a deal will be considered terminated. It is also possible that this is a reflection of initial investor sentiment: in many transactions that have failed, investors were skeptical from the beginning that the deal would complete, and the post-announcement target stock price reflected this. Thus, the transactions that terminated were much easier to identify early on than the transactions that completed. Furthermore, at $t = 0.5$, the model predicted that 1,769 deals would complete, out of which 1,649 actually completed, while 120 actually terminated. Thus, an arbitrageur can expect a 93% accuracy rate for predictions of deal completion.

The standard Brownian motion model has a number of shortcomings. Firstly, $\mathrm{E}[W_t] = x$, meaning the expected value of the target stock price at any point in time is equal to the initial value. Under this condition, a target stock would not trend towards the offer price and a deal would never complete or terminate. Furthermore, Brownian motion assumes constant volatility and linearly rising price variance, neither of which we found from the data to be true for target stock prices in completed deals. Lastly, we know from the data that the distribution of instantaneous returns is not Gaussian, but rather fat-tailed. Therefore, we proceed to more sophisticated models.

3.3   Geometric Brownian Motion

Geometric Brownian motion (GBM) is another continuous-time stochastic process, popular-
ized by Black and Scholes (1973) in the pricing of options and corporate liabilities. Geometric
Brownian motion is essentially standard Brownian motion with a modified drift and volatil-
ity component. In general, the stochastic differential equation for a GBM process is defined
as

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.5}$$

where, μ is the percentage drift expected per annum, dt is the time step, σ is the expected
daily volatility in stock price, and W_t is a standard Wiener process. The SDE can be solved
by dividing both sides by S_t, applying Itô's formula, and simplifying. The resulting solution
can be expressed as

$$S_t = S_0 e^{\left(\left(\mu - \sigma^2/2\right)t + \sigma W_t\right)}. \tag{3.6}$$

Thus, in GBM, the logarithm of the randomly varying component follows a Wiener process
with drift. The probability density function of a GBM price process is

$$f_{S_t}(s; \mu, t, \sigma) = \frac{1}{s\sigma\sqrt{2\pi t}} \exp\left(-\frac{\left(\ln s - \ln S_0 - \left(\mu - \frac{1}{2}\sigma^2\right)t\right)^2}{2\sigma^2 t}\right). \tag{3.7}$$

GBM is arguably the most widely used model of stock price movement due to a number
of its properties:

1. For any point in time $t_i < t_{i+n}$, $\frac{S_{t_i}}{S_{t_{i+n}}}$ is an independent random variable. If we interpret
   this ratio as the stock price return over a time horizon $[t_i, t_{i+n}]$, we find that returns

are independent of the stock prices and any previous returns, matching our real-world

expectations.

2. The expected value of the stock, given by $E[S_t|S_0 = x] = xe^{\mu t}$, allows stock prices

to drift deterministically. Furthermore, by definition, the stock price under GBM can

never be less than zero.

3. GBM is computationally convenient, as it relies on a standard Wiener process, and has

well-defined higher moments.

These properties can be advantageous when modeling target stocks in completed transac-

tions: the drift component, μ, can address the mean-reversion of the stock price to the offer

price, and the Wiener process, σW_t, reflects stochastic daily movements. We calibrated this

model to deals in our dataset by using the discretized version of the solution in Eq. 3.6,

$$S_{t_{i+1}} = S_{t_i} \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta}W_{i+1}\right), \tag{3.8}$$

and using a common calibration method for the discrete time GBM, which involves setting

the parameters μ and σ to equal

$$\mu = \frac{\mathrm{E}[R]}{t} + \frac{\sigma^2}{2}, \qquad \sigma = \frac{\mathrm{std}(R)}{\sqrt{\Delta t}}. \tag{3.9}$$

where R is a series of log returns such that $R_{i+1} = log(S_{t_{i+1}}/S_{t_i})$, and $\Delta = 0.01$ (per our

standardization process described in Section 2.1). Figure 3.3 shows the path of actual target

stocks along with 100 calibrated GBM simulations. The advantage over standard Brownian

motion is immediately obvious; the drift component μ allows the model to account for upward

or downward trends.

Again, we computed the simulated probability of completion for all deals, at the various

points in the deal's duration. For completed deals, we found that 46% of simulated GBM

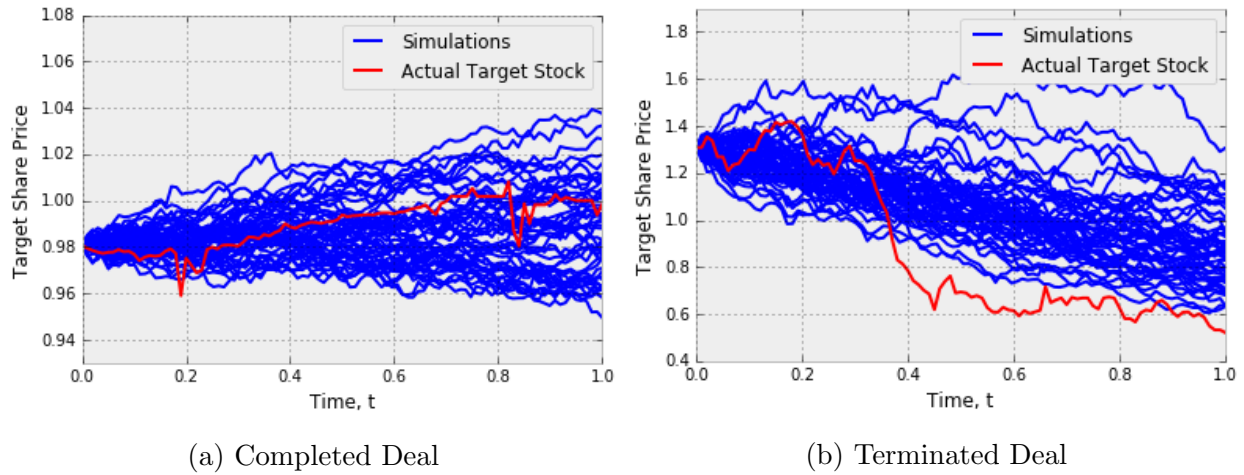(a) Completed Deal   (b) Terminated Deal

Figure 3.3: Geometric Brownian motion simulations of target stocks in transactions.

paths successfully completed. For terminated deals, we found that only 6% of simulated

paths completed. We calculated an optimal cutoff of around 14%. We then ran simulations

on our dataset and found that, on average, the model was able to successfully predict the

outcome for 72% of completed transactions and 76% of terminated transactions (halfway

through the deal, at $t = 0.5$). This represents a 2 percentage point improvement over

standard Brownian motion in success with completed transactions, and a 5 percentage point

improvement with respect to terminated transactions. Figure 3.4 shows the development

in prediction success rate over time. Generally, we find that the model's prediction success

rates are comparable to those of standard Brownian motion, with slightly better success

rates for terminated transactions. In the middle portion of the deal, the GBM model had slightly higher prediction rates for completed deals. Of the 1,814 deals that the GBM model predicted would complete, 1,711 completed, while 103 actually terminated, representing a 94% accuracy rate in prediction of completed outcomes. Given these results, the GBM model is not a significant improvement over the Brownian motion model, and thus we seek a better stochastic process. Figure 3.3(b) demonstrates the clear issue with using a Gaussian distribution for the stochastic component of the model: large, random movements are very difficult to replicate.
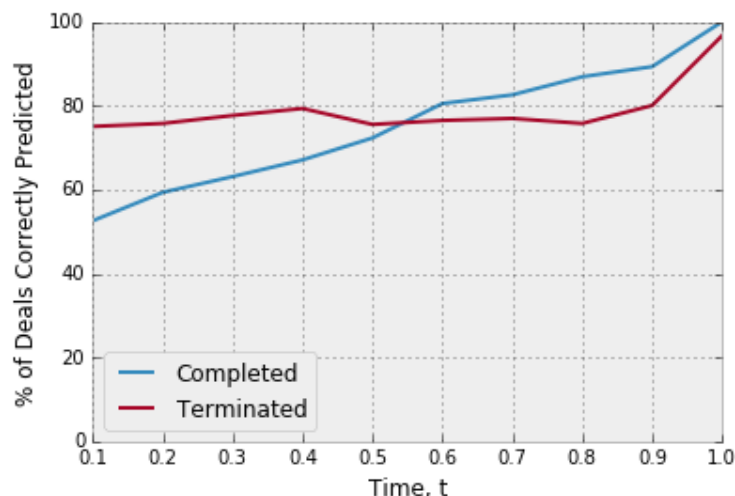

Figure 3.4: Geometric Brownian Motion - Deal outcome prediction success rate.

3.4   Merton Jump Diffusion Process

We next consider the Merton Jump Diffusion (MJD) process. Presented by Merton (1976), the MJD model was one of the first to address the limitations of the original GBM Black-

Scholes model. The model extended the original GBM process for stock price modeling by adding a Poisson Jump Diffusion process in order to capture the excess kurtosis of price returns. We begin by introducing the modified stochastic differential equation for an MJD process,

$$dS_t = \mu S_t dt + \sigma S_t dW_t + dJ_t, \tag{3.10}$$

where μ is the drift coefficient, σ is the diffusion coefficient, and W_t is of course our standard Wiener process. The key addition to GBM is the dJ_t term, which models jumps as

$$dJ_t = S_t d(\sum_{i=0}^{N_t}(V_i - 1)), \tag{3.11}$$

where $\log V_i$ is a normally distributed random variable and N_t is a Poisson process with rate λ. The difference $(V_j - 1)$ reflects the relative price change caused by a Poisson jump. The solution to Eq. 3.10, determined via Itô's Lemma, is

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right]\prod_{i=0}^{N_t}V_i. \tag{3.12}$$

Modifying the solution further, we consider $Y_i = log(V_i)$, yielding

$$X_t = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t + \sum_{i=0}^{N_t}(Y_i), \tag{3.13}$$

in which X_t now refers to the log returns of S_t (i.e. $\log\frac{S_t}{S_0}$) and the random variable $Y_i \sim N(\alpha, \delta^2)$. As we did for GBM in Eq. 3.8, we discretize the MJD solution yielding

$$\Delta X_t = \left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma \Delta W_t + \sum_{i=0}^{\Delta N_t}Y_i, \tag{3.14}$$

where $\Delta W_t \sim N(0, \Delta)$ and ΔN_t represents the number of jumps in $[t, t+\Delta]$. From Khaldi, Djeddour, and Meddahi (2014), the probability density function of ΔS_t is

$$f_{X_t}(x; \mu, \sigma, n, \delta) = \sum_{n=0}^{\infty} \frac{(\Delta\lambda)^n}{n!} e^{-\Delta\lambda} \times \left[\frac{1}{\sqrt{2\pi(\Delta\sigma^2 + n\delta^2)}} \times \exp\left(-\frac{(x - \Delta(\mu - \sigma^2/2))^2}{2(\Delta\sigma^2 + n\delta^2)} \right) \right],$$
(3.15)

where n represents the possible number of jumps and $\Delta = 0.01$ (per our time standardization).

The advantage of the MJD model over GBM is fairly straightforward: by allowing for discontinuous jumps in the target share price, we are effectively fattening the tails of the returns distribution (in order to better approximate what we saw in Figure 2.7). In Figure 3.5, we demonstrate a side-by-side comparison of MJD model simulations and GBM model simulations for the same target stock in a terminated transaction. While the GBM simulated paths do drift downwards, they still tend to overestimate the final target stock price. Meanwhile, the MJD simulations accurately account for larger price movements, and through the jump diffusion process, do a better job of simulating end results. The only downside of this model is the difficulty in calibrating model parameters. A number of studies, including Press (1967), Beckers (1981), Tankov and Coltchkova (1992), and Askari and Krichene (2008) present methods for parameter estimation with varying degrees of success. We ultimately use a modified version of the moments method used in Khadi, Djeddour, and Meddahi (2014), which builds on Askari and Krichene (2008). The calibration technique works by equalizing four theoretical cumulants (κ_i), which can be analytically determined,

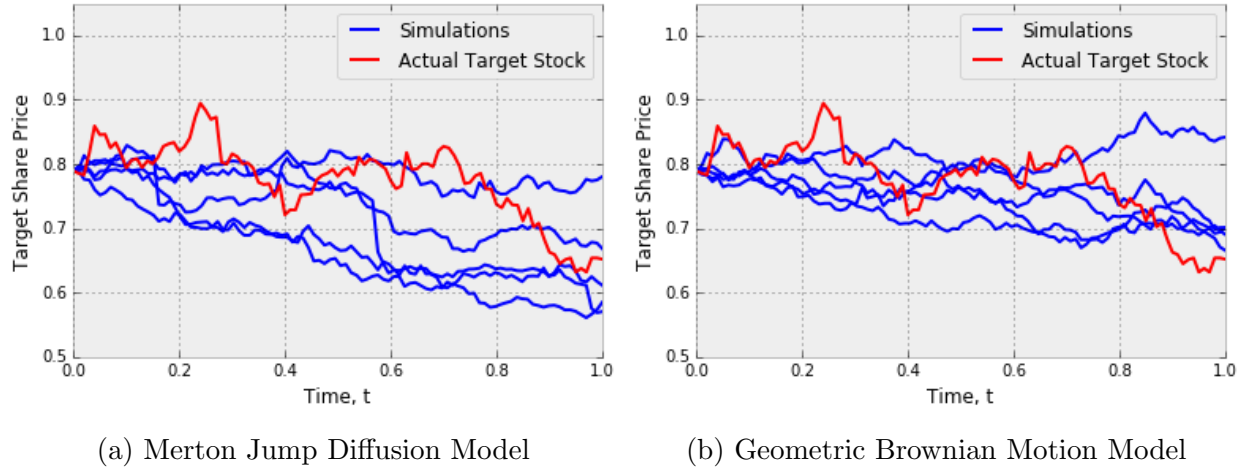(a) Merton Jump Diffusion Model   (b) Geometric Brownian Motion Model

Figure 3.5: Comparison of MJD and GBM Model Simulations for a Target Stock in a Terminated Deal.

with the corresponding empirical moments, which can be computationally found from the data. The theoretical cumulants are determined via the relation

$$E\left[(X - E(X))^2\right] = \frac{2\kappa}{2^\kappa \kappa!} \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} (\sigma^2 + n\delta^2)^\kappa. \tag{3.16}$$

From this relation, the cumulants $\kappa_1, \kappa_2, \kappa_4,$ and κ_6 can be determined:

$$\kappa_1 = E(X) = \mu - \frac{\sigma^2}{2},$$

$$\kappa_2 = E\left[(X - E(X))^2\right] = \sigma^2 + \lambda\delta^2,$$

$$\kappa_4 = E\left[(X - E(X))^4\right] = 3\left((\sigma^2 + \lambda\delta^2)^2 + \lambda\delta^4\right),$$

$$\kappa_6 = E\left[(X - E(X))^6\right] = 15\left((\sigma^2 + \lambda\delta^2)^3 + 3\lambda\delta^4(\sigma^2 + \lambda\delta^2) + \lambda\sigma^6\right). \tag{3.17}$$

We compute the corresponding empirical moments from the data, set them equal to the theoretical cumulants, and then solve the system of four equations to yield estimated parameters for μ, σ, λ, and δ. While this method is computationally expensive, it is feasible

in our discretized time field. We find that the estimated parameters are fairly accurate: the

MJD estimated parameters μ and σ, for most deals, very closely match the GBM estimated

parameters (both models share common drift and diffusion coefficients).

Using this calibration technique, we computed the simulated probabilities of completion

for all deals, at the various points in the deals' durations. For completed deals, we found that

roughly 45% of simulated MJB paths successfully completed. For terminated deals, roughly

6% of simulated price paths completed. We computed an optimal cutoff point of around 13%

of paths. Running simulations with this cutoff, our model was able to successfully predict

the outcome of 67% of completed transactions and 82% of terminated transactions, halfway

through the deal at $t = 0.5$. The development of deal outcome prediction success rate is

presented in Figure 3.6. We found that the MJD model was noticeably worse at predicting
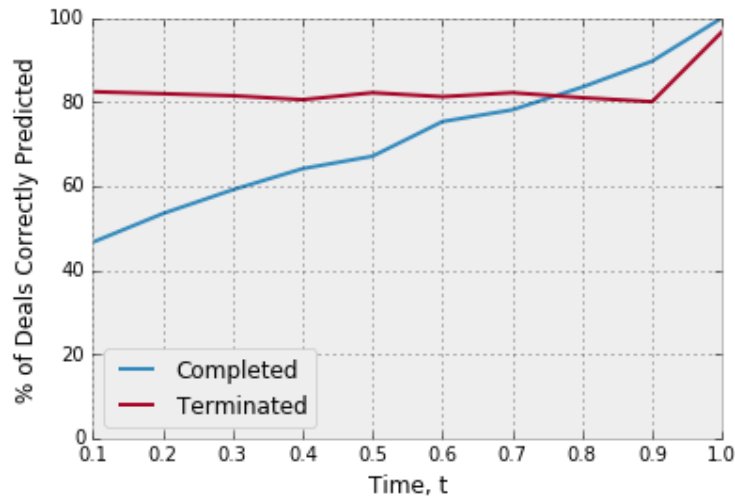

Figure 3.6: Merton Jump Diffusion Model - Deal outcome prediction success rate.

outcomes for completed deals in the first quarter of the transaction, and thereafter performed

comparable to previous models. This is likely due to difficulty in correctly calibrating the model parameters without a large amount of price data. On the other hand, we found that the MJD model was substantially better than previous models at predicting outcomes for terminated deals, at all points in the transaction. We presume that this is the jump process at work; because the standard Brownian model and the GBM model did not allow for discontinuous price jumps, they were incapable of capturing the large price movements of target stocks in terminated deals. Of the 1,662 deals that the MJD model predicted would complete, 1,587 actually completed, while only 75 terminated. This suggests a 95% accuracy rate in prediction of completion.

3.5   Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck (OU) is a stochastic process that is frequently used in the modeling of interest rates and commodity prices. Developed by Leonard Ornstein and George Uhlenbeck, the process is a modification of the standard Wiener process and is both Gaussian and Markovian. The process involves two components: a deterministic mean-reverting process and a stochastic diffusion process. The stochastic differential equation for the OU process is thus given by

$$dS_t = \lambda(\mu - S_t)dt + \sigma dW_t, \tag{3.18}$$

where μ is the steady state price level, λ is the mean reversion rate, and σ is the volatility of the process. The SDE can be solved via variation of parameters and basic integration, to

yield the continuous-time solution

$$S_t = S_0 e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \sigma \int_0^t e^{-\lambda(t-u)} W_u. \tag{3.19}$$

As with previous models, we employ the modified discrete solution to the OU process for our data

$$S_{t+i} = S_t e^{-\lambda\Delta} + \mu(1 - e^{-\lambda\Delta}) + \mu\sqrt{\frac{1 - e^{-2\lambda\Delta}}{2\lambda}} N_{0,1}, \tag{3.20}$$

where $\Delta = 0.01$ for a single time-step.

The OU process is especially useful to us for its property of long-term mean-reversion. Though the previous models we worked with did allow for deterministic trends, the drift coefficients were simply constants; thus, models with positive drift coefficients suggested unconstrained price growth in the long run. The SDE of the OU process, however, allows for a long-term steady state price. In the context of completed transactions, this steady state is, of course, the offer price. Furthermore, the drift coefficient, $\lambda(\mu - S_t)$, is a function of the current price of the target stock. Therefore, the further away a target stock is from the offer price, the more quickly it will mean revert. This is roughly consistent with our understanding of price dynamics in merger arbitrage.

There are multiple calibration methods for the OU process. We test two of these methods, following Van der Berg (2011). In the first approach, we take a least squares regression approach to estimating the parameters μ, λ, and σ. Given the relationship between consecutive price levels is linear with a random normal term ϵ, we can rewrite Eq. 3.20 as

$$S_{t+i} = aS_t + b + \epsilon, \tag{3.21}$$

in which the coefficients of the linear fit are

$$a = e^{-0.01\lambda},$$

$$b = \mu(1 - e^{-0.01\lambda}),$$

$$\epsilon = \mu\sqrt{\frac{1 - e^{-2(0.01)\lambda}}{2\lambda}} N_{0,1}. \tag{3.22}$$

To solve for the parameters of interest, we rewrite the equations as

$$\lambda = \frac{-\ln a}{0.01},$$

$$\mu = \frac{b}{1 - a},$$

$$\sigma = \text{std}(\epsilon)\sqrt{\frac{-2\ln a}{\delta(1 - a^2)}}. \tag{3.23}$$

The second approach uses the maximum likelihood estimation (MLE) method. By combining Eq. 3.20 and the probability density function of the normal distribution, we get the conditional probability density function,

$$f_{S_{i+1}}(S_i; \mu, \lambda, \hat{\sigma}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left[-\frac{(S_i - S_{i-1}e^{-\lambda\delta} - \mu(1 - e^{-\lambda\delta}))^2}{2\hat{\sigma}^2}\right], \tag{3.24}$$

where

$$\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2\lambda\delta}}{2\lambda}. \tag{3.25}$$

The log-likelihood function for the MLE process is therefore given by

$$\mathcal{L}(\mu, \lambda, \hat{\sigma}^2) = \sum_{i=1}^{n} \ln f_{S_{i+1}}(S_i; \mu, \lambda, \hat{\sigma})$$

$$= \frac{n}{2}\ln(2\pi) - n\ln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left[S_i - S_{i-1}e^{-\lambda\delta} - \mu(1 - e^{-\lambda\delta})\right]^2 \tag{3.26}$$

for a set of n price observations. Solving for the roots of the partial derivatives then gives the optimal parameters:

$$\frac{\partial \mathcal{L}(\mu, \lambda, \hat{\sigma})}{\partial \mu} = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^{n} \left[S_i - S_{i-1}e^{-\lambda\delta} - n(1 - e^{-\lambda\delta}) \right] = 0$$

$$\therefore \mu = \frac{\sum_{i=1}^{n} \left[S_i - S_{i-1}e^{-\lambda\delta} \right]}{n \left(1 - e^{-\lambda\delta} \right)}, \tag{3.27}$$

$$\frac{\partial \mathcal{L}(\mu, \lambda, \hat{\sigma})}{\partial \lambda} = -\frac{\delta e^{-\lambda\delta}}{\hat{\sigma}^2} \sum_{i=1}^{n} \left[(S_i - \mu)(S_{i-1} - \mu) - e^{-\lambda\delta} (S_{i-1} - \mu)^2 \right] = 0$$

$$\therefore \lambda = -\frac{1}{\delta} \ln \frac{\sum_{i=1}^{n} (S_i - \mu)(S_{i-1} - \mu)}{\sum_{i=1}^{n} (S_{i-1} - \mu)^2}, \tag{3.28}$$

$$\frac{\partial \mathcal{L}(\mu, \lambda, \hat{\sigma})}{\partial \hat{\sigma}^2} = \frac{n}{\hat{\sigma}} - \frac{1}{\hat{\sigma}^3} \sum_{i=1}^{n} \left[(S_i - \mu - e^{-\lambda\delta} (S_{i-1} - \mu) \right]^2 = 0$$

$$\therefore \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[(S_i - \mu - e^{-\lambda\delta} (S_{i-1} - \mu) \right]^2. \tag{3.29}$$

The issue that arises is that either of λ or μ is required to solve the equations. Van der Berg (2011) provides the following solution. We first change notation, so that

$$S_x = \sum_{i=1}^{n} S_{i-1},$$

$$S_y = \sum_{i=1}^{n} S_i,$$

$$S_{xx} = \sum_{i=1}^{n} S_{i-1}^2,$$

$$S_{xy} = \sum_{i=1}^{n} S_{i-1}S_i,$$

$$S_{yy} = \sum_{i=1}^{n} S_i^2, \tag{3.30}$$

which gives us equations for our parameters μ and λ:

$$\mu = \frac{S_y - e^{-\lambda\delta}S_x}{n\left(1 - e^{-\lambda\delta}\right)}, \tag{3.31}$$

$$\lambda = -\frac{1}{\delta}\ln\frac{S_{xy} - \mu S_x - \mu S_y + n\mu^2}{S_{xx} - 2\mu S_x + n\mu^2}. \tag{3.32}$$

By substituting Eq. 3.32 into Eq. 3.31 and simplifying, we get the solution for μ,

$$\mu = \frac{S_y S_{xx} - S_x S_{xy}}{n(S_{xx} - S_{xy}) - (S_x^2 - S_x S_y)}, \tag{3.33}$$

which can then be used to solve for σ and λ from Eq. 3.29 and Eq. 3.32, respectively.

We compared the effectiveness of these two calibration methods and found that, while both produced positive results, the least squares regression approach generally overestimated σ. As an example, Figure 3.7 shows simulations of a given target stock in a completed transaction using both approaches (Note: the y-axis range is narrower for the MLE plot). The simulated price paths for the LSR-calibrated model stray quite far from the actual price; the range of closing prices is $S_1 \in [0.85, 1.05]$. The simulated price paths for the MLE-calibrated model, however, remain true to the target stock, with closing prices in the range $S_1 \in [0.98, 1.01]$. Thus, we perform our Monte Carlo method using the MLE method for deal calibration.

We proceeded with our standard method. We computed the simulated probability of completion for all deals, at the various points in the deals' durations. For completed deals, roughly 39% of simulated OU paths successfully completed. For terminated deals, roughly 6% of paths reached completion. Our optimal cutoff varied with time, but was around 14%.

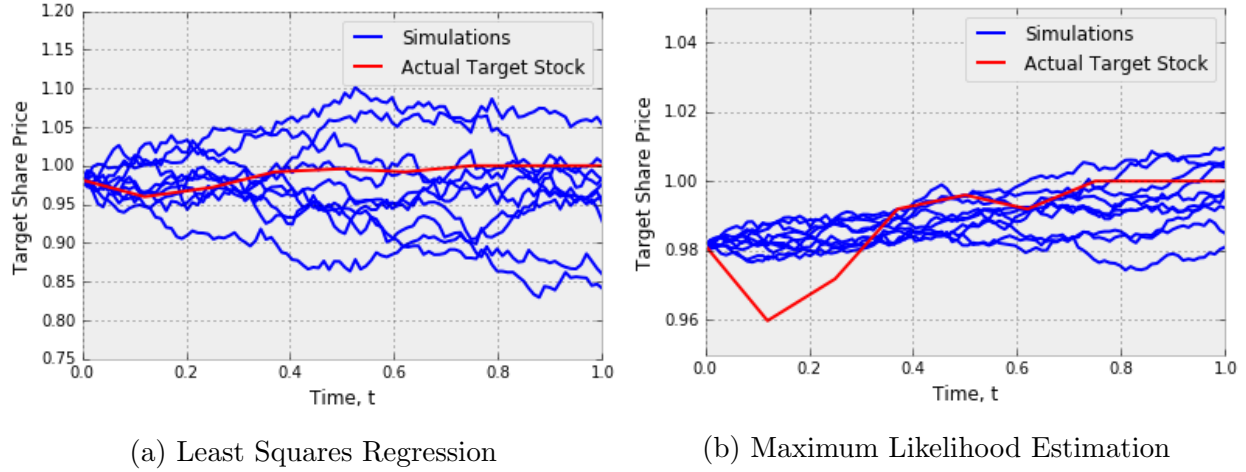(a) Least Squares Regression   (b) Maximum Likelihood Estimation

Figure 3.7: Comparison of Ornstein-Uhlenbeck process calibration techniques, on a target stock in a completed transaction.

Running simulations with this cutoff, our model was able to correctly predict the outcome of 71% of completed transactions and 81% of terminated transactions (halfway through the deal, at time $t = 0.5$). The development of the prediction success rate over time is shown in Figure 3.8. Of the 1,760 deals that the OU model predicted would complete, 1,679 actually completed, while 81 actually terminated (95% accuracy rate). There are several key observations here. Firstly, the OU prediction rate for completed deals represented a significant improvement over the MJD model, and a modest improvement over the standard Brownian model. The OU model's prediction rate for terminated deals was marginally lower than the MJD model's prediction rate, but was still substantially better than the standard Brownian and GBM models. We attribute superior prediction rates for completed deals to the model's mean reversion property; as the deal progresses, the model improves its estimation of μ (the steady state price) and simulations of target stocks in completed deals

are more likely to fall in the threshold range. Thus, the MLE-calibrated OU process is a fairly strong model for predicting deal outcomes. The principal shortcoming that we address in the next two sections, however, is the model's reliance on a Gaussian distribution to represent instantaneous returns.
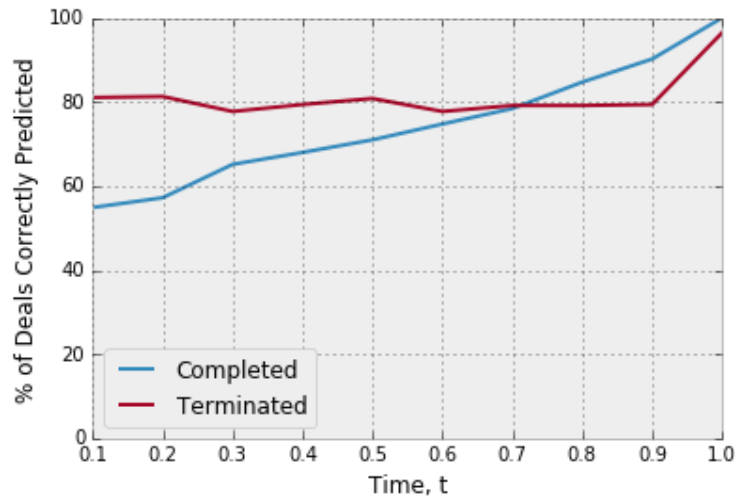

Figure 3.8: Ornstein-Uhlenbeck Model - Deal outcome prediction success rate.

3.6 Cauchy Process

Recalling the results of our data analysis in Section 2.3, we found that target stock price movements, in the short-term, were not normally distributed. Of the models we have considered thus far, however, only the Merton Jump Diffusion attempts to correct for the leptokurtic returns distribution. While the MJD model does allow for discontinuous jumps, it is an artificial solution; jumps are essentially randomly inserted into a standard GBM process. We, however, wish to construct a model that truly represents target stocks, and thus, we

require a fat-tailed distribution. In this section, we introduce the Cauchy distribution and the Cauchy process. While we don't use the Cauchy alone to predict deal outcomes, we do use it to replace the assumption of Gaussian stochastic movements in further models.

Named after Augustin Cauchy, the Cauchy is a stable, continuous, fat-tailed probability distribution. It has a number of important applications across physics and mathematics. Due to the nature of its fat tails, the distribution has no defined expected value or variance, or any moment generating function. In general, the distribution is defined by its location parameter (the statistical median), x_0, and its scale parameter (half the interquartile range), γ. The Cauchy is one of few stable distributions that has a probability density function which can be expressed analytically. It is given by

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]}. \tag{3.34}$$

The standard Cauchy distribution, similar to the standard Normal distribution, is the special case in which $x_0 = 0$ and $\gamma = 1$, with the simple probability density function

$$f(x; 0, 1) = \frac{1}{\pi(1 + x^2)}. \tag{3.35}$$

If we let X denote a Cauchy random variable, the characteristic function is given by

$$\phi_X(t; x_0, \gamma) = \int_{-\infty}^{\infty} f(x; x_0, \gamma)e^{ixt}dx = e^{ix_0 t - \gamma|t|}, \tag{3.36}$$

which is the Fourier transform of the PDF in Eq. 3.34. The characteristic function can be used to generate the PDF via the inverse Fourier transform. Lastly, for later use, we note

that the log-likelihood function of the standard Cauchy distribution is

$$\mathcal{L}(x; 0, 1) = -n \ln \pi - \sum_{i=1}^{n} \ln(1 + x_i^2). \tag{3.37}$$

We next introduce the Cauchy process – a stable, lévy stochastic process. The marginal probability distribution of the symmetric Cauchy process is, of course, the Cauchy distribution (similar to the relationship between the Wiener process and the standard Normal distribution). Due to the fat tails of the distribution, the Cauchy is a pure jump process. While we do not employ the Cauchy process alone in our analysis, we do use it to replace the Wiener process in the Ornstein-Uhlenbeck model.

3.7 Ornstein-Uhlenbeck-Cauchy Process

In this section, we present a modification of the standard Ornstein-Uhlenbeck process described in Section 3.5, in order to address the fat-tailed distribution of short-term target stock price returns. By substituting the Wiener process in the OU model with a Cauchy process, denoted by C_t, the modified stochastic differential equation is given by

$$dS_t = \lambda(\mu - S_t)dt + \sigma dC_t, \tag{3.38}$$

where μ is the steady state price level, λ is the mean reversion rate, and σ is the volatility of the process. Garbaczewski and Olkiewicz (2008) refer to this as the Ornstein-Uhlenbeck-Cauchy (OUC) process. The solution to the SDE, similar to the solution of the OU SDE

(Eq. 3.18), is given by

$$S_t = S_0 e^{-\lambda t} + \mu(1 - e^{-\lambda t}) + \sigma \int_0^t e^{-\lambda(t-u)} C_u, \tag{3.39}$$

and the discrete form of this solution is

$$S_{t+i} = S_t e^{-\lambda\Delta} + \mu(1 - e^{-\lambda\Delta}) + \mu\sqrt{\frac{1 - e^{-2\lambda\Delta}}{2\lambda}} C_{0,1}, \tag{3.40}$$

where $C_{0,1}$ is a random variable from a standard Cauchy distribution with $x_0 = 0$ and $\gamma = 1$.

To calibrate this model, we again use the MLE method. First, we derive the conditional

probability density function, which is a combination of the Cauchy PDF (Eq. 3.34) and the

OUC discrete solution (Eq. 3.40). We have

$$\begin{aligned}
f_{S_{i+1}}(S_i; \mu, \lambda, \hat{\sigma}) &= \frac{1}{\pi \left[1 + \frac{\left(S_i - S_{i-1}e^{-\lambda\delta} - \mu(1-e^{-\lambda\delta})\right)^2}{\hat{\sigma}^2} \right]} \\
&= \frac{\hat{\sigma}^2}{\pi \left[\hat{\sigma}^2 + (S_i - S_{i-1}e^{-\lambda\delta} - \mu(1 - e^{-\lambda\delta}))^2 \right]},
\end{aligned} \tag{3.41}$$

in which the parameter $\hat{\sigma}^2$ is again

$$\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2\lambda\delta}}{2\lambda}. \tag{3.42}$$

The log-likelihood function for a set of n price observations is then given by

$$\begin{aligned}
\mathcal{L}(\mu, \lambda, \hat{\sigma}^2) &= \sum_{i=1}^n \ln f_{S_{i+1}}(S_i; \mu, \lambda, \hat{\sigma}) \\
&= -n \ln \pi - \sum_{i=1}^n \ln \left[1 + \frac{\left(S_i - S_{i-1}e^{-\lambda\delta} - \mu(1 - e^{-\lambda\delta})\right)^2}{\hat{\sigma}^2} \right].
\end{aligned} \tag{3.43}$$

Determining the roots of the partial derivatives $\frac{\partial \mathcal{L}(\mu,\lambda,\hat{\sigma})}{\partial \mu}$, $\frac{\partial \mathcal{L}(\mu,\lambda,\hat{\sigma})}{\partial \lambda}$, and $\frac{\partial \mathcal{L}(\mu,\lambda,\hat{\sigma})}{\partial \hat{\sigma}^2}$ and solving yields the parameters of interest, μ, λ, and $\hat{\sigma}^2$. As the analytical solution is not straightforward, we computationally solve the system of equations when calibrating the model to each deal.

We demonstrate simulations of a calibrated OUC model for a target stock from a completed transaction in Figure 3.9. We find that, while the simulated outcomes are accurate, the price movement is not. The price jumps are realistic and well calibrated, but the price paths are overly linear. This is because the short-term stochastic price movements generated by the Cauchy are being masked by the mean reverting process, $\lambda(\mu - S_t)dt$. It is unrealistic to assume that the target stock's mean reversion towards the offer price will always be a strict function of the delta between the offer price and the current share price. To deal with this issue in a simple manner, we allow the mean reversion component of the SDE to vary
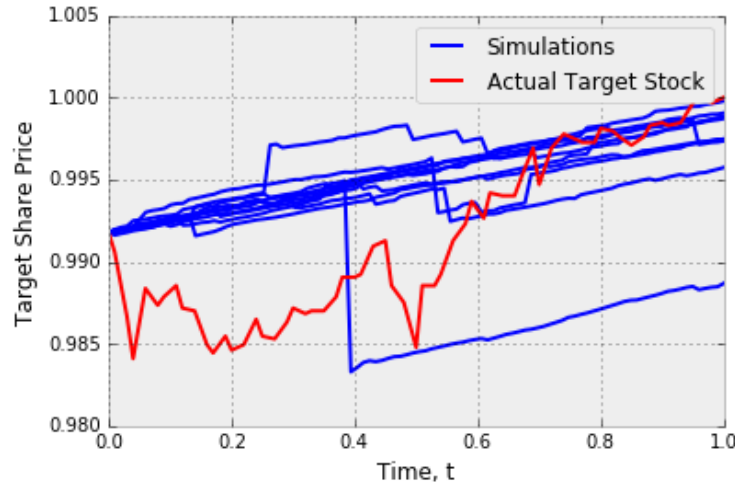

Figure 3.9: Ornstein-Uhlenbeck-Cauchy simulations of a target stock in a completed deal.

slightly at each time according to a modified Wiener process. This yields the modified SDE

$$dS_t = \lambda(\mu - S_t)dZ_t + \sigma dC_t, \tag{3.44}$$

in which Z_t represents the lognormal equivalent of the Wiener process. The PDF of this modified Wiener process is simply an extension of the lognormal distribution:

$$f_{Z_t}(y; x, \sigma) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2t}(\ln y - x)^2\right). \tag{3.45}$$

We use this modified process, rather than the ordinary Wiener process, because it allows us to set the expected value $\mathrm{E}(dZ_t) = 1$. Thus, we can have fluctuations in the mean-reversion component, while still maintaining the same long-run expected preferential drift towards μ. We refer to the model in Eq. 3.44 as the Ornstein-Uhlenbeck-Cauchy with lognormal Wiener mean reversion (OUCW). In Figure 3.10, we present simulations of the same target stock in Figure 3.9 using the OUCW model. The simulated price paths maintain the price jumps from the Cauchy process and now also include realistic short-term stochastic movements.

We then performed our Monte Carlo method on this model. For each deal, we computed the simulated probability at the various points in the deal's duration. For completed deals, roughly 38% of simulated OUCW paths successfully completed. For terminated deals, roughly 6% of paths reached completion. Our optimal cutoff varied around 13%. Running simulations with this cutoff, our model was able to correctly predict the outcome of roughly 64% of completed transactions and 86% of terminated transactions (halfway through the deal, at $t = 0.5$). Figure 3.11 shows the development of the prediction success rate over time. Of the 1,564 deals that the OUCW model predicted would complete, 1,505 actually
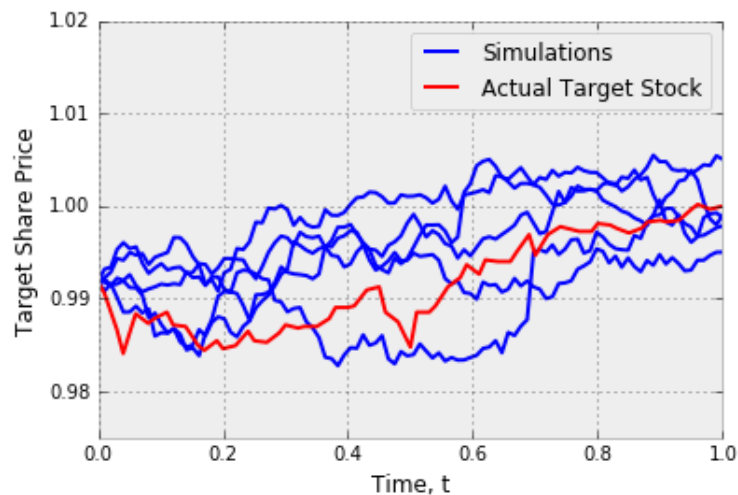
Figure 3.10: Ornstein-Uhlenbeck-Cauchy (with lognormal Wiener mean reversion) simulations of a target stock in a completed deal.

completed, while only 58 terminated (yielding an exceptional 96% accuracy rate for predictions of completion). There are a number of interesting observations here. Firstly, among all the models we have seen thus far, this model presents us with the worst average prediction
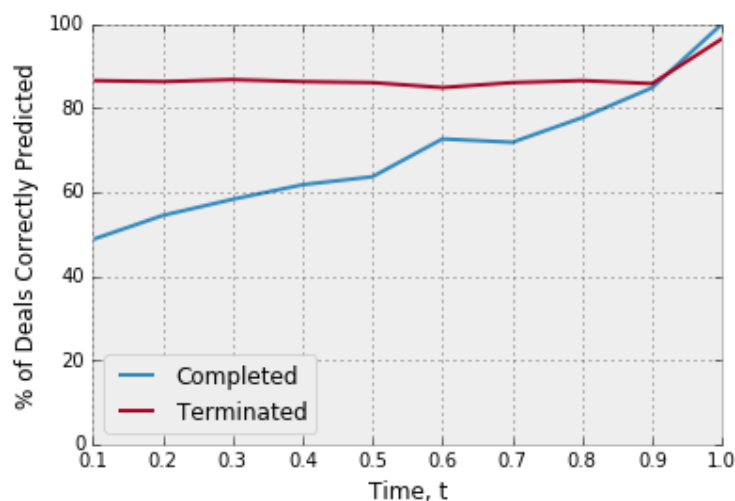

Figure 3.11: Ornstein-Uhlenbeck-Cauchy Model (with lognormal Wiener mean reversion) - Deal outcome prediction success rate.

success rate for completed deals. Very importantly, however, this model also significantly outperforms all previous models in predicting outcomes for terminated deals, presenting an average 5 percentage point improvement over the MJD and OU models, and a 10+ percentage point improvement over the standard Brownian and GBM models. Thus, we conclude that the Cauchy distribution does a significantly better job of modeling infrequent large price fluctuations than the Gaussian and Poisson processes we used in previous models.

3.8   Discussion

In this chapter, we demonstrated our first approach to estimating the probability that a transaction will successfully complete. We employed five stochastic processes of varying complexity to model the price of a target stock in a transaction: standard Brownian motion, geometric Brownian motion, Merton jump diffusion, Ornstein-Uhlenbeck, and Ornstein-Uhlenbeck-Cauchy[4]. For each model, we introduced calibration methods and performed Monte Carlo simulations to test for predictive power. In Figure 3.12, we present (a) the portion of all completed deals correctly predicted and (b) the portion of all terminated deals correctly predicted. Table 3.1 and Table 3.2 present these same results in color sensitized data tables.

With respect to prediction rates for completed deals, we found that all models had higher success rates over time. This means that as the deals approached completion, the models

---

[4]We are referring here to the Ornstein-Uhlenbeck-Cauchy model with lognormal Wiener mean reversion.

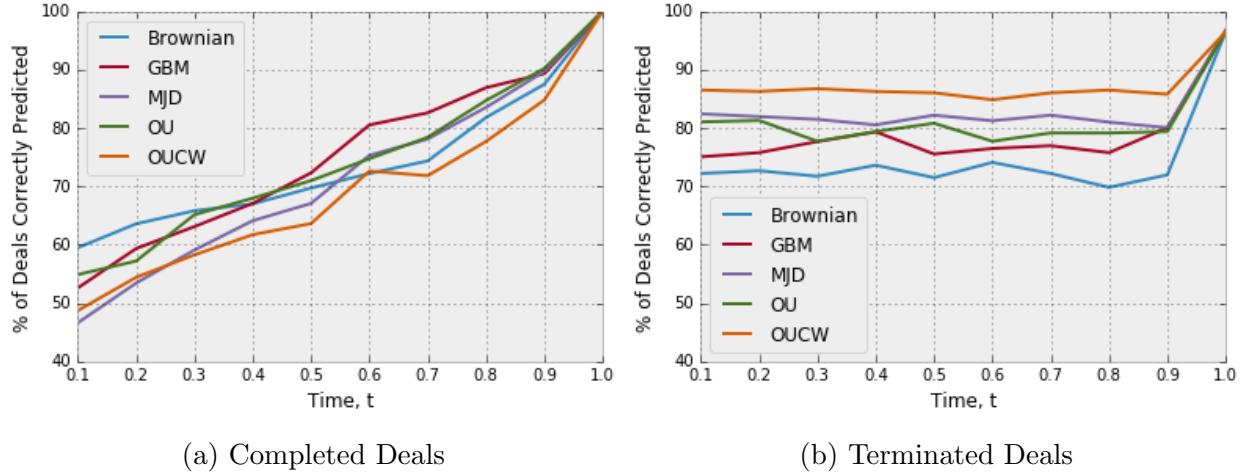(a) Completed Deals   (b) Terminated Deals

Figure 3.12: Deal outcome prediction success rate by model.

were able to make better predictions of the outcome. On average, halfway through the deal, the models correctly predicted outcomes for around 70% of completed deals. We found that the GBM, standard Brownian, and OU models had marginally higher success rates than did the MJD and OUCW models, especially in the early parts of the deal. Looking at prediction rates for terminated deals, however, we did not find that the models improved their prediction rates over time. Furthermore, it was abundantly clear from the results that the MJD and OUCW models provide much better predictions. In fact, the OUCW model predictions on average provided a 5 percentage point advantage over the MJD and OU models and an 8-14 percentage point advantage over other models.

In Table 3.3 and Figure 3.13, we present perhaps the most important results for an arbitrageur: the portion of deals predicted to complete that actually completed. As an example, at $t = 0.50$, the OUCW model had a 96% success rate when predicting successful outcomes, meaning that for every 100 deals the OUCW model predicted would complete, 96

Table 3.1: Portion of all completed transactions for which the outcome was correctly predicted.

Point in Time	Standard Brownian	Geometric Brownian	Merton Jump Diffusion	Ornstein- Uhlenbeck	Ornstein- Uhlenbeck- Cauchy
0.10	59%	53%	47%	55%	49%
0.20	64%	59%	53%	57%	54%
0.30	66%	63%	59%	65%	58%
0.40	67%	67%	64%	68%	62%
0.50	70%	72%	67%	71%	64%
0.60	72%	81%	75%	75%	73%
0.70	74%	83%	78%	78%	72%
0.80	82%	87%	84%	85%	78%
0.90	87%	89%	90%	90%	85%
1.00	100%	100%	100%	100%	100%

Table 3.2: Portion of all terminated transactions for which the outcome was correctly predicted.

Point in Time	Standard Brownian	Geometric Brownian	Merton Jump Diffusion	Ornstein- Uhlenbeck	Ornstein- Uhlenbeck- Cauchy
0.10	72%	75%	82%	81%	86%
0.20	73%	76%	82%	81%	86%
0.30	72%	78%	81%	78%	87%
0.40	74%	79%	81%	79%	86%
0.50	71%	76%	82%	81%	86%
0.60	74%	76%	81%	78%	85%
0.70	72%	77%	82%	79%	86%
0.80	70%	76%	81%	79%	86%
0.90	72%	80%	80%	79%	86%
1.00	96%	97%	97%	96%	96%

of them actually completed, while 4 of them terminated. This is, of course, important because an arbitrageur faces an asymmetric risk-reward profile and has a much larger downside than upside on any given merger arbitrage position. It is also important that we give more weight to these results than the basic results presented in Table 3.1 or Table 3.2, because the true test of a model is its ability to separate deals that will complete from deals that will terminate; a model that predicts 100% of completed deals is useless if it also claims that many terminated deals will complete. We find that the results here actually diverge quite a bit from what we saw in Table 3.1. While all models provide fairly high values (average of ∼94%), the OUCW model is the clear winner, with a consistent proportion of ∼96%. This is mostly due to the fact that the OUCW model presented significantly lower type II error than the other models; very few terminated transactions were incorrectly predicted to complete. The basic OU model and the MJD model also exhibited strong results.

Table 3.3: Portion of deals predicted to complete that actually completed.

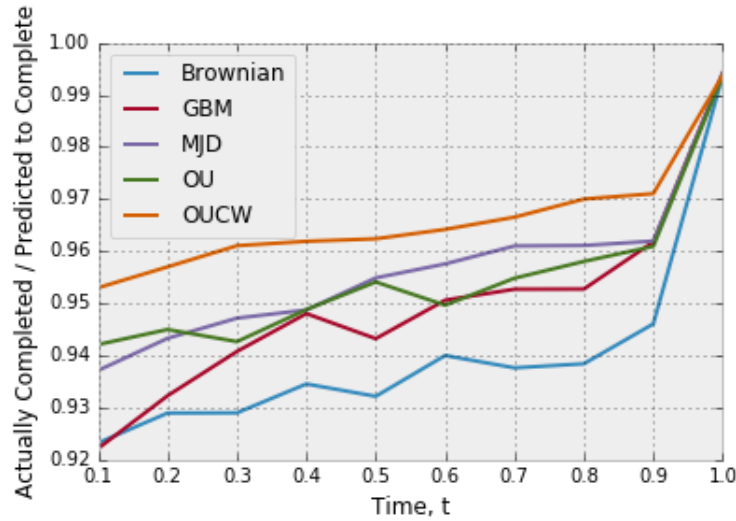| Point in Time	Standard Brownian	Geometric Brownian	Merton Jump Diffusion	Ornstein- Uhlenbeck	Ornstein- Uhlenbeck- Cauchy
0.10	92%	92%	94%	94%	95%
0.20	93%	93%	94%	94%	96%
0.30	93%	94%	95%	94%	96%
0.40	93%	95%	95%	95%	96%
0.50	93%	94%	95%	95%	96%
0.60	94%	95%	96%	95%	96%
0.70	94%	95%	96%	95%	97%
0.80	94%	95%	96%	96%	97%
0.90	95%	96%	96%	96%	97%
1.00	99%	99%	99%	99%	99%

Figure 3.13: Portion of deals predicted to complete that actually completed.

All of these findings are fairly consistent with our understanding of the data from Section 2.3. We postulated that target stock price movement in completed deals would be best described by a deterministic trend towards the offer price, combined with a short-term stochastic process. The decent predictive power of these models suggests that our breakdown of the forces on a target stock was valid. Regarding the deterministic trend, our results demonstrate that a mean-reverting process towards the offer price is more representative than a simple drift coefficient. We see from the results in Table 3.3 that the mean-reverting processes outperformed the simple drift models: the OU gave better results than the Brownian and GBM models, and the OUCW improved upon the MJD model. Regarding the stochastic component, we expected that short-term price returns could be better represented by a fat-tailed distribution than a Gaussian. Our results verify that the OUCW process is indeed better equipped to predict outcomes for terminated deals, in which unexpected, large price

fluctuations are common. We did, however, expect the OUCW and MJD models to have better success predicting outcomes for completed deals specifically. We postulate that the lower prediction success rates for the MJD and OUCW models in Table 3.1 from $t \in [0.1, 0.5]$ are a result of 1) poor calibration of parameters, and 2) the lack of a stochastic volatility component. Our parameter calibration methods, especially for the fat-tailed models, were rather complicated and may not have been as effective for shorter price histories. Additionally, we found in Section 2.3 that volatility actually decreases over the course of the transaction for target stocks in completed deals. Our models only allowed for constant volatility/scale, a definite limitation. Perhaps allowing the scale of the fat-tailed distributions to decrease over time would have reduced the number of completed deals that were assumed to have terminated.

In conclusion, our work in this chapter provides a framework through which arbitrageurs can approach the stochastic modeling of stocks in M&A deals. We demonstrated that reasonable predictive power can be attained by multiple models through the appropriate calibration of parameters using available target stock price data. Using an Ornstein-Uhlenbeck-Cauchy process with a lognormal Wiener mean reversion, we were able to successfully predict the outcome for roughly three out of every four completed deals in our dataset. We further showed that model predictions were robust; roughly 96% of deals the model predicted to complete actually ended up completing. Our empirical findings support the use of fat-tailed distributions, rather than the traditional Gaussian, to represent instantaneous returns.

Chapter 4

A General Model for Target Stocks in Transactions

4.1 Overview

In our first approach to predicting deal outcomes, we analyzed historical price data to determine the representative characteristics of target stock price movements in completed deals. We then presented various stochastic processes that met these characteristics and compared their effectiveness in predicting deal outcomes. In this chapter, we broaden our scope and construct general models to represent all target stocks in M&A transactions. The processes we present in this chapter differ from those in the previous chapter by now explicitly considering the expected probability that the deal will complete as a factor in the price of the stock: we assume that the price of the stock reflects the likelihood of deal completion, which itself

can be considered a function of stock price. We computationally test the predictive power of each model that we build, and again compare the effectiveness of a standard Gaussian distribution against a fat-tailed distribution in modeling stochastic returns.

4.2 Model Construction

To begin, we reintroduce notation from previous chapters. We assume that all deals occur on a normalized time scale $t \in [0, 1]$. We refer to the price of a target stock at time t as S_t. Let the probability of the deal successfully completing be $p(t, S_t)$, where K is the deal offer price. In constructing our models, we apply the following constraints:

1. There are only two outcomes in a given transaction: the deal is successful or it is unsuccessful.

2. We do not consider situations with competing offers, revised bids, or hostile/unsolicited offers.

3. In a successful transaction, the target stock price becomes K.

4. In an unsuccessful transaction, the target stock follows some unconstrained stochastic process y_t.

5. The target stock pays no dividends during the transaction.

6. There are no immediate arbitrage opportunities.

Assuming a Brownian Unconstrained Process

In this first model, we assume that the unconstrained process is simply Brownian motion, meaning

$$dy_t = \sigma dW_t, \tag{4.1}$$

where σ is the volatility of the process and W_t is a standard Wiener process. Thus, under our constraints, we assume that the value of the target stock at time t can be given by

$$S_t = (1 - p)y_t + pK. \tag{4.2}$$

It follows that the change in S_t is

$$dS_t = -p_t y_t \, dt + (1 - p)\sigma \, dW_t + p_t K \, dt, \tag{4.3}$$

where p_t is defined to be the partial $\partial p / \partial t$. Eq. 4.2 can be rewritten as

$$y_t = \frac{S_t - pK}{1 - p} \tag{4.4}$$

and substituted into Eq. 4.3 to yield

$$\begin{aligned} dS_t &= -\frac{p_t(S_t - pK)}{1 - p} dt + (1 - p)\sigma \, dW_t + p_t K \, dt \\ &= \frac{p_t(K - S_t)}{1 - p} dt + (1 - p)\sigma \, dW_t. \end{aligned} \tag{4.5}$$

This is the stochastic differential equation that defines this model. The SDE rationally suggests that a target stock price will increase if p_t is increasing; in other words, if the likelihood of deal completion increases, the target stock price will move closer to the offer

price. Additionally, as likelihood of completion increases, the stochastic component of the price movement becomes less significant; this matches our observation of decreasing price volatility over time in completed deals, from Section 2.3. Conversely, as likelihood of deal completion decreases, the stock price will approach the unconstrained process dy_t. We define

$$\mu = \frac{p_t(K - S_t)}{1 - p}, \qquad \Sigma = (1 - p)\sigma \tag{4.6}$$

in order to represent Eq. 4.5 in the common form

$$dS_t = \mu \, dt + \Sigma \, dW_t. \tag{4.7}$$

Our goal is to find an equation for p which reflects the probability of deal completion given all potential future price paths following this SDE. In order to obtain this solution, we first use Itô's lemma to define the change in probability p for a specific instance of S_t as

$$dp = \left(\frac{\partial p}{\partial t} + \mu \frac{\partial p}{\partial S} + \frac{\Sigma^2}{2} \frac{\partial^2 p}{\partial S^2}\right) dt + \Sigma \frac{\partial p}{\partial S} \, dW_t. \tag{4.8}$$

Given its dependence on only share price and time to completion, we claim that $p(t, S_t)$ can be treated as a contingent claim (similar to a call or put option). Thus, following similar logic to the Black-Scholes derivation, we construct an artificial portfolio, Π, given by

$$\Pi = -p + \Delta S_t, \tag{4.9}$$

whose variation over a small time period, dt, is wholly deterministic. The portfolio is short one derivative security (p) and long Δ shares of the underlying target stock. The change in the value of our portfolio given a small fluctuation is

$$d\Pi = -dp + \Delta dS_t. \tag{4.10}$$

We substitute Eq. 4.7 and Eq. 4.8 into Eq. 4.10 to yield

$$d\Pi = -\left[\left(\frac{\partial p}{\partial t} + \mu\frac{\partial p}{\partial S} + \frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}\right)dt + \Sigma\frac{\partial p}{\partial S}\,dW_t\right] + \Delta\left(\mu\,dt + \Sigma\,dW_t\right)$$

$$= -\frac{\partial p}{\partial t}dt - \mu\frac{\partial p}{\partial S}dt - \frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}dt - \Sigma\frac{\partial p}{\partial S}\,dW_t + \Delta\left(\mu\,dt + \Sigma\,dW_t\right). \tag{4.11}$$

To eliminate the stochastic component dW_t, we set $\Delta = \partial p/\partial S$, giving us

$$d\Pi = -\frac{\partial p}{\partial t}dt - \mu\frac{\partial p}{\partial S}dt - \frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}dt - \Sigma\frac{\partial p}{\partial S}\,dW_t + \frac{\partial p}{\partial S}\mu\,dt + \frac{\partial p}{\partial S}\Sigma\,dW_t$$

$$= -\frac{\partial p}{\partial t}dt - \frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}dt. \tag{4.12}$$

As this equation has no dependence on dW_t, it must be risk-less at dt. Further, under the assumption of no arbitrage, Π must have the same rate of return as any other risk-free short-term security component of the portfolio. Thus, we have

$$d\Pi = \mu\frac{\partial p}{\partial S}dt, \tag{4.13}$$

where the right-hand side of the equation is the deterministic return of Δ shares, from Eq. 4.7. Expanding $d\Pi$ gives us

$$-\frac{\partial p}{\partial t}dt - \frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}dt = \mu\frac{\partial p}{\partial S}dt$$

$$\therefore \frac{\partial p}{\partial t} + \mu\frac{\partial p}{\partial S} = -\frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}. \tag{4.14}$$

Substituting the definitions from Eq. 4.6 into this drift-diffusion equation yields

$$p_t + \frac{p_t(K - S_t)}{1 - p}\frac{\partial p}{\partial S} = -\frac{\sigma^2(1 - p)^2}{2}\frac{\partial^2 p}{\partial S^2}$$

$$\therefore \frac{\partial p}{\partial t} = -\frac{\sigma^2(1 - p)^3}{2\left[1 - p + (K - S_t)\frac{\partial p}{\partial S}\right]}\frac{\partial^2 p}{\partial S^2}. \tag{4.15}$$

This partial differential equation in Eq. 4.15 can now be solved to yield the probability of deal success given the parameters σ, K, S_t, and a differentiable function for $p(t, S_t)$. Though a solution could potentially be reached via transformation to a more basic diffusion equation using appropriate terminal and boundary conditions (in a similar procedure to solving the Black-Scholes equation), we opt for more straightforward computational methods. We represent p on a discretized grid and principally use two approaches: 1) solving p backwards in time using finite-difference approximation, and 2) integrating over all random walk paths to obtain a solution. Regarding the second method, to integrate over all paths, we use a backwards iterative procedure in which we solve for $p(t-\epsilon, S_{t-\epsilon})$ in terms of the probabilities at $p(t, S_t)$, beginning with the known $p(T, S_T)$, where $T = 1$. Thus, we have

$$p(t - \epsilon, S_{t-\epsilon}) = \int_{-\infty}^{\infty} p(t, S_{t-\epsilon} + y) f_\epsilon(y; S_{t-\epsilon}) dy, \qquad (4.16)$$

where $f_\epsilon(y; S_{t-\epsilon})$ refers to the probability density function of step size when starting at $S_{t-\epsilon}$ over a time interval ϵ. We assume the following differentiable function for $p(T, S_T)$, which will give us a value near 0.9 if the target stock price is at or near the offer price, and a value near 0 if the difference between the two is significant (only at $t = T$)[1]:

$$p(T, S_T) = 0.9 \exp\left(\frac{-(S_T - K)^2}{2\lambda^2}\right). \qquad (4.17)$$

We constrain the maximum value of Eq. 4.17 to be 0.9 in order to avoid division by zero errors in our solution approximation method. Thus, the maximum probability we allow the model to assign for deal completion is 90%. Regarding the probability density function, we

---

[1] λ is used to calibrate the initial probability function; we set $\lambda = 0.1$.

can calculate f_ϵ from the stochastic differential equation in Eq. 4.7 and the Brownian motion probability density function in Eq. 3.4. We find that

$$
\begin{aligned}
f_\epsilon(y; S_{t-\epsilon}) &= \frac{1}{\Sigma\sqrt{2\pi\epsilon}} \exp\left[\frac{-1}{2\Sigma^2\epsilon}(y - \mu\epsilon)^2\right] \\
&= \frac{1}{(1-p)\sigma\sqrt{2\pi\epsilon}} \exp\left[\frac{-1}{2(1-p)^2\sigma^2\epsilon}\left(y - \frac{p_t(K - S_t)}{1-p}\epsilon\right)^2\right].
\end{aligned}
\tag{4.18}
$$

Thus, the first iteration can be solved for by

$$
p(T - \epsilon, S_{T-\epsilon}) = \int_{-\infty}^{\infty} \exp\left(\frac{-((S_{T-\epsilon} + y) - K)^2}{2\lambda^2}\right) f_\epsilon(y; S_{T-\epsilon})dy,
\tag{4.19}
$$

and then the second iteration can be solved using the solutions from the first iteration:

$$
p(T - 2\epsilon, S_{T-2\epsilon}) = \int_{-\infty}^{\infty} p(T - \epsilon, S_{T-2\epsilon} + y)f_\epsilon(y; S_{T-2\epsilon})dy.
\tag{4.20}
$$

We approximate a solution for $p(t, S_t)$ by repeating this iterative integration in a discretized field, using small timesteps over the complete duration of a transaction, while constantly updating values of p based on the most recent timesteps. Given appropriate parameters σ, K, and S_t, we can estimate the probability the deal will complete at any point in time. Figure 4.1 shows estimated deal completion probability curves at different points in time for a sample completed transaction. The results match our expectations: the probability curves narrow and grow taller as the deal progresses.
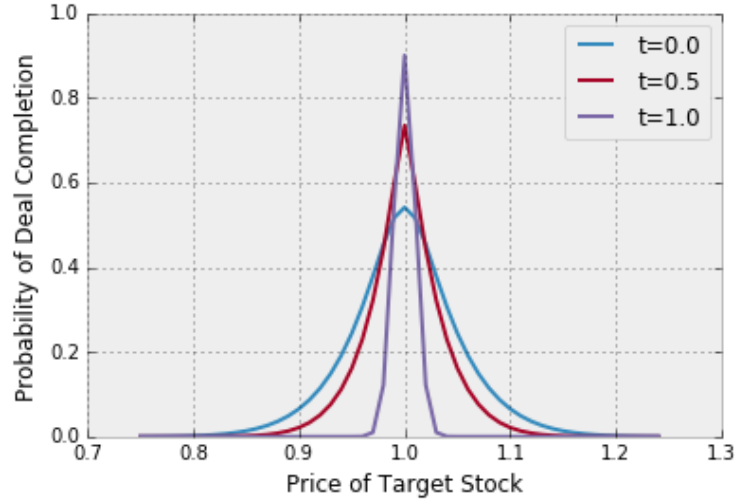
Figure 4.1: Probability of deal completion at various points in the transaction, based on target stock price, assuming a Brownian unconstrained process.

Assuming a Geometric-Brownian Unconstrained Process

We improve upon the model produced in the previous section by now using geometric Brownian motion to represent the unconstrained process, meaning

$$dy_t = \mu S_t + \sigma S_t dW_t, \tag{4.21}$$

where μ is the drift coefficient, σ is the volatility of the process, and W_t is a standard Wiener process. Thus, the target stock price at time t is given by

$$S_t = (1 - p)y_t + pK. \tag{4.22}$$

A change in S_t is thus given by

$$dS_t = -p_t y_t \, dt + (1 - p) \left(\mu S_t \, dt + \sigma S_t \, dW_t \right) + p_t K \, dt. \tag{4.23}$$

Eq. 4.22 can be rewritten as

$$y_t = \frac{S_t - pK}{1 - p},$$

(4.24)

and substituted into Eq. 4.23 to yield

$$
\begin{aligned}
dS_t &= -\frac{p_t(S_t - pK)}{1 - p}dt + (1 - p)\left(\mu S_t\, dt + \sigma S_t\, dW_t\right) + p_t K\, dt \\
&= \left[\frac{p_t(K - S_t) + (1 - p)^2 \mu S_t}{1 - p}\right] dt + (1 - p)\sigma S_t\, dW_t.
\end{aligned}
$$

(4.25)

This modified stochastic differential equation suggests that

1. When the expected probability of deal completion is high, the stochastic process is mostly a deterministic drift towards the offer price.

2. When the expected probability of deal completion is low, the process has a deterministic GBM drift of $\mu S_t dt$ and a substantial stochastic component $\sigma S_t dW_t$.

We define

$$M = \frac{p_t(K - S_t) + (1 - p)^2 \mu S_t}{1 - p}, \qquad \Sigma = (1 - p)\sigma$$

(4.26)

in order to represent Eq. 4.25 in the common form

$$dS_t = M\, dt + \Sigma\, dW_t.$$

(4.27)

Repeating the same derivation as in the previous section, we find the drift-diffusion equation is given by

$$\frac{\partial p}{\partial t} + M\frac{\partial p}{\partial S} = -\frac{\Sigma^2}{2}\frac{\partial^2 p}{\partial S^2}.$$

(4.28)

Substituting Eq. 4.26 into this equation yields the partial differential equation of interest,

$$p_t + \frac{p_t(K - S_t) + (1 - p)^2 \mu S_t}{1 - p} \frac{\partial p}{\partial S} = -\frac{\sigma^2 (1 - p)^2}{2} \frac{\partial^2 p}{\partial S^2}$$

$$\therefore \frac{\partial p}{\partial t} = -\frac{\sigma^2 (1 - p)^3}{2 \left[1 - p + (K - S_t) + (1 - p)^2 \mu S_t \right] \frac{\partial p}{\partial S}} \frac{\partial^2 p}{\partial S^2}. \tag{4.29}$$

We apply the same iterative computational method to solve this equation as was described

in Eq. 4.16, only with a modified conditional probability density function, given by

$$f_\epsilon(y; S_{t-\epsilon}) = \frac{1}{\Sigma \sqrt{2\pi\epsilon}} exp\left[\frac{-1}{2\Sigma^2 \epsilon} (y - M\epsilon)^2 \right]$$

$$= \frac{1}{(1 - p)\sigma \sqrt{2\pi\epsilon}} exp\left[\frac{-1}{2(1 - p)^2 \sigma^2 \epsilon} \left(y - \frac{p_t(K - S_{t-\epsilon}) + (1 - p)^2 \mu S_{t-\epsilon}}{1 - p} \epsilon \right)^2 \right].$$

$$\tag{4.30}$$

Therefore, given appropriate parameters μ, σ, K, and S_t, we can estimate the probability

that a given deal will complete at any point in its duration. For any given target stock, we

assume that the drift coefficient in the unconstrained process, μ, is simply the prevailing risk-

free rate. This assumption is rational because, as we learned in Sections 1.2 and 2.3, target

stocks in successful transactions exhibit less exposure to broader market conditions, while

stocks outside of transactions tend to move with the market. Figure 4.2 shows estimated deal

completion probability curves at different points in time for a sample completed transaction.

There are a few interesting observations here.

First, we note that the center of the probability curve actually shifts from $S_0 \sim 0.96$ to

$S_1 \sim 1.00$ over the course of the transaction. This is the drift component of the GBM process

at work and it provides our model with a very important feature. Recall from Section 2.3

that the average target share price on the day after deal announcement is $S_0 \sim 0.97$; part of the reason that we typically see this discount to the offer price is time discounting. Under no-arbitrage, even if there is a guaranteed 100% chance that the deal will complete, the initial share price of the target stock must be K/e^{rt}, where K is the offer price, r is the risk-free rate of return, and t is the anticipated length of the transaction. Thus, very rationally, our model attributes a lower expected probability of success to deals in which the target stock begins trading above K/e^{rt}. We emphasize that this feature is distinct to this model; the model that assumed a Brownian unconstrained process did not account for this.

Second, as with the previous model, we note that the probability curves become narrower and taller over the course of the deal. This is expected: assuming constant volatility, if there is little time left in the transaction there is a low probability that the target stock will be able to move significantly.
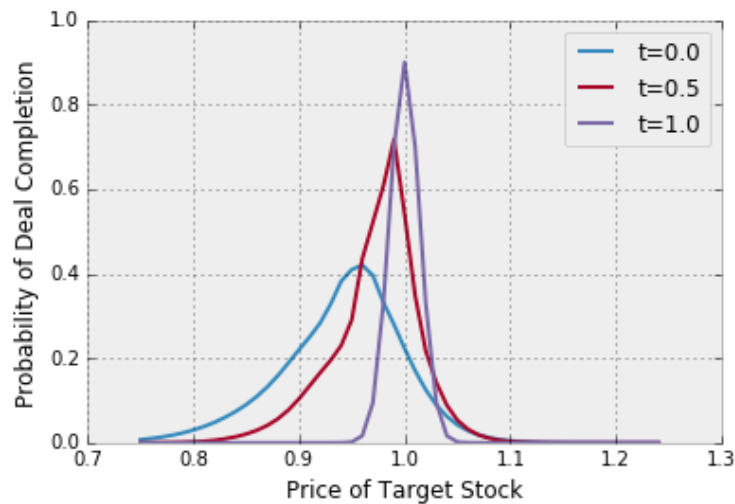

Figure 4.2: Probability of deal completion at various points in the transaction, based on target stock price, assuming a geometric Brownian unconstrained process.

Assuming a Cauchy Unconstrained Process with Drift

In this last iteration, we improve upon previous models by now allowing for instantaneous returns in the unconstrained process to fit a fat-tailed Cauchy distribution. Thus, the unconstrained process is given by

$$dy_t = \mu S_t + \sigma S_t dC_t, \tag{4.31}$$

where μ is the drift coefficient, σ is the volatility of the process, and dC_t is a standard Cauchy process. The target stock price at time t is still given by

$$S_t = (1 - p)y_t + pK, \tag{4.32}$$

and a change in S_t is thus given by

$$dS_t = -p_t y_t \, dt + (1 - p) \left(\mu S_t \, dt + \sigma S_t \, dC_t \right) + p_t K \, dt. \tag{4.33}$$

As before, rewriting Eq. 4.32 yields

$$y_t = \frac{S_t - pK}{1 - p}, \tag{4.34}$$

which can be substituted into Eq. 4.33 to yield

$$
\begin{aligned}
dS_t &= -\frac{p_t(S_t - pK)}{1 - p} dt + (1 - p)\left(\mu S_t \, dt + \sigma S_t \, dC_t\right) + p_t K \, dt \\
&= \left[\frac{p_t(K - S_t) + (1 - p)^2 \mu S_t}{1 - p} \right] dt + (1 - p)\sigma S_t \, dC_t.
\end{aligned}
\tag{4.35}
$$

The stochastic differential equation here is, of course, very similar to that of the GBM model. We can again define M and Σ to be

$$M = \frac{p_t(K - S_t) + (1 - p)^2 \mu S_t}{1 - p}, \qquad \Sigma = (1 - p)\sigma \tag{4.36}$$

to give us the common form

$$dS_t = M\,dt + \Sigma\,dC_t, \tag{4.37}$$

and apply our iterative computational method to solve this partial differential equation. The key difference in this model, however, is the fat-tailed conditional probability density function. Using the Cauchy PDF from Eq. 3.34, we find that $f_\epsilon(y; S_{t-\epsilon})$ is given by

$$
\begin{aligned}
f_\epsilon(y; S_{t-\epsilon}) &= \frac{1}{\pi\Sigma\left[1 + \left(\frac{y-M}{\Sigma}\right)^2\right]} \\
&= \frac{\Sigma}{\pi\left[\Sigma^2 + (y-M)^2\right]} \\
&= \frac{(1-p)\sigma}{\pi\left[(1-p)^2\sigma^2 + \left(y - \frac{p_t(K-S_t)+(1-p)^2\mu S_t}{1-p}\right)^2\right]}
\end{aligned} \tag{4.38}
$$

We again assume that μ is the prevailing risk-free rate of return. Then, given appropriate parameters for σ, K, and S_t, we can estimate the probability that a deal will complete at any point in its duration. Figure 4.3 shows estimated deal completion probability curves at different points in time for a sample completed transaction. The probability curves are similar in appearance to those we saw for the GBM-based process in Figure 4.2, except for a few subtleties:

1. The probability curves are narrower under the Cauchy model than the GBM model. This makes sense given our understanding of the Cauchy distribution: even when using comparable distribution scale parameters, the Cauchy mostly suggests very small, near-zero stochastic returns. Thus, the model attributes a lower likelihood of success to a deal in which the target stock is further from the offer price.

2. The peaks of the probability curves are slightly lower than we saw in the GBM model. We expect the model to assign lower probabilities of completion, because it allows for large, occasional price fluctuations (essentially simulating spontaneous deal failure).
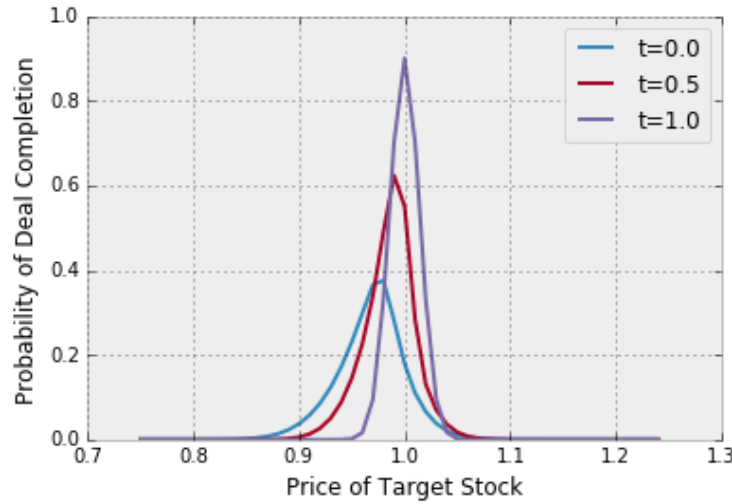

Figure 4.3: Probability of deal completion at various points in the transaction, based on target stock price, assuming a Cauchy unconstrained process with drift.

4.3   Model Testing

Now that we have discussed these three models in depth, we wish to assume the role of an arbitrageur and test the models for predictive power. For simplicity sake, we refer to the three models as 'Brownian', 'GBM', and 'Cauchy'. For each deal in our dataset, we calibrated each of the three models to the target stock price data, at 11 different uniformly distributed points in time $(t \in \{0.0, 0.1, ..., 0.9, 1.0\})^2$. We assigned model parameters as

_____

[2] We do not calculate probability of completion at each point in time due to computational burden.

follows:

- K: We set K to be the deal offer price; given our price normalization, $K = 1$ for all

  deals.

- μ: For the GBM and Cauchy models, we set μ to be the average risk-free rate of return

  in the given year that the deal was announced[3].

- σ: Estimating σ presents a unique difficulty in these models, due to the fact that

  volatility is given by $(1 - p)\sigma$, and p is the unknown variable that we are solving for.

  While we provide our best estimates for σ, based on the volatility of the target stock,

  we recognize that the models may not be perfectly calibrated. Thus, we encourage

  more rigorous parameter fitting in future studies.

- S_t: We simply used the target stock price at each of the 11 points in time that we

  considered $(S_t, t \in \{0.0, 0.1, ..., 0.9, 1.0\})$.

Thus, each of the three models yielded an estimate for expected probability of deal completion

at each of these points in time, for each transaction.

In Figure 4.4(a), we show the price path of a target stock from a completed transaction,

and in Figure 4.4(b), we present each model's estimated probability that the deal will com-

plete, over the course of the transaction. Very clearly, the upward trend in the stock price

is matched by increasing probability estimates that the deal will complete, across all three

models.

---

[3]We used the rate of the 10-Year Treasury Note as our risk-free rate.

(a) Target Share Price
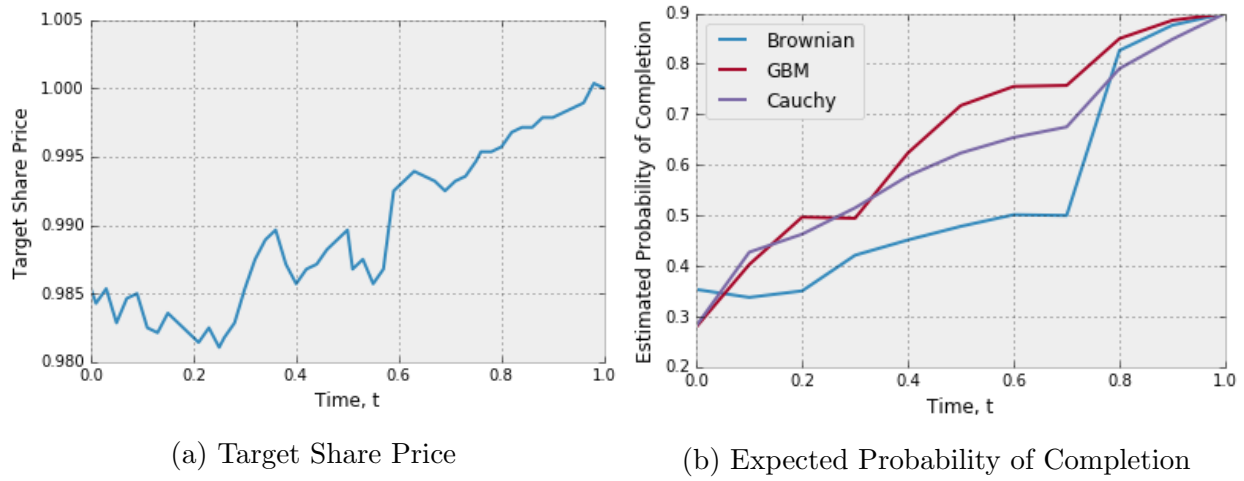

(b) Expected Probability of Completion

Figure 4.4: Comparison of the three models' estimates for probability of completion for a given completed deal.

Similarly, in Figure 4.5(a), we show the price path of a target stock from a terminated transaction, and in Figure 4.5(b), we present each model's estimated probability that the deal will complete. As we would expect, even though the price remains flat for the first half of the deal, the probability of completion decreases across all models; this is because the time until the expected completion date is decreasing. Additionally, we find that for this particular deal, the Brownian and Cauchy models are significantly more conservative in their probability estimation than the GBM model.

Furthermore, in Figure 4.6, we present the average estimated probability of completion across (a) all completed deals and (b) all terminated deals, by each model (note that the y-axis changes between the two plots). From this perspective, the GBM model seems to do the best job of predicting outcomes for completed deals, followed closely by the Cauchy model. On the flip-side, the Cauchy model is significantly better than both the GBM and Brownian

(a) Target Share Price
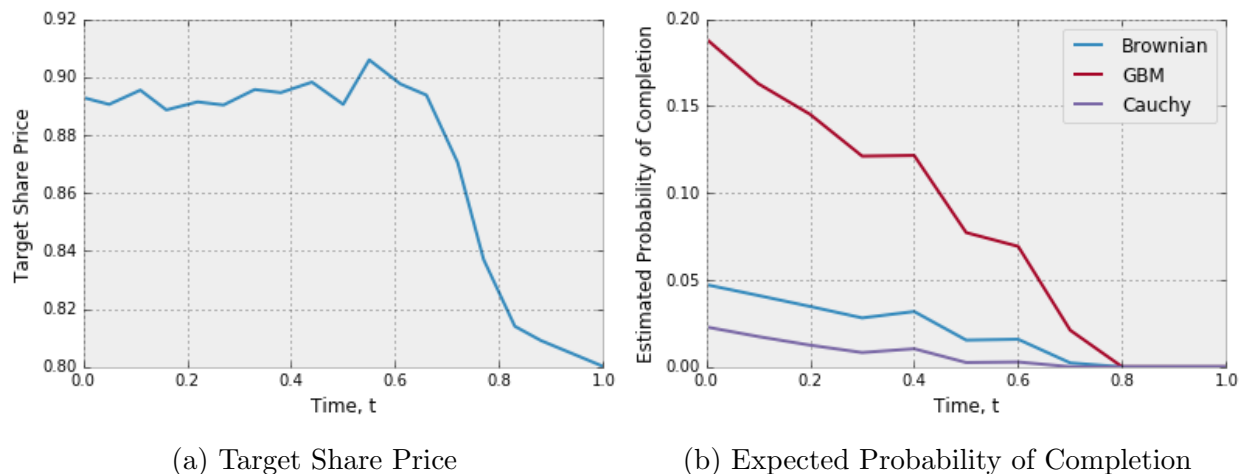
(b) Expected Probability of Completion

Figure 4.5: Comparison of the three models' estimates for probability of completion for a given terminated deal.

models at predicting outcomes for deals that terminated; on average, the Cauchy model gives terminated transactions estimated probabilities that are 5 percentage points lower than the GBM model. Comparing average estimated probability of completion across the dataset is a very rough approach, however, and therefore we proceed with further analysis of the models' performance.

(a) All Completed Deals
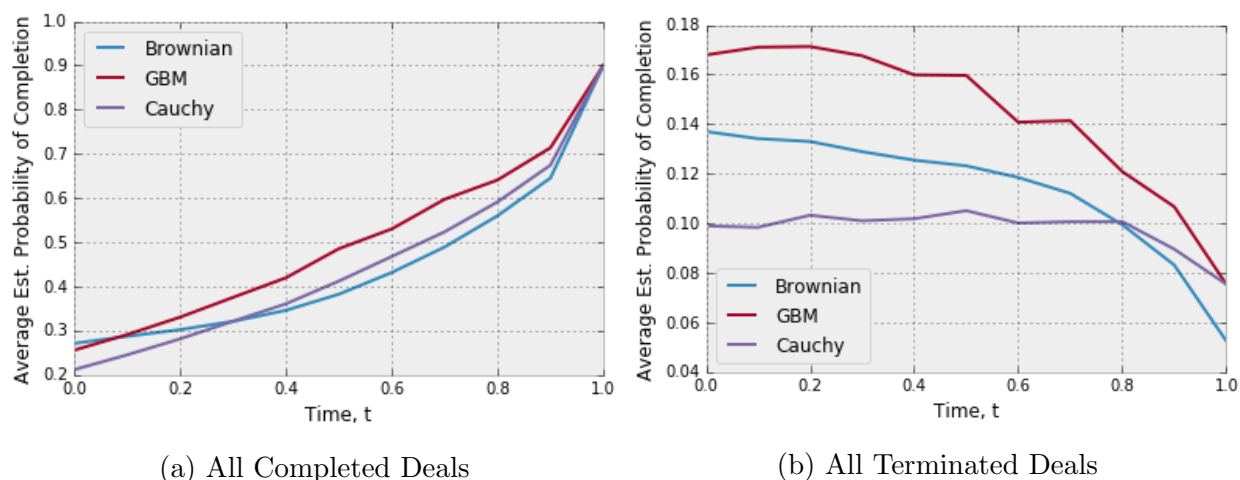
(b) All Terminated Deals

Figure 4.6: Average estimated probability of completion across all deals, presented by model.

To further test the power of these models, we establish probability cutoffs (similarly to the process in Section 3.1) so that we can compare predicted outcomes with actual outcomes. For each model, at each of the 11 points in time, we define the cutoff as halfway between the median estimated completion probability for completed deals and the median estimated completion probability for terminated deals[4]. As an example, under the Brownian model at time $t = 0.50$, the median estimated completion probability for completed deals is 47.8% and the median estimated completion probability for terminated deals is 3.6%. Thus, the probability cutoff is 25.7%; for testing purposes, we consider any transaction with an estimated probability of completion above 25.7% to be completed. In Table 4.1 and in Figure 4.7, we demonstrate the deal outcome prediction success rate for each model, for completed deals and for terminated deals.

Table 4.1: Deal outcome prediction success rate by model.

| Point in Time | Completed Deals | | | Terminated Deals | | |
	Brownian	Geometric Brownian	Cauchy	Brownian	Geometric Brownian	Cauchy
0.00	75%	76%	69%	69%	59%	71%
0.10	76%	76%	77%	70%	62%	70%
0.20	78%	74%	75%	71%	68%	72%
0.30	73%	81%	76%	79%	65%	76%
0.40	75%	80%	69%	80%	73%	82%
0.50	78%	77%	74%	79%	80%	83%
0.60	75%	73%	73%	84%	87%	87%
0.70	77%	79%	75%	85%	84%	88%
0.80	81%	80%	80%	85%	86%	86%
0.90	80%	85%	85%	91%	89%	89%
1.00	100%	100%	100%	93%	93%	93%

[4]This cutoff was determined based on repeated testing to maximize predictive power.

Considering the noise in the results, we find that all three models perform similarly well when predicting outcomes for deals that completed; the prediction success rate hovers between 75% and 80%. The Cauchy model performs marginally worse than the other two models at various points in the transaction. When it comes to predicting outcomes for deals that terminated, however, we see that the Cauchy model has a slightly higher success rate than the Brownian model and a significantly higher success rate than the GBM model. The prediction success rate for deals that terminated hovers between 70% and 90%.

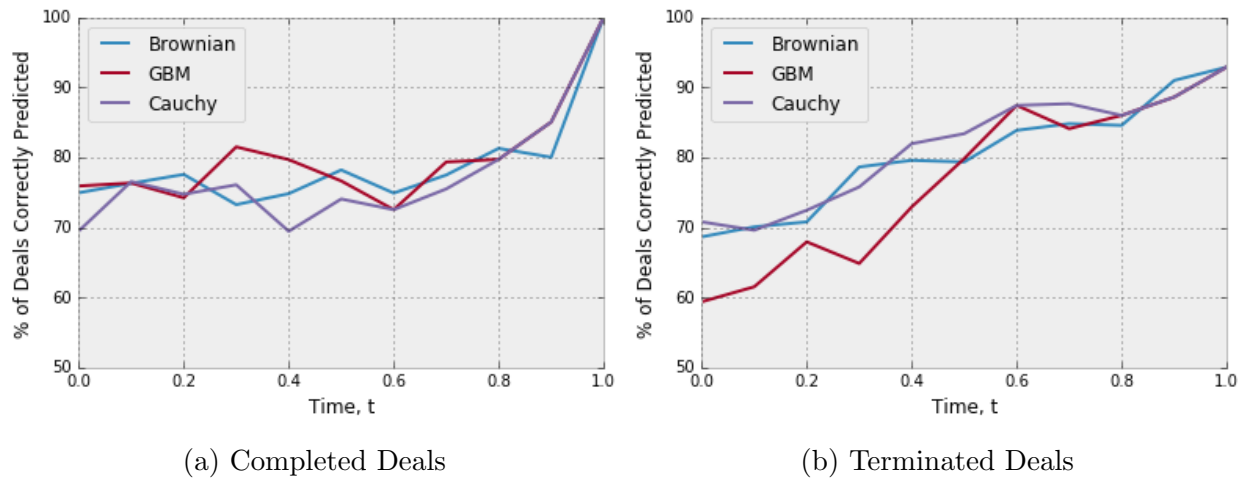(a) Completed Deals   (b) Terminated Deals

Figure 4.7: Deal outcome prediction success rate by model.

As we discussed in the previous chapter (see Table 3.3), a very important result for an arbitrageur is the portion of deals predicted to complete that actually completed; we reiterate, even if the model correctly predicted 100% of outcomes for completed deals, it is useless if it also claimed that the terminated deals would have completed. Thus, in Table 4.2 and in Figure 4.8, we present the portion of completed deals that actually completed out of those that were predicted to complete. We find that the ratios dramatically improve over

the course of the deal; this is due to increasing prediction success rates for both completed

deals and terminated deals. While all models perform well, the Cauchy model does produce

subtly higher ratios, likely due to its superior prediction of outcomes for terminated deals.

Table 4.2: Portion of deals predicted to complete that actually completed.

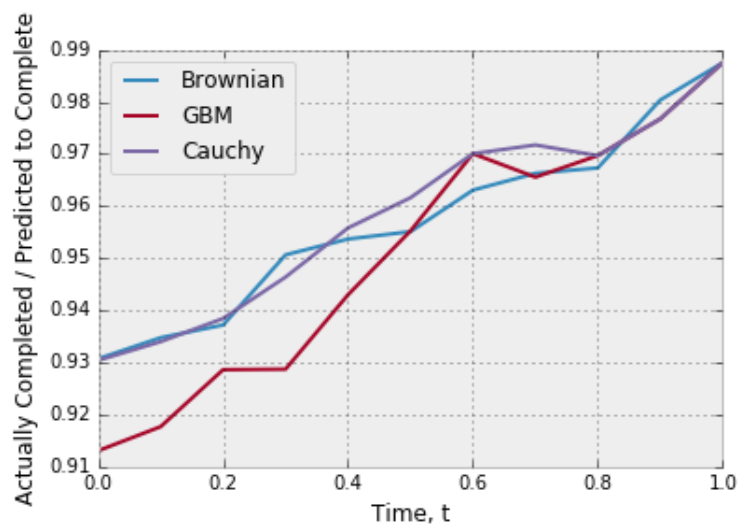| Point in Time	Brownian	Geometric Brownian	Cauchy
0.00	93.1%	91.3%	93.0%
0.10	93.5%	91.8%	93.4%
0.20	93.7%	92.9%	93.8%
0.30	95.1%	92.9%	94.6%
0.40	95.4%	94.3%	95.6%
0.50	95.5%	95.5%	96.2%
0.60	96.3%	97.0%	97.0%
0.70	96.6%	96.6%	97.2%
0.80	96.7%	97.0%	97.0%
0.90	98.0%	97.7%	97.7%
1.00	98.7%	98.7%	98.7%

Figure 4.8: Portion of deals predicted to complete that actually completed.

Analyzing Incorrect Predictions

Having established that these models can predict outcomes for the majority of transactions, we wish to understand when and why they fail. We find that the Brownian, GBM, and Cauchy models, halfway through the deal, respectively incorrectly predict 25%, 23%, and 26% of completed deals, as terminated. At the same point in time, 21%, 20%, and 16% of terminated deals, respectively, are incorrectly predicted as completed. Figure 4.9 demonstrates price paths for two target stocks that ended up completing. For both of these target stocks, the models predicted near 0 probability of completion for the majority of the deal. Looking at the price paths, it is very understandable that the models incorrectly predicted the outcomes: in both deals, there was a clear downward price trajectory for the first half of the deal. Looking at Figure 4.9(a) in particular, we see that the target stock price actually dropped as low as 70% of the offer price. This type of price movement suggests a sharp reversal in market sentiment due to new information, potentially related to regulatory approvals or management decisions. These circumstances are naturally very difficult to predict in a generalizable fashion, and instead require case-by-case consideration.

Similarly, in Figure 4.10, we demonstrate price paths for two target stocks in deals that eventually terminated. We see that, over the course of the transactions, the three models predicted probabilities of completion between 20% and 70%, only reverting towards 0% in the last stages of the deal. Again, it is fairly clear why the models failed to predict the deal outcomes earlier on. The target stock price paths followed very clear upward trends and

(a) Target Share Price

(b) Expected Probability of Completion

(c) Target Share Price
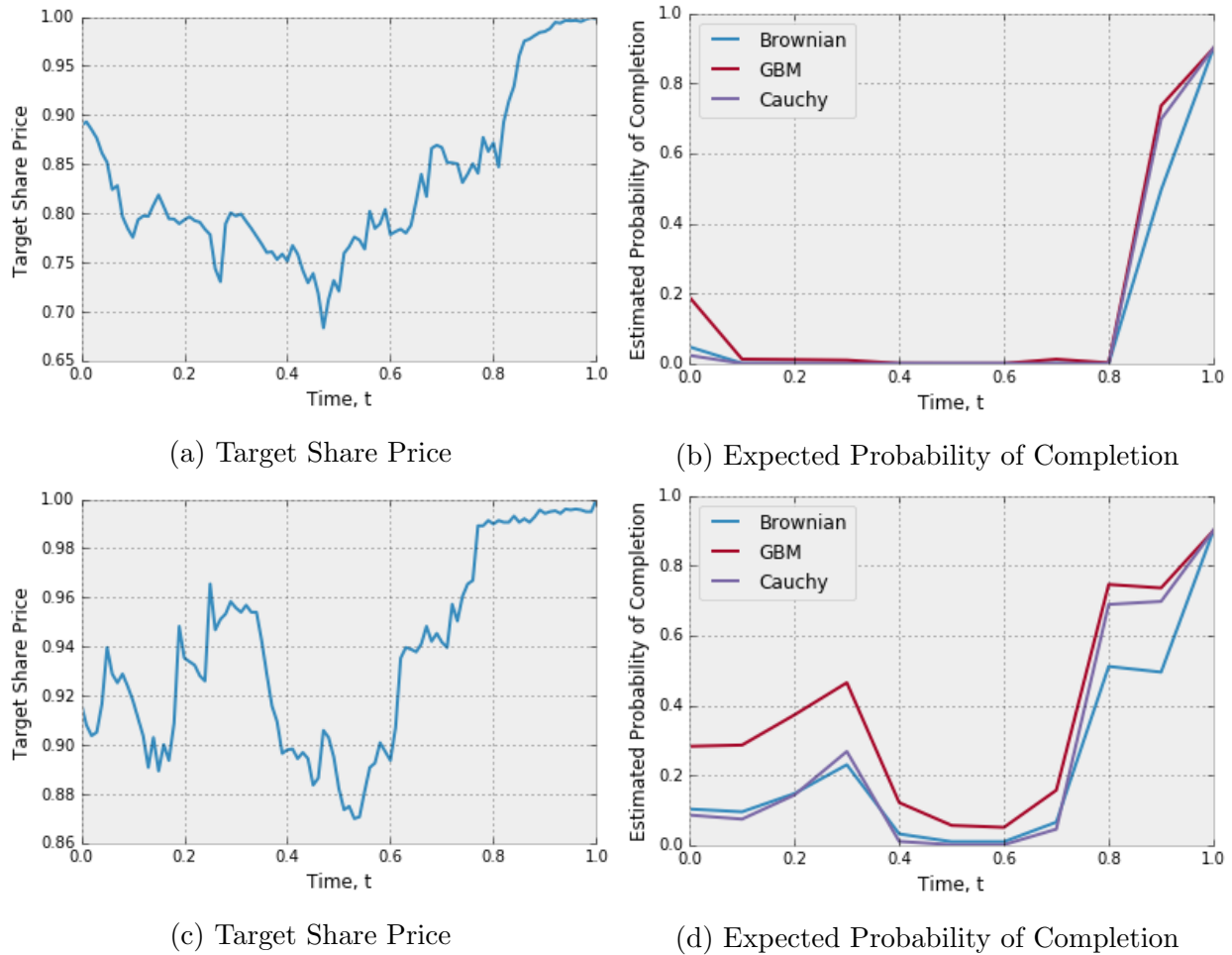

(d) Expected Probability of Completion

Figure 4.9: Comparison of the three models' incorrect estimates for probability of completion for two deals that actually completed.

hovered near the offer price for substantial portions of the deal duration. In fact, the target

stock shown in Figure 4.10(c) did not price correct away from the offer price until a massive

drop in the last days of the deal. This pattern was consistent across the terminated deals

that our models incorrectly predicted would complete: there was often no indication that the

deal would not complete until the last moments of the transaction. While these price shocks

are impossibly difficult to predict with accuracy in generalized models, using the fat-tailed

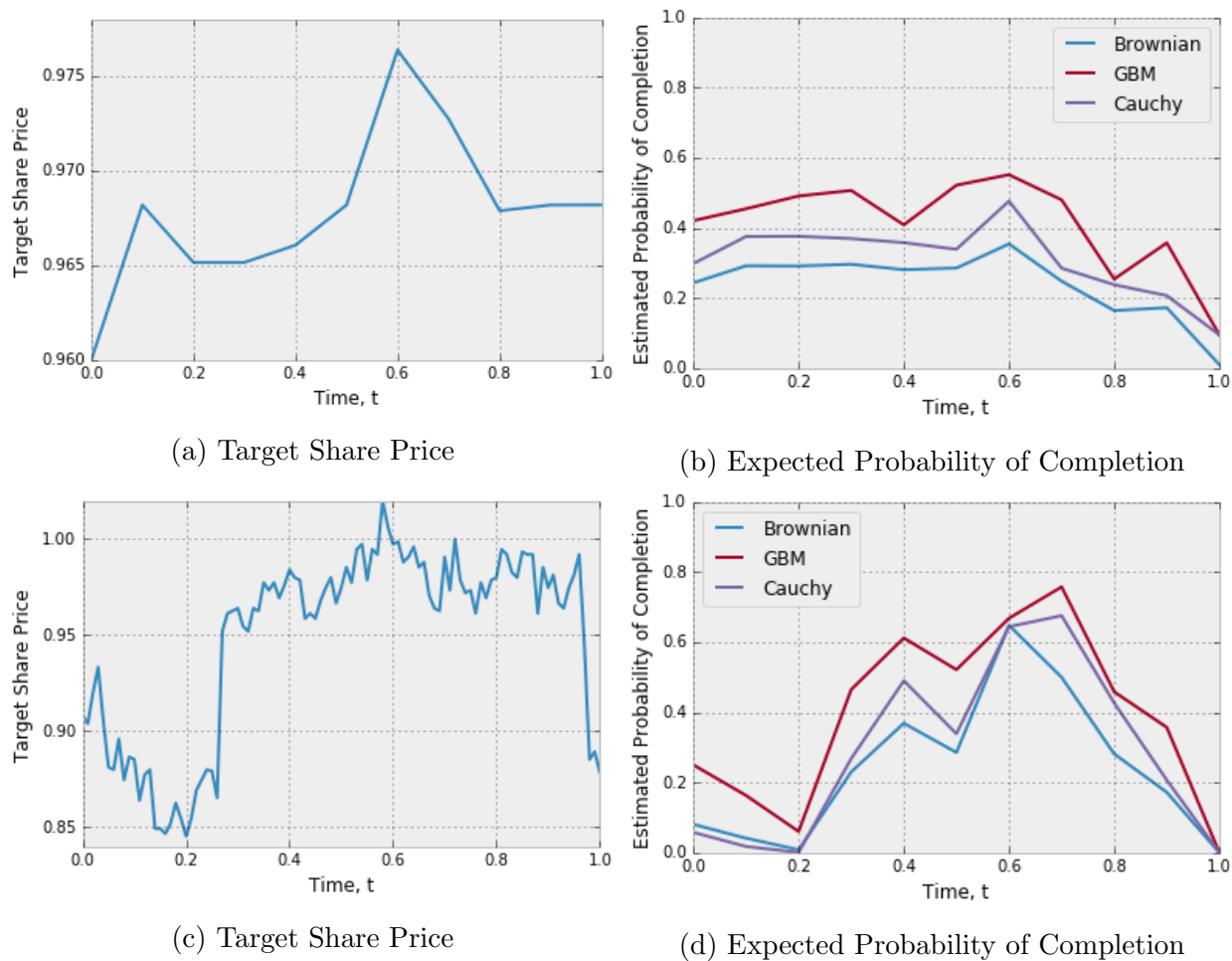distribution in the Cauchy model did yield better results.

(a) Target Share Price

(b) Expected Probability of Completion

(c) Target Share Price

(d) Expected Probability of Completion

Figure 4.10: Comparison of the three models' incorrect estimates for probability of completion for two deals that actually terminated.

4.4   Discussion

In this chapter, we demonstrated our second approach to estimating the probability that a transaction will successfully complete. Building on our work from previous chapters, we

developed a foundational model to represent all target stocks in transactions. Our model was built on the assumption that the target stock price necessarily reflects the expected probability that the deal will complete, and depending on this probability, the target stock price either approaches the offer price or follows an unconstrained stochastic process. We began by defining the unconstrained process as standard Brownian motion, and then extended the model to include geometric Brownian motion, as well as a Cauchy process with drift. By integrating over all paths in a discretized grid, we solved these equations to yield the probability that a deal would complete at a given time, given various specific target stock parameters.

In general, we found that all three models were fairly strong. The GBM model tended to have slightly better success predicting outcomes for completed transactions, due to its wider probability curves. The model thus attributed a higher probability of completion to target stocks that were not necessarily close to the offer price. The principal shortcoming of the Brownian and GBM models, however, was in predicting outcomes for transactions that terminated. Figure 4.6 showed that the GBM model, on average, gave deals that terminated a 17% chance of success at $t = 0.2$, whereas the Cauchy model accurately estimated only a 10% chance of success. Figure 4.7(b) provided further evidence for the Cauchy model's prowess in predicting outcomes for deals that would terminate. We attribute this success to the use of fat-tails to represent stochastic price movements; the standard Brownian and GBM models simply could not reproduce the price crashes seen in so many terminated deals. In fact, we showed in Figure 4.8 that these properties actually make the Cauchy a superior

model for arbitrageurs. We estimated a prediction success rate of roughly 75% for completed deals and roughly 85% for terminated deals. Of the deals predicted to complete by the model, ∼96% of them actually completed.

Chapter 5

Conclusion

In the universe of event-driven investment strategies, merger arbitrage has prevailed for decades. Attracted by excess returns that are minimally impacted by broader market fluctuations, institutional investors and hedge funds have long sought innovative strategies to predict M&A deal outcomes. The ensuing body of academic work has principally focused on the construction and analysis of logistic regression models. While the qualitative factors that may indicate deal success have been well studied, few attempts have been made to quantitatively model the dynamics of target company stock prices in a transaction. This paper seeks to fill this void; we employ stochastic processes to offer two novel approaches to predicting deal outcomes. Furthermore, we explore the use of fat-tailed distributions to model instantaneous stock returns, diverging from the majority of Brownian-based financial literature.

In our first approach, we construct various stochastic models to represent target stocks

in completed transactions. We begin by conducting a thorough analysis of a dataset of 2,787 historical transactions, from which we identify two principal forces acting on target stocks in completed deals: 1) a long-term mean-reversion effect, and 2) a short-term stochastic returns process. We find that the short-term stochastic component draws from a leptokurtic, fat-tailed distribution of returns, rather than a standard Normal. Additionally, we observe several particular characteristics of target stock price paths, including flat variance in price level across deals and decreasing volatility in stochastic movements over time. We present various models of increasing complexity to represent these characteristics and examine calibration methods for each. Our final model consists of an Ornstein-Uhlenbeck process modified with a Cauchy process and Wiener mean-reversion. Via Monte Carlo methods, we calibrate these models to target stocks in our dataset and test for predictive power. We demonstrate that a 75% prediction success rate can be achieved for ongoing completed transactions. The empirical findings support the use of fat-tailed distributions rather than the traditional Gaussian for short-term returns.

In our second approach, we broaden our scope and construct a general model to represent target stocks in all transactions. Our model endogenously considers the expected probability of deal completion as a factor in the target share price: the probability of deal completion, when large, motivates movement towards the deal offer price, and when small, towards an unconstrained process. We present various options for the unconstrained process, including standard Brownian motion, geometric Brownian motion, and a fat-tailed Cauchy process with drift. Using computational methods, we present solutions that yield the expected

probability of deal completion at a given time. We find that the Cauchy process demonstrates a significant advantage in predicting outcomes for deals that will terminate. Our final model has a roughly 75% prediction success rate for completed transactions and roughly 85% for terminated transactions.

In sum, the empirical results of this paper present two key findings:

1. Stochastic processes can effectively model the share price of a target company in an M&A transaction, and can further be used to robustly predict deal outcomes.

2. Diverging from the body of financial literature, we find that fat-tailed distributions represent an improved model for stochastic price movements, relative to standard Gaussian distributions. This especially holds true in situations where large price fluctuations are expected (e.g. M&A transactions).

Further research on the applicability of other models, particularly stochastic volatility processes, to merger arbitrage situations could yield greater predictive power. While this paper principally worked with the Cauchy distribution, an examination of other steady, fat-tailed distributions may provide further insight. Additionally, critical to our results were the calibration processes and approximate computational methods we used; further investigation into analytical solutions is encouraged. Lastly, our work did not consider the actual returns to an investment strategy based on the models we produce; assessing the potential profitability of these models using portfolio back-testing is a natural next step.

Bibliography

Bachelier, Louis et al. (2006). *Louis Bachelier's Theory of speculation: the origins of modern finance.* Princeton University Press.

Baker, Malcolm and Serkan Savasoglu (2002). "Limited arbitrage in mergers and acquisitions". In: *Journal of Financial Economics* 64.1, pp. 91–115. DOI: `10.1016/s0304-405x(02)00072-7`.

Barndorff-Nielsen, Ole E. and Neil Shephard (2001). "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.2, pp. 167–241. DOI: `10.1111/1467-9868.00282`.

Barone-Adesi, Giovanni, Keith Brown, and W. V. Harlow (1994). *On the Use of Implied Stock Volatilities in the Prediction of Successful Corporate Takeovers.*

Beckers, Stan (1981). "A Note on Estimating the Parameters of the Diffusion-Jump Model of Stock Returns". In: *The Journal of Financial and Quantitative Analysis* 16.1, p. 127. DOI: `10.2307/2330670`.

Bester, Alan, Victor Hugo Martinez, and Ioanid Rosu (2011). "Option Pricing and the Probability of Success of Cash Mergers". In: *SSRN Electronic Journal*. DOI: `10.2139/ssrn.1364491`.

Black, Fischer and Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3, pp. 637–654. DOI: `10.1086/260062`.

Branch, Ben S and Jia Wang (2008). "Risk-Arbitrage Spreads and Performance of Risk Arbitrage". In: *The Journal of Alternative Investments* 11.1, pp. 9–22. DOI: `10.3905/jai.2008.708847`.

Branch, Ben and Jia Wang (2009). "Takeover Success Prediction and Performance of Risk Arbitrage". In: *Journal of Business & Economic Studies* 15.2, pp. 10–25.

Branch, Ben and Taewon Yang (2010). "Estimating the Profit Potential of Risk Arbitrage Opportunities". In: *Pan-Pacific Journal of Business Research* 1.1, pp. 22–40.

Brown, Keith C. and Michael V. Raymond (1986). "Risk Arbitrage and the Prediction of Successful Corporate Takeovers". In: *Financial Management* 15.3, p. 54. DOI: `10.2307/3664844`.

Christensen, Clayton et al. (2011). "The Big Idea: The New M&A Playbook". In: *Harvard Business Review*.

Cont, Rama and Peter Tankov (n.d.). "Calibration of Jump-Diffusion Option Pricing Models: A Robust Non-Parametric Approach". In: *SSRN Electronic Journal*. DOI: `10.2139/ssrn.332400`.

Fama, Eugene F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2, pp. 383–417. DOI: 10.2307/2325486.

Garbaczewski, Piotr and Robert Olkiewicz (2000). "Ornstein-Uhlenbeck-Cauchy process". In: *Journal of Mathematical Physics* 41.10, p. 6843. DOI: 10.1063/1.1290054.

Hoffmeister, J. Ronald and Edward A. Dyl (1981). "Predicting Outcomes of Cash Tender Offers". In: *Financial Management* 10.5, pp. 50–58. DOI: 10.2307/3664856.

Hsieh, Jim and Ralph A. Walkling (2005). "Determinants and implications of arbitrage holdings in acquisitions". In: *Journal of Financial Economics* 77, pp. 605–648. DOI: 10.2139/ssrn.571022.

Jetley, Gaurav and Xinyu Ji (2010). "The Shrinking Merger Arbitrage Spread: Reasons and Implications". In: *Financial Analysts Journal* 66.2, pp. 54–68. DOI: 10.2469/faj.v66.n2.3.

Jindra, Jan and Ralph A. Walkling (2004). "Speculation spreads and the market pricing of proposed acquisitions". In: *Journal of Corporate Finance* 10.4, pp. 495–526. DOI: 10.1016/s0929-1199(03)00030-0.

Khaldi, K., K. Djeddour, and S. Meddahi (2014). "Comparison of Jump-Diffusion Parameters Using Passage Times Estimation". In: *Journal of Applied Mathematics* 2014, pp. 1–5. DOI: 10.1155/2014/975418.

Larcker, David F. and Thomas Lys (1987). "An empirical analysis of the incentives to engage in costly information acquisition". In: *Journal of Financial Economics* 18.1, pp. 111–126. DOI: 10.1016/0304-405x(87)90063-8.

Lin, Luke, Li-Huei Lan, and Shuang-Shii Chuang (2012). "An Option-Based Approach to Risk Arbitrage in Emerging Markets: Evidence from Taiwan Takeover Attempts". In: *Journal of Forecasting* 32.6, pp. 512–521. DOI: `10.1002/for.2250`.

Lu, Chaoxiao and Stephen Yau (2008). "Pricing Stock Options in Mergers and Acquisitions with Jump-diffusion Model". In: *2008 American Control Conference*, pp. 1008–1012. DOI: `10.1109/acc.2008.4586623`.

Mahdizadeh, M. and Ehsan Zamanzade (2016). "Goodness-of-fit testing for the Cauchy distribution with application to financial modeling". In: pp. 1–20.

Maheswaran, Krishnan and Soon Chin Yeoh (2005). "The Profitability of Merger Arbitrage: Some Australian Evidence". In: *Australian Journal of Management* 30.1, pp. 111–126. DOI: `10.1177/031289620503000106`.

Matsuda, Kazuhisa (2004). *Introduction to Merton Jump Diffusion Model*.

Merton, Robert C. (1976). "Option pricing when underlying stock returns are discontinuous". In: *Journal of Financial Economics* 3.1-2, pp. 125–144. DOI: `10.1016/0304-405x(76) 90022-2`.

Mitchell, Mark and Todd Pulvino (2001). "Characteristics of Risk and Return in Risk Arbitrage". In: *The Journal of Finance* 56.6, pp. 2135–2175. DOI: `10.1111/0022-1082.00401`.

Mudchanatongsuk, Supakorn, James A. Primbs, and Wilfred Wong (2008). "Optimal Pairs Trading: A Stochastic Control Approach". In: *2008 American Control Conference*. DOI: `10.1109/acc.2008.4586628`.

Press, S. James (1967). "A Compound Events Model for Security Prices". In: *The Journal of Business* 40.3, pp. 317–335. DOI: `10.1086/294980`.

Reid, Stuart (2015). *Random walks down Wall Street, Stochastic Processes in Python.* URL: `http://www.turingfinance.com/random-walks-down-wall-street-stochastic-processes-in-python/`.

Subramanian, Ajay (2004). "Option Pricing on Stocks in Mergers and Acquisitions". In: *The Journal of Finance* 59.2, pp. 795–829. DOI: `10.1111/j.1540-6261.2004.00649.x`.

Tankov, Peter and Ekaterina Voltchkova (2009). "Jump-diffusion models: a practitioner's guide". In: *Banque et Marches* 99.

Tassel, Peter Van (2016). "Merger Options and Risk Arbitrage". In: *Federal Reserve Bank of New York Staff Reports* 761.

Taylor, Howard M. and Samuel Karlin (2009). *An introduction to stochastic modeling.* Academic Press.

Van den Berg, Thijs (2011). *Calibrating the Ornstein-Uhlenbeck (Vasicek) model.* URL: `https://www.sitmo.com/?p=134`.

Wilfling, Bernd and Sergey Gelman (2009). "Option Pricing on Target Stock Under Multiple Decision Reversions". In: *SSRN Electronic Journal.* DOI: `10.2139/ssrn.1343976`.

Wilmott, P., S. Howison, and J. Dewynne (1995). *Mathematical models of financial derivative products: a student introduction.* Cambridge University Press.