Models for measuring and predicting value creation during merger and acquisitions: a study of bank industry

Guo-yi Chen*
Chongqing three gorges university, Wanzhou City, Chongqing, China, 404120
Received 6 September 2014, www.cmnt.lv

Abstract

This paper employs event study methodology with a 36-day event window to assess the value effects of the US bank mergers occurring between 1994 and 2003. A 38-transaction sample is chosen from the top fifty US bank mergers (according to the assets of targets) occurred during the period from 1994 to 2003. Through analysis, result indicates that the average cumulative abnormal return of the bidders in the chosen sample is negative (-0.99%), while the targets and combined firms are both positive (15.07% and 2.57% respectively). Significance testing also verified that the negative bidder return is confirmed to be insignificant, whilst the positive return of the target and combined firm are both significant. Combined together, It indicates that the 3,517 US bank mergers occurred between 1994 and 2003 create insignificantly negative value for bidders, whilst benefit the target and the integrated banks with significant positive gains.

Keywords: mergers, acquisitions, event study, value effects, measurement model

1 Introduction

The US bank sector has experienced a couple of merger waves. The first one occurred in the 1960s and the second started from the late 1980s and has not finished yet [1]. During the second wave, till 1998, the number of US bank mergers accounted for 15 percent of all public mergers compared to 8 percent before [2]. Such outstanding proportion encourages researchers to search for the reasons of the popularity of M&A in the US bank sector, particularly in terms of the value effects, i.e. whether M&A creates value for the participants and its magnitude.

This study chooses stock market reaction approach as the method to assess the value effects of US bank mergers. A 38-transaction sample is chosen from the top fifty mega-bank mergers occurring from 1994 to 2003, which has not been analysed comprehensively by previous researchers. A 36-day event window is adopted for the event study considering the trade-off between the market efficiency and avoiding the impacts of irrelative noisy events. The stock market data are collected through software Datastream Advance 3.5. This paper observes negative average cumulative abnormal return for bidders and positive ones for both targets and their combinations. Further, t-test, sign test and Wilcoxon signed rank test are used to test the statistical significance of the above results.

2 Establishing value creation measurement model

The primary methodology of this research is event study. Event study is an empirical study of prices of an asset just before and after some event, like an announcement, of merger or dividend. It contains three main steps, namely identifying the event, calculating the average cumulative abnormal return and testing its statistical significance [3]. This paper follows Becher’s (2000) 36-day event window (30 trading days prior to and 5 trading days after the announcement date) to test the value effects of bank mergers [4]. 5-day after the announcement date takes account of the under reaction of the stock market and, on the other hand, beneficial for avoiding the unwanted noisy events. 30-day prior to the event aims to involve the advance market reaction caused by the semi-strong form efficiency of the realistic stock market.

2.1 EVENT IDENTIFICATION

The sample used in this study initially contains fifty US bank mergers occurred between 1994 and 2003. They are the top fifty US bank mergers according to the ranking of the targets’ assets. Through searching for relative data, including announcement date, share price, market value and S&P 500 Composite Index, there are eventually 38 transactions whose required data are completely available. The rest 12 cases are abandoned due to the lack of data of either the merging or merged firm. In this case, in Datastream Advance 3.5, which is the main source of stock market data for this study, the record under the current name actually reflects the historical data of the merging bank before and after the combination. The record of the acquired bank has thus been covered and cannot be found any more. This situation is severe especially for those banks that experienced several mergers and changed their name in the preceding way. The typical examples include Firstar Corporation and JPMorgan Chase. The main reason of choosing a sample comprised of large bank mergers lies in that it represents the most influential bank mergers in this ten-year period. The distribution of the sampled 38
bank mergers across years is illustrated by the following Table.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>2</td>
</tr>
<tr>
<td>1995</td>
<td>3</td>
</tr>
<tr>
<td>1996</td>
<td>7</td>
</tr>
<tr>
<td>1997</td>
<td>5</td>
</tr>
<tr>
<td>1998</td>
<td>8</td>
</tr>
<tr>
<td>1999</td>
<td>2</td>
</tr>
<tr>
<td>2000</td>
<td>1</td>
</tr>
<tr>
<td>2001</td>
<td>7</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
</tr>
<tr>
<td>2003</td>
<td>1</td>
</tr>
</tbody>
</table>

2.2 CUMULATIVE ABNORMAL RETURNS

After data collection, the next step is to calculate abnormal returns. As explained by Frame and Lastrapes, abnormal return is calculated by subtracting estimated normal return from actual return during the event period [5]. Estimated normal return may be attained through three most widely used methods, namely single-index model, Capital Asset Pricing Model (CAPM), and Market Model [6]. This study follows Becher (2000) to employ single-index model in estimating normal return. The standard and poor (S&P) 500 Composite Index is adopted in this study. S&P 500 Composite Index is a market value-weighted index containing 500 widely held common US stocks that measures the general performance of the stock market. Through observing its daily figures over the event period, the estimated normal return can be obtained by calculating the difference of the S&P 500 Composite Index between a certain day (n) and its previous day (n-1), and then dividing the difference by the index of the previous day (n-1),

\[
NR_n = \frac{I_n - I_{n-1}}{I_{n-1}},
\]

where \(NR_n\) is the normal return of day \(n\), \(I_n\) is the S&P 500 Composite Index of day \(n\) and \(I_{n-1}\) is that of day (n-1). The similar Equation is used to compute the change of share price (the actual return).

\[
\Delta P = \frac{P_n - P_{n-1}}{P_{n-1}},
\]

where \(\Delta P\) refers to the change in share price of day \(n\) from that of its previous day (n-1); \(P_n\) is the share price in day \(n\) and \(P_{n-1}\) is that in day (n-1). Based on these calculations, the abnormal return (\(AR_n\)) can be attained through subtracting the normal return from the change of share price.

\[
AR_n = \Delta P_n - NR_n.
\]

The daily data of both S&P 500 Composite Index and stock price are collected through the software Datastream Advance 3.5. The cumulative abnormal return (CAR) can be computed using either arithmetic or geometric process. The arithmetic process computes the sum of the daily abnormal returns.

\[
CAR = \sum_{n=30}^{5} AR_n.
\]

(4)

Compared to the simply summing, geometric process multiplies the abnormal return of a certain day (n) with the geometric CAR of its previous day (n-1), and then subtract 1 (Becher 2000).

\[
CAR_n = (1 + AR_n) \times (1 + CAR_{n-1}) - 1.
\]

(5)

Therefore, the geometric CAR over the selected event window, i.e. the geometric CAR of day (+5), is actually the product of every daily abnormal return minus 1.

\[
CAR = \prod_{n=30}^{5} (1 + AR_n) - 1.
\]

(6)

In order to calculate the cumulative abnormal return of the combined bidder and target (CCAR), this study follows the method of Houston and Rygaert (1994), which is also employed by Becher (2000). The Equation is as below:

\[
CCAR = \frac{MV_b \times CAR_b + MV_t \times CAR_t}{MV_b + MV_t}.
\]

(7)

where \(MV_b\) and \(MV_t\) are respectively the market value of the bidder and target 30 days prior to the announcement date, which are also collected from Datastream Advance 3.5. The CCAR is the market value-weighted CAR of the combined bidder and target. It takes account of not only their separate CAR over the event window, but also their joint contributions towards the integrated CAR from the stock market. Being weighted by their market values at the beginning of the 36-day event window, the combined cumulative abnormal return reflects the cumulative abnormal increase in terms of the combined market value. This is an effective representative of the market reaction to the bank merger.

From the mathematic perspective, this method exactly calculates the abnormal increase of the combined market value over the event period. It can be proved through the following calculations. Firstly, the change of market value (\(\Delta MV\)) equals the market value of the last day minus that of the first day of the event window.

\[
\Delta MV = MV_f - MV_i.
\]

(8)

where \(MV_f\) represents the market value at the end of the event window (in this case the day +5), and \(MV_i\) represents that at the beginning (day –30). Subtracting the normal increase of the market value (\(NMV\), the increase derived from the normal return) from it, \(\Delta MV\) can be transferred to abnormal increase of market value (\(\Delta AMV\)), that is:
The abnormal increase of market value (AMV) can be expressed by their abnormal increase in market value, which reflects the integrated market reaction to the bank merger. Alternatively, this relationship can also be proved by simply combining Equation (13) and Equation (14):

\[ \text{CAMV} = \frac{\text{AMV}}{\text{MV}} = \frac{\text{MV}_0 \times \text{CAR}_0 + \text{MV}_t \times \text{CAR}_t}{\text{MV}_0 + \text{MV}_t} \]

Equation (15) is identical with Equation (8) which is used to determine the cumulative combined abnormal return. Therefore, the combined cumulative abnormal return determined by weighting their CARs is equivalent to the combined abnormal increase of market value, which reflects the integrated market reaction to the bank merger. Alternatively, this relationship can also be proved by simply combining Equation (13) and Equation (14):

\[ \text{CAMV} = \frac{\text{AMV}}{\text{MV}_0} = \frac{\text{MV}_0 \times \text{CAR}_0}{\text{MV}_0} = \text{CAR}_0 \]

Most of these sampled transactions occurred at different time. Even if being announced at the same day, such as NationsBank’ merger with BankAmerica and BankOne acquiring First Chicago NBD, the transactions did not dependent on each other. From this perspective, all the transactions in the sample can be regarded as independent and therefore the average combined cumulative abnormal returns of the sample (CCAR) can be expressed by their arithmetic mean.

\[ \text{CCAR} = \frac{1}{N} \sum_{n=1}^{N} \text{CCAR}_n \]

where \( N \) stands for the sample size.

2.3 SIGNIFICANCE TESTING

The final stage is to test the statistical significance of the sample mean (CCAR). There are 3,517 bank mergers occurring in America from 1994 to 2003 [7]. What is chosen is thus a small sample relative to the population. After calculating the mean of the cumulative abnormal returns of all the 38 transactions, it is necessary to test the statistical significance of the result. This study applies t-test, sign test and Wilcoxon signed rank test in this stage. The reasons of choosing these three methods and their respective procedures are demonstrated as follow.

The t-test is the standard and most sensitive test for interval data [8]. It deals with the issues associated with the inference based on small samples. However, the t-test is subject to the assumption of a normal distribution. According to the central limited theorem that the mean of a sample is normally distributed if the size of the sample is over 30 [9], the 38-transaction sample in this study is large enough to consist with the required assumption and hence the result of the t-test should be valid.

The standard Equation of t-statistic is:

\[ t = \frac{\bar{x} - u}{s \sqrt{n}} \]

where \( \bar{x} \) is the sample mean and \( u \) is the population mean; \( n \) is the sample size and \( s \) is the standard deviation [10]. Another parameter is p-value, which indicates how accurately the regression coefficient has been estimated. The p-value can be obtained through checking the Confidence Limits Table according to the t-statistic.

These two parameters above will assist the hypothesis testing. The null hypothesis will be set as that the population mean equals zero:

\[ H_0: \mu = 0 \]

Using the t-test and the corresponding p-value, we can decide whether to reject or accept the null hypothesis as well as how accurate the decision is. If the t-statistic is located within the confidence interval (under the condition of 37, \( N-1 \)), degrees of freedom and 5 percent significance level, i.e. \( p \leq 5\% \), we will accept the null hypothesis [11]. Otherwise, if \( p < 5\% \) the null hypothesis is unacceptable, i.e. \( u \neq 0 \). The p-value will provide the exact degree of significance of this decision. To calculate the p-value, according to Watsham and Parramore (1997), this paper firstly identify that the t-statistic can be interpolated in the interval between \( \beta_1 \) and \( \beta_2 \) [12].

816
Through checking the Confidence Limits Table, this interval corresponds to a interval of significance level between $a_1$ and $a_2$. Then the calculation of p-value can be based on the following Equation:

$$p = a_1 - (t - \beta) \times (a_2 - a_1).$$  

(19)

In this study, both the t-test and p-value will be calculated with the assistance of the statistical software Stata-8.

In order to ensure the accuracy of the statistical significance testing, this study also employs two nonparametric statistics, Wilcoxon signed rank test and sign test. Although the 38-transaction sample can be treated to consist with the normal distribution assumption, the actual distribution is still unknown. Under this circumstance, it is necessary to apply nonparametric statistics which are most appropriate when the sample is small. In contrast with t-test, sign test is not subject to any assumption on distribution and hence treated as the most insensitive test and also the most convincing and easiest test to be applied. It only uses the signs to test the null hypothesis that the number of + signs (the probability of observing a + sign is denoted as $p$, + sign represents the positive abnormal return) is equal to the number of - signs (the probability of observing a - sign is denoted as $q$, - sign represents the negative abnormal return), i.e. H0: $p=q=0.5$. The relationship between $p$ and $q$ consists with binomial distribution because the possibility of observing a + sign plus that of observing a - sign is equal to 1, i.e. $p+q=1$. Therefore we can employ the Equation of binomial distribution to calculate the p-value of each probability of observing + signs ($x$, $x=n$, $n-1$,…,0) in this sample:

$$p(x) = \frac{x!}{x!(n-x)!} \cdot (p^x \cdot q^{n-x})$$  

(20)

We choose a certain $p_i$ which makes $p(x \geq p_i)$, i.e. $p(n)+p(n-1)+\cdots+p(p_i)$, the closest probability to 5 percent level(degree of significance). If the p-value of real number of + signs in the sample ($p(x \geq p_i)$) falls in the critical region, i.e. $p(x \geq p_i) \leq p(x \geq p_i)$ (or <5%), we reject the null hypothesis. Otherwise, if $p(x \geq p_i) \geq p(x \geq p_i)$ (or >5%) the null hypothesis will be accepted.

Wilcoxon test takes account of not only the signs but also the magnitude of each sign through calculating the rank sums. Thereby, it is regarded as a more sensitive test than sign test. Moreover, it is more sensitive than t-test for small samples with unknown distributions [13] as it does not require estimating the parameters of the values (such as the standard deviation of the population) and the distribution of the sample.

In Wilcoxon test, the elements are ranked according to their absolute values (ascending sort). Then the ranks of the positive elements are summed (the sum is denoted by V). The null hypothesis (H0: $u=0$) here indicates that there is no difference between V and the rank sum of the negative elements ($W$), i.e. $V=W=\frac{1}{2} \sum_{n=1}^{N} n$ ($N$ is the sample size). The mean ($u_i$) and standard deviation ($\sigma_i$) of V are calculated by the following Equations:

$$u_i = \frac{1}{2} \sum_{n=1}^{N} n = \frac{N(N+1)}{4},$$  

(21)

$$\sigma_i = \sqrt{\frac{N(N+1)(2N+1)}{24}}.$$  

(22)

The z-value is computed by the Equation below:

$$z = \frac{V - u_i}{\sigma_i}.$$  

(23)

Compared the z-value with the critical values under the condition of 5 percent significance level and 37, 38-1, degrees of freedom, if the z-value falls in the critical region the null hypothesis will be rejected, otherwise we accept it. The p-value is available from the table of Rank-Sum Critical Values, and its precise value can be calculated by the same method as that in t-test (Equation (19)). In this study, Stata-8 will calculate the corresponding p-value of each z-value. We will compare it with 5 percent significance degree: if $p \leq 5\%$, we will accept the null hypothesis; Otherwise, if $p < 5\%$, the null hypothesis will be rejected.

According to the hypothesis testing assisted by t-test, sign test and Wilcoxon signed rank test, the mean of the cumulative abnormal returns will be justified to be acceptable or not as well as at what significance level. These three methods are all used for testing small samples. However, the nonparametric statistics sign test and Wilcoxon signed rank test relax the normality assumption of the t-test. Moreover, the Wilcoxon signed rank test take the magnitude of the signs into account and thereby can further confirm the result of sign test.

### 3 Empirical analysis of US bank M&A

#### 3.1 SAMPLE MEAN

Through the operations based on the Equations (1-7) and assisted by Microsoft Excel, the cumulative abnormal returns and combined cumulative abnormal returns are computed from the collected data about the sampled 38 transactions. The results of the cumulative abnormal returns of the bidders (CAR(b)) and targets (CAR(t)) and their CCARs are summarized and illustrated by the following Figure 1.

**FIGURE 1** CAR Distributions
Figure 1 shows that most of the sampled mergers provide positive cumulative abnormal returns to the targets. There are two outstanding exceptions that characterized by negative target returns but positive bidder returns. One is No. 13—NationsBank (CAR: 6.34%) acquires Boatmen’s Bancshares (CAR: -8.44%); the other is No. 24—Wells Fargo & Company (CAR: 12.04%) acquires First Security Corporation (CAR: -43.62%). However, their combined cumulative returns maintain positive.

The average cumulative abnormal returns are listed in the following Table 2.

<table>
<thead>
<tr>
<th>CAR(b)</th>
<th>CAR(t)</th>
<th>CCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.99%</td>
<td>15.07%</td>
<td>2.57%</td>
</tr>
</tbody>
</table>

Therefore, within the chosen 38-transaction sample, the observed value effect of the US bank mergers on the bidders is negative (-0.99%) but positive on targets (15.07%) and combined banks (2.57%). Using Stata-8, the statistical significance of the above results can be tested by t-test, sign test and Wilcoxon signed rank test respectively.

3.2 RESULTS OF SIGNIFICANCE TESTING

The process of Stata-8 based significance testing is recorded and analyzed in this part. The following analysis interprets the result of the significance testing.

3.2.1 t-test

As explained in Methodology, the null hypothesis of t-test is that the population mean of the cumulative abnormal returns is zero (H0: \( \mu = 0 \)). The result of t-test is demonstrated by the Table 3 below.

<table>
<thead>
<tr>
<th>t</th>
<th>p</th>
<th>Accept H0?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CARb</td>
<td>-0.863</td>
<td>0.394</td>
</tr>
<tr>
<td>CARt</td>
<td>5.181</td>
<td>0.000</td>
</tr>
<tr>
<td>CCAR</td>
<td>2.334</td>
<td>0.025</td>
</tr>
</tbody>
</table>

On the bidder side, its t-value (-0.863) corresponds to a p-value of 39.4%, which is much bigger than the selected degree of significance, 5%. Therefore, this t-value is located in the acceptance region that means the null hypothesis cannot be rejected. Under this circumstance, the negative sample mean is insignificant for the population. Based on the result of t-test, it is concluded that the value effect of the US bank mergers on bidders is negative, but insignificantly.

In contrast to the situation of the bidder, both the t-statistics of target (5.181) and combined bank (2.334) fall in the critical region, as their p-values, 0.00% and 2.5% respectively, are smaller than 5%. The H0 for the target and combined firm are therefore rejected. As the conclusion, the value effects of the US bank mergers on target and combined firm are significantly positive.

3.2.2 Sign test

There are 17 bidders with observed positive cumulative abnormal returns whilst 21 targets with negative ones. The null hypothesis here is \( p=q=0.5 \), i.e. the probabilities of observing a positive and a negative cumulative abnormal return are equivalent. According to the results of sign test shown in Table 4, \( p(x \geq 21) = 62.7\% \) which is much larger than 5%, which means that the null hypothesis should not be rejected. Consists with that of t-test, the corresponding conclusion is that the value effect on bidder is insignificant negative.

<table>
<thead>
<tr>
<th>CARb</th>
<th>CARt</th>
<th>CCAR</th>
<th>p</th>
<th>Accept H0?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.627</td>
<td>0.000</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As for targets, the number of observed + signs is 31 while that of - signs is 7. \( p(x \geq 31) = 0.0\% \) which is smaller than 5% and H0 therefore can be rejected even at 1% level. Similarly, there are 26 pluses and 12 minuses observed within combined banks. In this case, \( p(x \leq 26) = 3.4\% \) that is also smaller than 5%. The H0 is therefore can be rejected at 5% significance level. Based on the couple of rejections of H0, the sign test obtains the same results as t-test: the positive value effects on targets and combine firms are both significant.

3.2.3 Wilcoxon signed rank test

The result of Wilcoxon test is demonstrated by the Table 5 below.

<table>
<thead>
<tr>
<th>CARb</th>
<th>CARt</th>
<th>CCAR</th>
<th>z</th>
<th>p</th>
<th>Accept H0?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.812</td>
<td>4.343</td>
<td>2.110</td>
<td>0.417</td>
<td>0.035</td>
<td></td>
</tr>
</tbody>
</table>

According to the result attained by Stata-8, the z-values of the CAR’s rank sums of the bidder, target and combined firm are -0.812, 4.343 and 2.110 respectively. Their corresponding p-values are 41.7%, 0.0% and 3.5%. Therefore, the z-value of the bidder is located in acceptance region while that of the other two fall in critical region. It indicates that the null hypothesis, H0: \( V = W = \frac{1}{2} n(n+1) \), should be accepted for the bidder while rejected for both the target and the combined firms. As the conclusion, the same as that of t-test and sign test, the negative CAR of bidders is insignificant while the positive CAR of targets and combined firms are both significant.

4 Research finding

The significance testing through t-test, sign test and Wilcoxon signed rank test thus draw the same conclusion, that is, over a 36-day event window the US bank mergers occurring between 1994 and 2003 create insignificantly negative gains for bidders, and significant positive value for target banks and their combinations.

This conclusion firstly supports the previous studies that find positive gains for targets. Such studies include
those conducted by Becher (2000), Houston and Ryngaert (1994), Zhang (1995) and other six studies quoted by Becher (2000) [14], which calculate 20.48% of target bank gains on average. However, although a negative bidder gain is observed (−0.99%) in the 38-transaction sample, it is insignificantly different from zero. This is different from most of the previous studies. Two-third of studies conducted before 1994, found significant negative gains of acquiring banks. Through analysing the bank mergers in the 1990s, Becher (2000) calculate significantly positive bidder returns over a 36-day event window that is similar as what has been used by this study, while significant negative bidder returns are found over an 11-day event window. As for the combined firms, this study documents a significant positive return which supports the conclusion of Becher (2000). Nevertheless, more previous studies suggest the insignificant effect of bank mergers on combined firms due to the offset of negative bidder returns against the positive targets gains. Such studies include those conducted by Hannan and Wolken (1989) and Houston and Ryngaert (1994) [15].

According to Becher’s (2000) criteria of clarifying merger motivations, the results of the event study consist of the characteristics of mixed hypotheses of hubris and synergy—positive target returns, negative bidder returns and positive combined returns. However, the negative bidder return is insignificant, therefore, its attribute is much closer to synergy hypothesis, which is characterised by non-negative bidder returns compared to the mixed hypothesis which implies negative bidder returns. In this case, the US bank mergers occurring between 1994 and 2003 are incline to be motivated by synergy, rather than hubris.

5 Conclusions

References


Author

Guo-yi Chen, born on January 20, 1982, Wanzhou district, Chongqing Province, P.R. China.

Current position, grades: the lecturer of business administration, Chongqing three gorges University, Wanzhou district, Chongqing city, China.

Scientific interest: business management and regional economy.

Publications: 8 papers.

Experience: teaching experience of 5 years, 6 scientific research projects.