Estimating the Physical Probability for Successful Stock Swap Mergers: An Application of MCMC Methods*

By Giuseppe Corvasce†

Abstract

The paper proposes a Bayesian logistic regression, based on MCMC methods with a Gibbs sampling algorithm, with the aim to estimate the probability of a deal completion, for U.S. stock swap mergers and acquisitions (M&As). The estimated unknown parameters rely on several informative priors, able to consider the mean and the dispersion of the prior distributions. The Relative Operating Characteristics (**ROC**) curves compare in sample and pseudo out-of-sample results, among maximum likelihood and Bayesian estimates, based on 50% and 70% of the analyzed M&As. The empirical analysis also reports the diagnostic tests and the metrics of accuracy for evaluating the equality of the mean and variance between chains, for estimating the unknown parameters, across specifications of the prior distributions.

*The author would like to thank several participants of the conferences and seminars organized by GARP New York and Chicago Chapters, New York University, the University of Alberta, The Wharton School at the University of Pennsylvania, Luxembourg School of Finance, the Allied Social Science Association, The Paris EUROPLACE International Financial Forum, the Office of Financial Research (OFR), the Center for Financial Statistics and Risk Management at Rutgers University, its advisors and the staff members. The author also thanks several consultants, advisors, hedge/private equity and family funds as well as banks located in Chicago, New York, Washington, Philadelphia, London and Hong Kong areas for comments received.

†Rutgers University - The State University of New Jersey; Address: 57 U.S. Highway, 1 New Brunswick, NJ 08901-8554; Society for Financial Studies; Email: giuseppecorvasce@gmail.com

1. Introduction

The announcement of a merger process generally causes the target stock's price to rise, usually trading below the offer price, thereby, producing an operation able to create a risky opportunity for speculating, based on the deal risk. In merger attempts involving stock payments, risk arbitrageurs generally perform a trading strategy, offsetting a long position in the target firm's stock with a short position in the acquiring firm's stock. These trading positions are designed to capture the difference between the target stock's market price and the offered price that depends on the exchange ratio and the bidder stock's market price. This difference, defined as the "risk arbitrage" spread, represents the expected gain for arbitrageurs if a deal succeeds and plays an important role for the "risk arbitrage" process. If the merger is successfully consummated, risk arbitrageurs pocket this spread. However, when a merger fails, the spread will generally widen instead of converging to zero, causing risk arbitrageurs to incur huge losses. The "risk arbitrage" spread also depends on the probability of a deal completion that represents a crucial component able to characterize the expected gain pocket by speculators of a merger transaction.

The paper proposes a Bayesian logistic regression, based on MCMC methods with a Gibbs sampling algorithm, with the aim to estimate the probability of a deal completion, for a sample of 1090 U.S. stock swap mergers and acquisitions (M&As). The estimated unknown parameters, that allow the determination of the physical probability of a deal process, rely on several informative prior distributions (priors), able to incorporate the information regarding the mean and the dispersion of the priors. The Relative Operating Characteristics (**ROC**) curves compare in sample and pseudo out-of-sample results, among maximum likelihood and Bayesian estimates, based on 50% and 70% of the analyzed M&As.

In line with the literature (Gelman and Rubin 1992, Geweke 1992, Brooks and Gelman 1997) and the remarks pointed out by Cowles and Carlin (1996), the empirical analysis also reports some diagnostic tests, with the aim to evaluate the equality of the mean and variance between chains involved with the Bayesian estimation and across specifications of the priors and proposes the Monte Carlo error, the Root Mean Square Divergence (RMSD) and the Mean Absolute Percentage Divergence (MAPD), as metrics for evaluating the accuracy of the estimated unknown parameters.

The rest of the paper is organized as follows: section 2 provides an overview of the literature; section 3 provides a summary of the data and some descriptive statistics. Section 4 derives the model and the estimation procedure. Section 5 is devoted to the description of the diagnostic tests, with the aim to evaluate the convergence of the chains. Section 6 discusses the empirical results and the convergence of the chains. Section 7 provides the concluding remarks.

2. Overview of the literature

The literature on the topic of merger arbitrage proposes several frameworks for estimating the probability of a deal completion. Walkling (1985) uses a logit probability model with the aim to predict the success of a tender offer. Samuelson and Rosenthal (1986) show how stock price variations for the target company are informative for estimating the success probability of a tender offer that monotonically improves over time. Brown and Raymond (1986) estimate the physical probability of a deal, assuming the hypothesis of the convergence, between the target stock price and the bid offer, in case of resolution, using a sample of 35 failed mergers, for the period from January 1980 to December 1984.

De Bodt et al. (2014) propose an empirical evidence for testing the hubris hypothesis (Roll 1986). The first test involves the first order condition (**FOC**) of a bidder expected profit maximization, where, the probability of a deal completion is related to the bid premium, while, profits conditional on the completion of a transaction are decreasing on it; whereas, the second relation is between the probability of success and bidder returns, with a probability for a deal completion that is computed with a probit framework or roughly estimated as the difference between the bid premium and the observed price reaction, around the announcement.

Betton et al. (2014) propose a theoretical framework able to justify and empirically test the relation between the informativeness of a takeover signal, the target run-ups and the probability of a deal completion. The framework is tested on a sample of 6000 initial takeover bids for U.S. public targets from 1980 to 2008 and recognizes the role of the takeover probability and the conditional deal synergies, for estimating the run-up for the target stock prices (Schwert 1996).

Subramanian (2004) develops a jump diffusion model with the aim to reduce the risk for a merger arbitrage portfolio. The author infers the implied probability that the market assigns to the resolution of a deal, from the implied parameters of the proposed framework. Bester, Martinez and Rosu (2007) propose a framework able to compute the probability of a deal completion, the fallback price and estimate the synergies from the option prices.

Another part of the literature studies the relation between risk and return as well as the effects of the stock market and the business conditions on risk arbitrage activities are discussed in several academic studies, where, the Capital Asset Pricing Model (Sharpe 1964, Lintner 1965 and Mossin 1966) and the three factors Fama-French model (Fama and French 1993) have the aim to study the profitability of the merger processes.

Larker and Lys (1987) suggest that arbitrageurs are passive, with a superior ability for predicting the completion of a deal. According to this study, the average excess return is estimated around 3.75%, during a merger period. Theoretical models, proposed by Cornelli and Li (2002) and Gomes (2001), suggest the active role of arbitrageurs. Cornelli and Li (2002) document the role of arbitrageurs able to affect the value of the target shares and the procedure that allows to determine the number of arbitrageurs and the paid price, pointing out several empirical implications concerned about the trading volume and the takeover premium.

Dukes et al. (1992) examine a sample of 761 cash tender offers that realized an abnormal return of 171%. Jindra and Walkling (1999) point out several aspects regarding the profitability of risk arbitrage. According to this paper, an arbitrageur earns an annualized return greater than 100%. Further, the authors document a positive relation with the duration of the offer, the size of the bid premium and a negative relation with the target size and the revision of the offer. Karolyi and Shannon (1999) find an annualized return in excess of the TSE 300 for about 26%.

Baker and Savasoglu (2000, 2002) point out that merger arbitrage strategies do not earn a return on a continual basis; as such, the Capital Asset Pricing Model (CAPM) and the Fama-French (F&F) three factors model are not adequate for computing the annualized returns. Mitchell and Pulvino (2001) and Mitchell et al. (2002) propose a non linear pricing model able to study the relation between merger arbitrage portfolio returns and market returns.

The authors construct a setting able to consider the transaction costs and the real-world constraints, showing an annualized excess return for about 4%. The performance reflects a risk premium for arbitrageurs that provide liquidity during a merger process. During negative market periods, merger arbitrage portfolios suffer huge losses with a beta of roughly 0.50%; whereas, in flat and appreciating periods, the merger arbitrage strategy yields a positive return with a zero beta.

3. Data and descriptive statistics

The data-set relies on a sample of 1090 U.S. stock swap mergers and acquisitions (M&As), where, both the target and the bidder are publicly U.S. traded companies. The list of M&As is from Securities Data Company (SDC) Platinum that provides information on bond and equity new issues, financial securities, syndicated loans, project finance as well as mergers and acquisitions (M&As). The empirical analysis relies on the following screening criteria:

1. The announcement date and the resolution date for a merger or an acquisition are between January 1st, 1992 and December 31st, 2008;
2. The percentage of shares acquired by the bidder is greater than 50%;
3. The value of the transaction is greater than $ 1 million USD;

4. The accounting information and the stock prices for the target and the bidder are available on the joint database CRSP/COMPUSTAT;

5. The attitude of a merger or an acquisition can be friendly or hostile;

6. The bidder and the target firms do not belong to the same parent groups and are listed on NYSE, AMEX and NASDAQ;

7. SDC Platinum provides the values for the exchange ratio, concerned about each stock swap merger transaction.

The Dow Jones News Services, The Barron's and The Wall Street Journal allow to validate some critical information, concerned about the announcement date, the resolution date, the exchange ratio as well as the date for the revision of the offer, from a new bidder (*multiple offer bids*).

[Please Insert Table 1 around here]

Table 1 provides a summary of Mergers and Acquisitions (M&As) and some descriptive statistics regarding the number of announced deals, the average and the median market values for the target and the bidder, the ratio between the average market value for the target and the average market value for the bidder as well as the average amount of the transaction value. The number of announced deals change over the time, from 1992 to 2008, achieving the maximum number around the years 1997-2000, when 499 M&As were announced.

For each period, the average market value for the bidder is greater than the average market value for the target. From a modest average transaction value of $ 196.74 million USD (in 1992), the average transaction value of M&As respectively marched steadily upward to $ 4060.85 million USD (in 1998) and $ 3926.39 million USD (in 2000). Most of the 1990s deals were strategically negotiated and a major part were settled via stock, with the largest transactions in the period from 1998 to 2000. The worldwide volume of M&As reached a level above $ 3.3 trillion USD in 2000, when, the year started with the announcement of the record, setting a merger of $ 165 billion USD for Time Warner and AOL.

After a five years of telecommunication, media and technology (TMT) mergers, there was a general decline of the stocks in the TMT sector, followed by the earnings and financial problems that lead to a decrease in volume of the TMT mergers. The era of the mega deals ended with great scandals, like Enron, where, the shareholders filed a $ 40 billion USD lawsuit and the company reported a dramatic reduction of the stock price, from $ 90.75 USD (in the middle of 2000) to $ 1 USD (by the end of November 2001).

During the period from January 2002 to December 2006, the average transaction value for the U.S. stock swap mergers experienced a reduction, with respect to the previous wave of M&As, due to an increase of the U.S. cross-borders transactions, the rise in commodity prices, the availability of low interest financing and the role of private equity funds, linked to an increase of buyouts.

4. The methodology

This section is devoted to the discussion of the empirical framework able to provide a specification, with the aim to estimate the probability of a deal completion. It assumes that a stock swap merger is announced at a certain time $\tau = 0$ and that each deal i is expected to be completed, at the horizon time T. The framework considers the latent variable MT_i that represents the merger time, measured from time $\tau = 0$, for each transaction i.

[Please insert Appendix A around here]

The difference between the resolution and the failure of a merger depends upon the covariates (FINANCIAL DUMMY, DIVERS, FRIENDLY, RET_TARGET_PRE, RET_ACQUIRER_PRE, RELSIZE, LEVE_FIN, LEVE_NO_FIN) able to provide the information regarding the bidder and the target as well as the merger process. As such, the physical probability p of a deal completion i, computed at time τ, can be defined in the following way:

$$p_{i,\tau} = P\left[MT_i < T \mid MT_i > \tau\right]. \tag{1}$$

Appendix A describes the list of the independent covariates considered in a binary choice model able to explain the dependent variable *SUCCESS,* for each deal i, that takes a value equals to 1 if the merger time is below the horizon time T and the value 0 otherwise. As such:

$$SUCCESS_i = \begin{cases} 1 \qquad if \qquad MT_i < T \\ 0 \qquad\qquad otherwise \end{cases}. \tag{2}$$

For simplicity of the notation, the framework denotes with y, the realization of a merger process at time T. As such, the definition of the likelihood function depends on the probability mass function that is based on the number of failed and completed merger processes. In general, the likelihood contributions for the deal i that is completed, i.e. $y_i = 1$, and the one that is failed, i.e. $y_i = 0$, are respectively defined in the following way: $(P\{y_i = 1|\,\mathbf{x}_{i,\tau},\, \beta_\tau\})$ and $(P\{y_i = 0|\,\mathbf{x}_{i,\tau},\, \beta_\tau\})$. The quantities are a function of the vector of covariates $(\mathbf{x})$ and the vector of unknown parameters β, observed at time τ.

Therefore, the conditional likelihood function $L(y_i\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau)$, for the entire sample of merger processes (N), can be defined in terms of the probability mass function and computed in the following way:

$$L(y_i\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau) = \prod_{i=1}^{N} P\{y_i = 1\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\}^{y_i} \cdot P\{y_i = 0\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\}^{1-y_i}\,, \tag{3}$$

where, the quantities $P\{y_i = 1\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\}$ and $P\{y_i = 0\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\}$ are respectively defined as follows:

$$P\{y_i = 1\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\} = \left(\frac{exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}{1 + exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}\right) \tag{4}$$

whereas,

$$P\{y_i = 0\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau\} = \left(1 - \frac{exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}{1 + exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}\right). \tag{5}$$

For simplicity of the specification, the methodology discusses the case in which each unknown parameter β_j from $j = 0, ..., p$, computed at time τ, where, p is the number of the covariates, has a normal prior distribution with a mean equals to μ_j and a variance equals to σ_j^2 as well as random effects for the error components. Therefore,

$$\beta_{\mathrm{j},\tau} \sim N\left(\mu_{j,\tau}\,;\, \sigma_{j,\tau}^2\right). \tag{6}$$

The application of Bayes' theorem implies that the derivation of the posterior distribution is a function of the prior distribution for each unknown parameter and the conditional distribution for each merger transaction. As such, the corresponding likelihood function associated to the conjugate posterior distribution $L\left(\beta_\tau\,|\,y_i,\, \mathbf{x}_{i,\tau}\right)$ is equal to the likelihood function for the informative prior distribution over all parameters, times the conditional likelihood function $L(y_i\,|\,\mathbf{x}_{i,\tau};\, \beta_\tau)$.

As such, the associated likelihood is derived in the following way:

$$L\left(\beta_\tau\,|\,y_i,\, \mathbf{x}_{i,\tau}\right) = \prod_{i=1}^{N}\left(\frac{exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}{1 + exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}\right)^{y_i} \cdot \left(1 - \frac{exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}{1 + exp^{\left(\beta_\tau' \cdot \mathbf{x}_{i,\tau}\right)}}\right)^{1-y_i} \cdot \left(\prod_{j=0}^{p} \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma_{j,\tau}} \cdot exp\left\{-\frac{1}{2}\left(\frac{\beta_{j,\tau} - \mu_{j,\tau}}{\sigma_{j,\tau}}\right)^2\right\}\right). \tag{7}$$

The estimation procedure for determining the unknown parameters implies the derivation of the log likelihood and the minimization of it. The unknown parameters are based on a MCMC methodology with a Gibbs sampling algorithm (Gelman and Gelman 1984, Gelfand et al. 1990). It allows to generate posterior distributions, starting from initial values that can be randomly determined or derived with some expectationmaximization algorithms and sampled from prior conditional distributions. The sampling procedure allows to approximate the joint distribution for all covariates and the expected value of any covariate can be approximated by averaging the values, over all simulated samples.

5. Diagnostic tests for evaluating the convergence of the chains

The section is devoted to the description of the diagnostic tests able to evaluate the convergence in terms of bias and variance of the chains, with the aim to depict the quality of the estimated unknown parameters. The tests proposed in the literature (Geweke 1992; Gelman and Rubin 1992a, 1992b; Liu, Liu and Rubin 1992; Roberts 1992, 1994; Garren and Smith 1993; Johnson 1994) and the concluding remarks (Cowles and Carlin 1996) call for caution since a detailed analysis, with the aim to evaluate the speed and the time for the convergence of the chains, based on reparameterization procedures (Hills and Smith 1992, Gelfand et al. 1995a,b), auxiliary variables (Besag and Green 1993, Swendsen and Wang 1987), resampling and adaptive switching of the transition kernel (Gelfand and Sahu 1994) and multi-chain annealing or "tempering" (Geyer and Thompson 1995) might definitely improve the diagnostic of the convergence between chains involved in the estimation procedure.

The specification of the methodology, the Gibbs-sampling algorithm, the *burn-in* procedure and the number of iterates are also crucial points able to provide a discussion around the Monte Carlo (MC) error regarding the mean, median and variance for the estimates of the unknown parameters. The limitations of the diagnostic tests for Gibbs samplers and other MCMC algorithms (Cowles and Carlin 1996) might consider further tests for the equality of the mean (t-test, Anova-F test, Satterwhaite-Welch test and Welch test) and variance (Anova-F test, Siegel-Tukey test, Bartlett test, Brown-Forsythe test), between chains involved[1], with the aim to estimate the unknown parameters[2]. The tests require the following steps:

1. Running m multiple parallel chains, started from dispersed initial values, relying on the Gibbs sampler algorithm;

2. Dividing the number of iterates in quartiles, with the aim to generate m posterior distributions for the multiple parallel chains, with a mean and a variance;

[1] Further tests, such as Wilcoxon-Mann-Whitney (1947), Median Chi-Square (Conover 1999), Kruskal-Wallis (1957), Van der Waerden (1952), evaluate the equality of the median between chains and consider the outliers that might jeopardize the stability and decrease the speed for the convergence of the chains.

[2] The tests for evaluating the convergence among the chains can be developed for deciles and percentiles, with the aim to study the evolution of the probability values over the subgroups, in which the number of the iterates are subdivided.

3. Performing the tests for the equality of the *mean* among m multiple parallel chains, for each quartile, with the aim to assess the distance of the estimates from the true values, concerned about the quantities of interest obtained on a particular iteration, under the target distribution and evaluate the dependence of the estimated unknown parameters;

4. Performing the tests for the equality of the *variance* among m multiple parallel chains, for each quartile, in order to assess the variance and the convergence rate, in terms of MC error, for the estimated unknown parameters and so assess the quality of the estimates;

5. Valuating the statistical significance of the tests, for each quartile, concerned about the number of iterates;

6. Providing conclusions, with the aim to evaluate the convergence for m multiple parallel chains, across the quartiles, considering that:

   (a) the first quartile of the iterates takes into account auto-correlated and cross-correlated m multiple parallel chains that might decrease the speed, the time to the convergence for the chains and the probability value of the statistical tests, implying the rejection of the null hypothesis;

   (b) the statistical significance of the tests does not necessarily imply that the first and the second moments of the distributions are equal;

   (c) the statistical significance of the tests might depend on the number of the subgroups for dividing the iterates, generated with the Gibbs sampling algorithm;

   (d) the higher moments for the distributions of the chains might create bias conclusions, around the diagnostic tests for studying the convergence.

Following Zellner and Min (1995), the proposed tests allow to assess not only whether the Gibbs sampler converges in distribution, but if it also converges to the correct posterior distribution. Further, the tests allow to graphically evaluate and test the monotonic convergence of the estimated unknown parameters. The statistical procedure relies on a Bayesian data analysis package based on S-Plus and R routines called CODA (Best et al. 1995), accompanied with the BUGS software and developed for Bayesian analysis Using Gibbs Sampling (Spiegelhalter, Thomas and Best 1994, 1995; Thomas, Spiegelhalther and Best 1992, 2009). It allowed the origination of the software OpenBUGS (Thomas 2005) and its updates. These software for

Bayesian data analysis are freely available from the MRC Biostatistic Unit, at the University of Cambridge and Imperial College.

6. Empirical Results

The section is devoted to the discussion of the empirical results that rely on the methodology developed in section 4. The framework accommodates several specifications able to consider prior distributions (or simply called the **priors**) for the unknown parameters that would express the beliefs of the investors about these quantities, before some empirical evidences are considered. For example, the **priors** could be the probability distributions able to represent the density functions of the unknown parameters based on similar experiments that, relying on initial values, allow the definition of the posterior distributions for the parameters.

For the purpose of the analysis, the empirical results rely on normal as well as truncated normal distributions and the level of the informativeness for the priors is also based on the values for the first and the second moments of the distributions. These two quantities define the average value and the dispersion that represents the error for the prior distributions. The initial values allow to set the Gibbs sampling algorithm and rely on the maximum likelihood estimates, based on the Berndt-Hall-Hall-Hausman (B-H-H-H) algorithm (Berndt et al. 1974), that is similar in spirit to the Gauss-Newton algorithm.

In particular, the empirical results discuss nine sets of alternative (informative) priors, divided in two groups. The first group of priors respectively considers normal distributions centered around 0 and levels of dispersion that are respectively equal to 1.0E-3 (**Bayesian with Priors 1**) and 1.0E-6 (**Bayesian with Priors 2**), for the prior distributions concerned about the covariates DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN as well as centered around the value of 0.5 for the regressors DIVERS and FRIENDLY and -0.5 for the regressors RELSIZE, LEVE_FIN and LEVE_NO_FIN (**Bayesian with Priors 3**).

The set of priors n. 4 (**Bayesian with Priors 4**) considers normal distributions that are centered around the value of 1, for the covariates DIVERS and FRIENDLY and the value of -1, for the covariates RELSIZE, LEVE_FIN and LEVE_NO_FIN, with levels of dispersion equal to 1.0E-3. The set of priors n. 5 (**Bayesian with Priors 5**) considers normal distributions with levels of dispersion equal to 1.0E-1, for the prior distributions related to the covariates DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN centered around the level of 0.5 (**Bayesian with Priors 5**) and the level of 1 (**Bayesian with Priors 6**) as well as around -0.5 and -1 for the distributions of the covariates RELSIZE, LEVE_FIN and LEVE_NO_FIN.

The second group of priors considers the truncation of the prior normal distributions that enriches the level of the informativeness. In particular, the empirical results discuss the estimates of the methodology, based on the prior distributions, with a level of dispersion equals to 1.0E-1 and respectively centered around 0 and truncated on the **left side** for the covariates DIVERS and FRIENDLY and on the **right side** for the covariates RELSIZE, LEVE_FIN and LEVE_NO_FIN (**Bayesian with priors 7**). The sets of priors n. 8 (**Bayesian with priors 8**) and n. 9 (**Bayesian with priors 9**) allow the estimation of the unknown parameters, based on the prior normal distributions centered around 0.5 and 1, truncated on the left side for the covariates DIVERS and FRIENDLY; whereas, the prior normal distributions for the covariates RELSIZE, LEVE_FIN and LEVE_NO_FIN are truncated on the right side and centered around -0.5 and -1.

The Bayesian estimates of the unknown parameters are based on the Gibbs sampling algorithm and able to generate two Markov Chains (Chain 1 and Chain 2), with n. 100000 sampled iterates. The descriptive statistics, regarding the first and the second moments of the posterior distributions, consider a "*burn-in*" procedure, able to discard 10% of the initial iterates. The "*burn-in*" procedure allows to ignore some numbers of Markov Chain samples at the beginning of the Gibbs sampling algorithm and then only considers the n-th Markov Chain of samples, when, averaging the values to compute the descriptive statistics.

The estimates of the unknown parameters also consider the ordered over-relaxation (Neal 1995) procedure, based on the heatbath algorithm, that allows to reduce the "*random walk*" behavior in the early part of the sampling processes, with the aim to consider the auto-correlation between Markov Chain samples, increasing the speed to the convergence for the Markov Chains and reducing the time for reaching it.

[Please Insert Table 2 around here]

The estimates of the unknown parameters, reported on Table 2, show a Monte Carlo error that is below 2.5%, across the sets of priors. This metrics of accuracy allows to compare the differences between the mean (median) of the Markov Chain sampled values, with the true posterior mean (median). In particular, there is a lower level of the precision and so a higher level of the standard deviation for the covariates RET_ACQUIRER_PRE and FRIENDLY, across the sets of priors. The unknown parameters, estimated via the maximum likelihood procedure (Table 3), for the covariates FINANCIAL DUMMY, FRIENDLY, RET_TARGET_PRE, LEVE_FIN, report a z-statistics that is significant at 1% level.

[Please Insert Table 3 around here]

The Relative Operating Characteristics (**ROC**) curves show in-sample results concerned about the percentage of correctly classified Mergers and Acquisitions (M&As), as a function of the probability space discretized from 0 to 1, across the sets of informative priors. In particular, Figure 1.1 reports the percentage of correctly classified *successful* M&As; whereas, Figure 1.2 shows the evolution for the correctly classified *withdrawn* M&As.

[Please Insert Figure 1 around here]

The set of priors n. 9 (**Bayesian with Priors 9**) allows a higher number of correctly classified successful M&As; whereas, the Maximum Likelihood Estimation (MLE) of the unknown parameters allows a higher percentage of correctly classified withdrawn M&As. The intersection of the **ROC** curves allows to determine the optimal numbers of withdrawn and successful M&As and the related level of probability (Figure 1), and so evaluate the type I and type II errors.

[Please Insert Figure 2 and Figure 3 around here]

The pseudo out-of-sample results, based on 50% and 70% of the training sample of M&As, are respectively reported on Figure 2 and Figure 3, that show the general superiority of the MLE technique for the classification of withdrawn M&As; whereas, for the classification of successful M&As, the Bayesian estimates of the unknown parameters allow to outperform the pseudo out-of-sample results, based on the MLE technique.

In particular, the superiority of the maximum likelihood estimates, for correctly classifying 50% of the sample concerned about the withdrawn M&As, is reported in Figure 2.2, with the cutoff of the probability that is below 0.84. A threshold of the probability, between 0.84 and 0.90, guarantees the out-performance of the Bayesian estimates with respect to the maximum likelihood estimates for correctly classifying the withdrawn M&As and the out-performance of the maximum likelihood estimates with respect to the Bayesian estimates, for correctly classifying the successful M&As. A level of the probability that is greater than 0.90 allows the out-performance of the maximum likelihood estimates, for correctly classifying the withdrawn M&As and the under-performance of the maximum likelihood estimates with respect to the Bayesian estimates, for correctly classifying the successful M&As.

6.1 Evaluating the convergence of the chains

The discussion for evaluating the convergence of the chains is elaborated in this section. It shows the reliability of the tests for the equality of the mean and variance, with the aim to measure the reliability of the Bayesian estimates, regarding the unknown parameters. For the purpose of the section, the empirical results discuss the tests for the Bayesian estimates, concerned about the sets of priors n. 1 and n. 7 and subdivide the number of the iterates in quartiles.

[Please Insert Table 4 around here]

Table 4 reports the t-test and the Anova F-test able to depict the coefficients and the levels of probability that allow to reject the null hypothesis, regarding the equality of the mean, between the generated Markov Chains. A high probability value implies that the null hypothesis is not rejected; whereas, a probability value smaller than 0.05 guarantees the rejection of the null hypothesis concerned about the equality of the mean, between the Chain 1 and the Chain 2 and a high statistical significance.

[Please Insert Appendix B around here]

The Bayesian estimates, generated with the Gibbs sampling algorithm, might improve with the increase of the sampled iterates. Appendix B reports some metrics of accuracy for the Bayesian estimates of the unknown parameters and based on the computation of the convergence rates, derived as a ratio between the MC error and the standard deviation of the posterior distributions, the Root Mean Square Divergence (RMSD) and the Mean Absolute Percentage Divergence (MAPD) ratios, computed with n. 1000000 sampled iterates. The convergence rates, for the Bayesian estimates of the unknown parameters, are below 0.01, implying a lower level of dispersion concerned about the estimated values. Although the increase of the sampled iterates, the difference of the estimated values, between the Bayesian approach and the maximum likelihood, is still evident.

The probability values, concerned about the first quartile, show how the t-test and the Anova F-test for evaluating the equality of the mean, between the Chain 1 and the Chain 2, computed for each unknown parameter, are in general not statistically significant, across the sets of priors n. 1 and n. 7, although, the mean of the sampled values turns out to be equal for the Chain 1 and the Chain 2. The results might change if the tests are performed for the first decile of the iterates, because of the higher order auto-regressive components concerned about the generated Markov Chains that might impact this decile and the reliability of the tests.

Further diagnostic tests, for evaluating the convergence of the chains, involve the reliability of the Satterthwaite-Welch test and the Welch test able to consider the quartile heterogeneity of the variances and derive some conclusions concerned about the speed to the convergence, among the chains involved with the Gibbs sampling estimation procedure.

[Please Insert Appendix C around here]

Appendix C reports the statistical values for the estimated unknown parameters. In particular, the first and the last quartiles, in which the tests are performed, report high probability values, implying the statistical non rejection of the null hypothesis, concerned about the equality of the mean. The speed to the convergence is evident for the second and the third quartiles, in which the number of the iterates for the Markov Chains are divided, where the performed tests are also statistically significant.

[Please Insert Table 5 around here]

In particular, the over-relaxation procedure guarantees an increase of the speed and so reduces the auto-regressive component for the Markov Chains. The coefficients of these tests drastically decrease, corroborating the statistical findings also based on the F-test, the Siegel-Tukey test, the Bartlett test and the Brown-Forsythe test. These tests are performed for the quartiles, in which the sampled iterates are divided and allow to evaluate the equality of the variance for each quartile, between the generated Markov Chains. The first and the last quartiles show high probability values that allow to not reject the null hypothesis regarding the equality of the variance.

7. Conclusion

The Gibbs sampling algorithm represents an alternative estimation procedure for evaluating and testing the physical probability of a deal completion for U.S. stock swap mergers and acquisitions (M&As), that represents a crucial component able to characterize the expected gain pocket by *"risk arbitrageurs"* of a merger transaction. The framework relies on a Bayesian logistic regression and compares the results derived via the Gibbs sampling algorithm with the ones generated via the MLE procedure, based on the algorithm developed by Berndt et al. (1974), where, the estimated unknown parameters rely on several sets of informative prior distributions (**priors**), able to consider the mean and the dispersion. The empirical results are corroborated with the Relative Operating Characteristics (**ROC**) curves that have the aim to compare in sample and pseudo out-of-sample results, among maximum likelihood and Bayesian estimates, based on 50% and 70% of the training sets, concerned about the analyzed M&As.

The concluding remarks show the superiority of the maximum likelihood estimates, for correctly classifying 50% of the sample, concerned about the *withdrawn* M&As, with a cutoff of the probability that is below 0.84; whereas, a threshold of the probability, between 0.84 and 0.90, guarantees the out-performance of the Bayesian estimates, with respect to the maximum likelihood estimates, for correctly classifying the *withdrawn* M&As and the out-performance of the maximum likelihood estimates with respect to the Bayesian estimates, for correctly classifying the *successful* M&As.

A level of the probability that is greater than 0.90 allows the superiority of the maximum likelihood estimates, for correctly classifying the *withdrawn* M&As and the under-performance of the maximum likelihood estimates with respect to the Bayesian estimates, for correctly classifying the *successful* M&As.

Following the remarks pointed out by Cowles and Carlin (1996), the tests for assessing the equality of the mean (t-test, Anova F-test, Satterwhaite-Welch test and Welch test) and variance (Anova F-test, Siegel-Tukey test, Bartlett test, Brown-Forsythe test), between chains involved with the Bayesian estimation procedure, allow to depict the quality of the estimated unknown parameters and provide some insights regarding the convergence to the posterior distributions.

Prof. Dr. Giuseppe CORVASCE is a professor and research scientist in financial economics. He holds a Ph.D. in Economics with Specialization in Finance at the Swiss Finance Institute, after his studies also undertaken at Università commerciale Luigi Bocconi, following a background in accounting. Prof. CORVASCE is or was a visiting research scholar and professor at the University of Calgary - Haskayne School of Business, University of Alberta, Luxembourg School of Finance - University of Luxembourg, Fordham University, Jacobs - Levy Equity Management Center for Quantitative Financial Research at The Wharton School of Business, Rutgers University - The State University of New Jersey and New York University, among many other institutions. He is also an advisor and consultant for several (international) financial and non-financial institutions, central and reserve banks, executive departments of several governments as well as The Officer of the Society for Financial Studies. Prof. CORVASCE is a member of many academic associations and acts as a reviewer and a referee for several peer-reviewed international academic journals and publishers. His research interests are: financial economics, systemic risk, financial intermediation, risk management and control, asset management and allocation, economics, international finance, volatility, financial regulation and law, financial stability, corporate governance and special situations.

Some References

Baker, P. M., and Savasoglu, S. (2002), "Limited Arbitrage in Mergers and Acquisitions", Journal of Financial Economics, 64(1), 91-116.

Bartlett, M. S. (1937), "Properties of sufficiency and statistical tests", Proceedings of the Royal Statistical Society Series A, 160, 268–282.

Besag, J. E., and Green, P. J. (1993), "Spatial statistics and Bayesian computation" (with Discussion), Journal of the Royal Statistical Society B, 55, 25–37.

Bester, A. C., Martinez, V. H., and Ioanid, R. (2007), "Option Prices and the Probability of Success of Cash Mergers", working paper.

Betton, S., Eckbo, B.E., Thompson, R., and Thorburn, K. (2015), "Merger Negotiations with Stock Market Feedback", Journal of Finance, 69, 1705-1745.

Brown, K. C., and Raymond, M. V. (1986), "Risk Arbitrage and the prediction of successful corporate takeovers", Financial Management, 15(3), 54-63.

Brown, M. B., and Forsythe, A. B. (1974), "Robust tests for the equality of variances", Journal of the American Statistical Association, 69, 364–367.

Cornelli, F., and Li, David D. (2002), "Risk Arbitrage in Takeovers", Review of Financial Studies, 15(3), 837-868.

Cowles, M. K., and Carlin, B. P. (1996), "Markov Chain Monte Carlo convergence diagnostics: A comparative review", Journal of the American Statistical Association, 91, 883–904.

De Bodt, E., Cousin, J. B., and Roll, R. (2014), "The Hubris Hypothesis: Empirical Evidence", working paper.

Efron, B., and Tibshirani, R. (1993), "An Introduction to the Bootstrap", Chapman & Hall, New York.

Fama, E., and French, K. (1993), "Common risk factors in the returns on stocks and bonds", Journal of Financial Economics, 33(1), 3-56.

Gelfand, A. E., and Smith, A. F. M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities", Journal of the American Statistical Association, 85, 398–409.

Gelfand, A. E., and Sahu, S. K. (1994), "On Markov Chain Monte Carlo Acceleration", Journal of Computational and Graphical Statistics, 3(3), 261-276.

Gelfand, A. E., Sahu, S. K., and Carlin, B. P. (1995a), "Efficient Parameterizations for Normal Linear Mixed Models", Biometrika, 82, 479-488.

Geyer, C. J., and Thompson, E. A. (1995), "Annealing Markov chain Monte Carlo with applications to ancestral inference", Journal of the American Statistical Association, 90, 909-920.

Green, P. J. (1995), "Markov Chain Monte Carlo computation and Bayesian model determination", Biometrika, 82, 711-732.

Hills, S. E., and Smith, A. F. M. (1992), "Parameterization Issues in Bayesian Inference" (with discussion), in Bayesian Statistics 4, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford, U.K.: Oxford University Press, 227–246.

Larcker, D. F., and Lys, T. (1987), "An empirical analysis of the incentives to engage in costly information acquisition", Journal of Financial Economics, 18(1), 111-126.

Liu, C., and Rubin, D. B. (1996), "Markov-normal analysis of iterative simulations before their convergence", Journal of Econometrics, 75(1), 69-78.

Liu, C., and Rubin, D. B. (2002), "Model based analysis to improve the performance of iterative simulations", Statistica Sinica, 12(3), 751-767.

Mitchell, M., and Pulvino, T. (2001), "Characteristics of Risk and Return in Risk Arbitrage", The Journal of Finance, 56(6), 2135-2175.

Mitchell, M., Pulvino, T., and Stafford, E. (2002), "Limited Arbitrage in Equity Markets", Journal of Finance, 57(2), 551-584.

Neal, M. R. (1995), "Suppressing Random Walks in Markov Chain Monte Carlo Using Ordered Overrelaxation", Technical Report n. 9508, Department of Statistics, University of Toronto.

Robert, C. P., and Casella, G. (2004), "Monte Carlo Statistical Methods" (second edition), Springer-Verlag.

Roll, R. (1986), "The Hubris Hypothesis of Corporate Takeovers", Journal of Business, 59(2), 197-216.

Samuelson, W., and Rosenthal, L. (1986), "Price movements as indicators of tender offer success", Journal of Finance, 41(2), 481-499.

Snedecor, G. W., and Cochran, W. G. (1989), "Statistical Methods", Eighth Edition, Iowa State University Press.

Spiegelhalter, D. J., Thomas, A., and Best, N. G. (1996), "Computation on Bayesian Graphical Models", Bayesian Statistics 5, ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford, U.K.: Oxford University Press, 5, 407-425.

Schwert, G. W. (1996), "Markup pricing in mergers and acquisitions", Journal of Financial Economics, 41(2), 153-192.

Shleifer, A., and Vishny, R. W. (1997), "The limits of arbitrage", Journal of Finance, 52(1), 35-55.

Subramanian, A. (2004), "Option Pricing on Stocks in Mergers and Acquisitions", Journal of Finance, 59(2), 795-831.

Swendsen, H. R., and Wang, J-S. (1987), "Non universal Critical Dynamics in Monte Carlo Simulations", Physics Review Letters, 58, 86-88.

Walkling, R. A. (1985), "Predicting Tender Offer Success: A Logistic Analysis", Journal of Financial and Quantitative Analysis, 20(4), 461-478.

Welch, B. L. (1951), "On the Comparison of Several Mean Values: An Alternative Approach", Biometrika, 38(3/4), 330–336.

Weston, J. F., Mitchell, M. L., and Mulherin, J. H. (2004), "Takeovers, restructuring and corporate governance", (4th, international edition), Pearson Education.

Wooldridge, J. M. (2002), "Econometric Analysis of Cross-Sectional Panel Data", MIT Press.

Zellner, A., and Min, C. K. (1995), "Gibbs Sampler Convergence Criteria", Journal of the American Statistical Association, 90(431), 921-927.

Table 1.

Summary of Mergers & Acquisitions (M&As) from 1992 to 2008

The dataset includes 1090 stock swap mergers and acquisitions (M&As), where both the target and the bidder are publicly traded firms in the U.S. stock exchanges. The list of M&As is from the **Securities Data Company**'s (SDC) U.S. Mergers and Acquisitions database. The criteria for screening the sample are the following: the initial announcement and the resolution date are between January 1st, 1992 and December 31st, 2008; the percentage of shares acquired by the bidder is greater than 50%; the deal value is greater than $ 1 million USD; the target and the bidder firms have data available on CRSP/Compustat database; the merger attitude can be friendly or hostile; the bidder and the target firms do not belong to the same group; **SDC Platinum** provides the exchange ratio for each stock swap merger. **The table shows**: (i) the number of announced deals for each year; (ii) the average and the median market value for the bidder, in $ million. USD (**AMKV**); (iii) the average and the median market value for the target company, in $ million. USD (**TMKV**); (iv) the ratio between the average market value for the target and the average market value for the bidder (**RELSIZE**); (v) the average amount of the transaction value (**TRANSACTION VALUE**), in $ million. USD.

YEAR	NUMBER OF ANNOUNCED DEALS	AMKV		TMKV		RELSIZE (TMKV/AMKV)	TRANSACTION VALUE (avg. $ Mil.)
		(avg. $ Mil.)	(median $ Mil.)	(avg. $ Mil.)	(median $ Mil.)		
1992	33	1710.50	663.41	137.59	76.78	0.08	196.74
1993	38	2522.80	565.66	392.74	76.21	0.16	404.98
1994	41	1386.78	838.69	337.78	79.07	0.24	379.64
1995	112	2440.35	653.62	670.15	124.01	0.27	852.47
1996	85	3841.56	1388.85	822.12	159.08	0.21	1042.40
1997	125	4099.71	1184.09	862.05	148.86	0.21	1145.06
1998	142	8992.65	967.01	3246.54	174.87	0.36	4060.85
1999	127	16855.99	1414.19	1415.99	133.86	0.08	2021.32
2000	105	16748.86	1208.22	2262.13	165.55	0.13	3926.39
2001	62	8292.15	815.10	683.94	88.82	0.08	795.01
2002	29	10171.58	720.28	2377.33	76.63	0.23	2721.31
2003	40	7520.87	630.19	1714.37	107.37	0.23	2247.82
2004	41	7448.83	486.86	2613.99	154.77	0.35	3255.09
2005	29	10649.74	709.34	2198.16	134.06	0.20	2681.73
2006	27	8320.27	1618.49	2166.50	313.44	0.26	2557.89
2007	32	10466.95	668.97	673.45	234.28	0.06	831.81
2008	22	19466.89	675.08	2166.02	126.61	0.11	1791.86
Total Average Value	**1090**	**8290.38**	**894.59**	**1455.34**	**139.66**	**0.19**	**1818.37**

Table 2.

The estimation results – Bayesian technique (MCMC with a Gibbs Sampler)

The table shows the estimation results for a logistic model, based on **MCMC** methods **with a Gibbs sampler**. The computations are performed with the software **OpenBUGS with the version 3.2.3** and rely on different specifications, based on (informative) priors for the parameters. (1) **Bayesian with Priors 1:** The prior distributions for the parameters are normals, centered around 0. There is a greater level of dispersion (1.0E-3) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (2) **Bayesian with Priors 2:** The prior distributions for the parameters are normals, centered around 0 and with a level of dispersion that is equal to 1.0E-6**.** (3) **Bayesian with Priors 3:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-3) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 0.5 for DIVERS and FRIENDLY and -0.5 for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (4) **Bayesian with Priors 4:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-3) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 1 for DIVERS and FRIENDLY and -1 for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (5) **Bayesian with Priors 5:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-1) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 0.5 for DIVERS and FRIENDLY and -0.5 for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (6) **Bayesian with Priors 6:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-1) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 1 for DIVERS and FRIENDLY and -1 for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (7) **Bayesian with Priors 7:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-1) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 0 and truncated on the **left side** for DIVERS and FRIENDLY and on the **right side** for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (8) **Bayesian with Priors 8:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-1) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 0.5 and truncated on the **left side** for DIVERS and FRIENDLY and centered around -0.5 and truncated on the **right side** for RELSIZE, LEVE_FIN and LEVE_NO_FIN**;** (9) **Bayesian with Priors 9:** The prior distributions for the parameters are normals. There is a greater level of dispersion (1.0E-1) for DIVERS, FRIENDLY, RELSIZE, LEVE_FIN and LEVE_NO_FIN and the distributions are centered around 1 and truncated on the **left side** for DIVERS and FRIENDLY as well as centered around -1 and truncated on the **right side** for RELSIZE, LEVE_FIN and LEVE_NO_FIN**. The table** reports the estimated mean (**mean**) and median (**median**) values of the parameters $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8$, for the following covariates: FINANCIAL DUMMY, DIVERS, FRIENDLY, RET_ACQUIROR_PRE, RET_TARGET_PRE, RELSIZE, LEVE_FIN, LEVE_NO_FIN; the standard deviations (**sd**) of the densities for each parameter; the Monte Carlo error (**MC error**) related to the parameters; the values of the densities for the parameters computed at the following levels: **5%**, **10%**, **90%** and **95%.**

Bayesian with Priors 1

Parameters	Covariates	Mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.1380**	0.6897	0.0245	-2.3340	-2.0400	**-1.1140**	-0.3025	-0.0623
β_1	**FINANCIAL DUMMY**	**2.4490**	0.5824	0.0058	1.5460	1.7380	**2.4240**	3.1990	3.4500
β_2	**DIVERS**	**0.4613**	0.2035	0.0016	0.1236	0.2014	**0.4624**	0.7215	0.7922
β_3	**FRIENDLY**	**2.6830**	0.6680	0.0237	1.6350	1.8720	**2.6580**	3.5500	3.8190
β_4	**RET_ACQUIRER_PRE**	**2.4760**	0.8186	0.0033	1.1430	1.4300	**2.4680**	3.5280	3.8160
β_5	**RET_TARGET_PRE**	**0.1053**	0.4562	0.0023	-0.6254	-0.4712	**0.0965**	0.6912	0.8696
β_6	**RELSIZE**	**-0.4583**	0.1319	0.0008	-0.6783	-0.6283	**-0.4552**	-0.2924	-0.2456
β_7	**LEVE_FIN**	**-0.0815**	0.0358	0.0004	-0.1423	-0.1276	**-0.0799**	-0.0380	-0.0260
β_8	**LEVE_NO_FIN**	**0.0482**	0.0367	0.0002	-0.0087	0.0028	**0.0462**	0.0966	0.1121

Bayesian with Priors 2

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.1400**	0.7076	0.0247	-2.3610	-2.0630	**-1.1150**	-0.2767	-0.0365
β_1	**FINANCIAL DUMMY**	**2.4510**	0.5876	0.0058	1.5380	1.7370	**2.4210**	3.2140	3.4610
β_2	**DIVERS**	**0.4616**	0.2031	0.0017	0.1241	0.2015	**0.4622**	0.7213	0.7924
β_3	**FRIENDLY**	**2.6850**	0.6848	0.0240	1.6170	1.8530	**2.6600**	3.5710	3.8560
β_4	**RET_ACQUIRER_PRE**	**2.4760**	0.8195	0.0034	1.1410	1.4300	**2.4640**	3.5350	3.8330
β_5	**RET_TARGET_PRE**	**0.1053**	0.4547	0.0024	-0.6262	-0.4690	**0.0985**	0.6887	0.8639
β_6	**RELSIZE**	**-0.4581**	0.1322	0.0009	-0.6783	-0.6285	**-0.4553**	-0.2909	-0.2436
β_7	**LEVE_FIN**	**-0.0817**	0.0361	0.0004	-0.1438	-0.1285	**-0.0798**	-0.0380	-0.0258
β_8	**LEVE_NO_FIN**	**0.0481**	0.0367	0.0002	-0.0087	0.0026	**0.0461**	0.0966	0.1118

Bayesian with Priors 3

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.1420**	0.7010	0.0246	-2.3560	-2.0520	**-1.1210**	-0.2870	-0.0483
β_1	**FINANCIAL DUMMY**	**2.4470**	0.5872	0.0055	1.5380	1.7330	**2.4190**	3.2050	3.4590
β_2	**DIVERS**	**0.4618**	0.2029	0.0017	0.1268	0.2027	**0.4623**	0.7210	0.7937
β_3	**FRIENDLY**	**2.6850**	0.6779	0.0238	1.6260	1.8650	**2.6620**	3.5600	3.8420
β_4	**RET_ACQUIRER_PRE**	**2.4740**	0.8201	0.0034	1.1340	1.4210	**2.4680**	3.5270	3.8160
β_5	**RET_TARGET_PRE**	**0.1067**	0.4557	0.0025	-0.6256	-0.4688	**0.0985**	0.6917	0.8664
β_6	**RELSIZE**	**-0.4580**	0.1322	0.0009	-0.6782	-0.6282	**-0.4552**	-0.2909	-0.2434
β_7	**LEVE_FIN**	**-0.0814**	0.0360	0.0003	-0.1432	-0.1283	**-0.0795**	-0.0378	-0.0257
β_8	**LEVE_NO_FIN**	**0.0482**	0.0367	0.0002	-0.0085	0.0030	**0.0462**	0.0966	0.1116

Bayesian with Priors 4

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.1420**	0.7010	0.0246	-2.3570	-2.0520	**-1.1210**	-0.2872	-0.0485
β_1	**FINANCIAL DUMMY**	**2.4470**	0.5872	0.0055	1.5380	1.7330	**2.4190**	3.2050	3.4590
β_2	**DIVERS**	**0.4619**	0.2029	0.0017	0.1268	0.2027	**0.4623**	0.7210	0.7937
β_3	**FRIENDLY**	**2.6860**	0.6779	0.0238	1.6260	1.8650	**2.6620**	3.5600	3.8430
β_4	**RET_ACQUIRER_PRE**	**2.4740**	0.8201	0.0034	1.1340	1.4210	**2.4680**	3.5280	3.8160
β_5	**RET_TARGET_PRE**	**0.1067**	0.4557	0.0025	-0.6256	-0.4688	**0.0985**	0.6917	0.8664
β_6	**RELSIZE**	**-0.4580**	0.1322	0.0009	-0.6782	-0.6282	**-0.4552**	-0.2909	-0.2434
β_7	**LEVE_FIN**	**-0.0814**	0.0360	0.0003	-0.1432	-0.1283	**-0.0795**	-0.0378	-0.0257
β_8	**LEVE_NO_FIN**	**0.0482**	0.0367	0.0002	-0.0085	0.0030	**0.0462**	0.0966	0.1116

Bayesian with Priors 5

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.0250**	0.6664	0.0217	-2.1670	-1.8960	**-0.9881**	-0.2159	0.0045
β_1	**FINANCIAL DUMMY**	**2.4490**	0.5720	0.0066	1.5580	1.7350	**2.4270**	3.1970	3.4240
β_2	**DIVERS**	**0.4601**	0.2024	0.0017	0.1260	0.2038	**0.4580**	0.7184	0.7942
β_3	**FRIENDLY**	**2.5680**	0.6430	0.0206	1.5660	1.7810	**2.5350**	3.4020	3.6650
β_4	**RET_ACQUIRER_PRE**	**2.4720**	0.8151	0.0035	1.1450	1.4370	**2.4700**	3.5210	3.8210
β_5	**RET_TARGET_PRE**	**0.1077**	0.4556	0.0027	-0.6260	-0.4626	**0.0988**	0.6919	0.8689
β_6	**RELSIZE**	**-0.4586**	0.1314	0.0009	-0.6753	-0.6263	**-0.4571**	-0.2923	-0.2452
β_7	**LEVE_FIN**	**-0.0816**	0.0352	0.0004	-0.1413	-0.1271	**-0.0801**	-0.0385	-0.0266
β_8	**LEVE_NO_FIN**	**0.0480**	0.0363	0.0002	-0.0082	0.0037	**0.0462**	0.0955	0.1110

Bayesian with Priors 6

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.0690**	0.7033	0.0229	-2.2600	-1.9730	**-1.0510**	-0.1846	0.0605
β_1	**FINANCIAL DUMMY**	**2.4430**	0.5790	0.0068	1.5210	1.7180	**2.4270**	3.1910	3.4230
β_2	**DIVERS**	**0.4616**	0.2032	0.0017	0.1268	0.2025	**0.4615**	0.7209	0.7971
β_3	**FRIENDLY**	**2.6140**	0.6790	0.0219	1.5200	1.7580	**2.6020**	3.4860	3.7700
β_4	**RET_ACQUIRER_PRE**	**2.4730**	0.8151	0.0033	1.1450	1.4300	**2.4720**	3.5170	3.8210
β_5	**RET_TARGET_PRE**	**0.1074**	0.4567	0.0027	-0.6296	-0.4687	**0.0966**	0.6895	0.8688
β_6	**RELSIZE**	**-0.4596**	0.1329	0.0009	-0.6799	-0.6293	**-0.4594**	-0.2907	-0.2445
β_7	**LEVE_FIN**	**-0.0812**	0.0355	0.0004	-0.1404	-0.1265	**-0.0801**	-0.0380	-0.0254
β_8	**LEVE_NO_FIN**	**0.0480**	0.0366	0.0002	-0.0086	0.0029	**0.0460**	0.0957	0.1114

Bayesian with Priors 7

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.0550**	0.6793	0.0191	-2.2030	-1.9120	**-1.0380**	-0.2069	0.0262
β_1	**FINANCIAL DUMMY**	**2.4440**	0.5795	0.0058	1.5280	1.7190	**2.4210**	3.2050	3.4350
β_2	**DIVERS**	**0.4579**	0.2057	0.0015	0.1212	0.1933	**0.4605**	0.7199	0.7956
β_3	**FRIENDLY**	**2.5980**	0.6548	0.0181	1.5590	1.7760	**2.5860**	3.4190	3.7110
β_4	**RET_ACQUIRER_PRE**	**2.4680**	0.8172	0.0033	1.1270	1.4280	**2.4690**	3.5180	3.8160
β_5	**RET_TARGET_PRE**	**0.1118**	0.4544	0.0024	-0.6172	-0.4626	**0.0988**	0.7074	0.8788
β_6	**RELSIZE**	**-0.4580**	0.1330	0.0006	-0.6795	-0.6309	**-0.4555**	-0.2911	-0.2445
β_7	**LEVE_FIN**	**-0.0811**	0.0356	0.0004	-0.1429	-0.1267	**-0.0798**	-0.0374	-0.0253
β_8	**LEVE_NO_FIN**	**0.0483**	0.0367	0.0002	-0.0086	0.0030	**0.0463**	0.0964	0.1112

Bayesian with Priors 8

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.0510**	0.6745	0.0186	-2.1950	-1.9270	**-1.0330**	-0.2115	0.0120
β_1	**FINANCIAL DUMMY**	**2.4350**	0.5776	0.0058	1.5200	1.7150	**2.4170**	3.1820	3.4240
β_2	**DIVERS**	**0.4597**	0.2038	0.0014	0.1219	0.1981	**0.4604**	0.7196	0.7969
β_3	**FRIENDLY**	**2.5960**	0.6514	0.0175	1.5590	1.7780	**2.5820**	3.4360	3.7010
β_4	**RET_ACQUIRER_PRE**	**2.4760**	0.8206	0.0034	1.1300	1.4280	**2.4750**	3.5270	3.8180
β_5	**RET_TARGET_PRE**	**0.1075**	0.4573	0.0022	-0.6301	-0.4757	**0.0983**	0.7049	0.8702
β_6	**RELSIZE**	**-0.4595**	0.1319	0.0006	-0.6813	-0.6298	**-0.4574**	-0.2918	-0.2453
β_7	**LEVE_FIN**	**-0.0807**	0.0354	0.0003	-0.1415	-0.1263	**-0.0790**	-0.0369	-0.0248
β_8	**LEVE_NO_FIN**	**0.0482**	0.0367	0.0002	-0.0080	0.0030	**0.0461**	0.0957	0.1111

Bayesian with Priors 9

Parameters	Covariates	mean	sd	MC error	5.00%	10.00%	Median	90.00%	95.00%
β_0		**-1.0960**	0.7132	0.0232	-2.3090	-2.0010	**-1.0660**	-0.2145	0.0369
β_1	**FINANCIAL DUMMY**	**2.4520**	0.5779	0.0063	1.5310	1.7320	**2.4320**	3.2020	3.4300
β_2	**DIVERS**	**0.4622**	0.2026	0.0016	0.1311	0.2016	**0.4612**	0.7225	0.7940
β_3	**FRIENDLY**	**2.6380**	0.6883	0.0220	1.5500	1.7930	**2.6160**	3.5260	3.8120
β_4	**RET_ACQUIRER_PRE**	**2.4700**	0.8190	0.0037	1.1260	1.4360	**2.4700**	3.5230	3.8260
β_5	**RET_TARGET_PRE**	**0.1080**	0.4558	0.0025	-0.6206	-0.4672	**0.1024**	0.6939	0.8714
β_6	**RELSIZE**	**-0.4597**	0.1324	0.0006	-0.6804	-0.6301	**-0.4565**	-0.2939	-0.2438
β_7	**LEVE_FIN**	**-0.0816**	0.0356	0.0004	-0.1417	-0.1275	**-0.0802**	-0.0381	-0.0257
β_8	**LEVE_NO_FIN**	**0.0484**	0.0366	0.0002	-0.0086	0.0029	**0.0466**	0.0961	0.1123

Table 3.

The estimation results – Maximum Likelihood technique

The table reports in sample results for a logistic regression. The dependent variable is the probability of resolution for a Merger or Acquisition. It reports the estimated values of the parameters $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8$, for the following covariates: FINANCIAL DUMMY, DIVERS, FRIENDLY, RET_ACQUIRER_PRE, RET_TARGET_PRE, RELSIZE, LEVE_FIN, LEVE_NO_FIN. The Table reports the values of the coefficients, the standard error, the z-statistic and the probability value (p-value). The sample contains 1090 U.S. Mergers and Acquisitions (M&As).

Parameters	**Covariates**	**Coefficient**	**Standard Error**	**z-statistic**	**p-value**
β_0		**-1.0626**	0.6835	-1.5546	0.1200
β_1	**FINANCIAL DUMMY**	**2.3284**	0.5301	4.3926	0.0000
β_2	**DIVERS**	**0.4604**	0.2034	2.2635	0.0236
β_3	**FRIENDLY**	**2.6080**	0.6603	3.9498	0.0001
β_4	**RET_ACQUIRER_PRE**	**2.4834**	0.8083	3.0722	0.0021
β_5	**RET_TARGET_PRE**	**0.0758**	0.4521	0.1676	0.8669
β_6	**RELSIZE**	**-0.4496**	0.1291	-3.4823	0.0005
β_7	**LEVE_FIN**	**-0.0758**	0.0314	-2.4107	0.0159
β_8	**LEVE_NO_FIN**	**0.0418**	0.0355	1.1803	0.2379

Table 4.
The convergence of the chains - Diagnostics tests for the equality of the MEAN

Table 4 reports the tests for the equality of the **mean** between the **Chain 1** and the **Chain 2,** with the aim to test the convergence of the chains. The tests for the equality of the **mean** are the following: t-test, Anova F-test. The brackets report the levels of the probability that allow to reject the null hypothesis.

Panel 4.1: t-test for evaluating the equality of the MEAN, between the Chain 1 and the Chain 2

Parameters	MEAN									
	t									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	-1.607 (0.108)	0.904 (0.366)	-4.692 (0.000)	2.750 (0.006)	-2.217 (0.027)	5.120 (0.000)	2.960 (0.003)	5.056 (0.000)	-5.519 (0.000)	7.742 (0.000)
β_1	-0.924 (0.356)	-1.590 (0.112)	-3.849 (0.000)	0.163 (0.871)	3.396 (0.001)	4.514 (0.000)	0.826 (0.409)	1.591 (0.112)	5.034 (0.000)	1.566 (0.117)
β_2	0.114 (0.909)	-1.465 (0.143)	0.504 (0.615)	2.643 (0.008)	-1.455 (0.146)	-1.830 (0.067)	-0.897 (0.370)	-1.012 (0.312)	0.949 (0.342)	-2.705 (0.007)
β_3	1.858 (0.063)	-1.107 (0.268)	4.462 (0.000)	-2.745 (0.006)	3.143 (0.002)	-4.936 (0.000)	-3.024 (0.003)	-5.074 (0.000)	5.375 (0.000)	-7.154 (0.000)
β_4	0.618 (0.536)	-0.517 (0.605)	-0.409 (0.682)	1.162 (0.245)	1.001 (0.317)	-1.292 (0.196)	-2.371 (0.018)	0.797 (0.425)	-0.707 (0.480)	-0.301 (0.763)
β_5	0.587 (0.557)	1.529 (0.126)	2.586 (0.010)	-1.881 (0.060)	-1.049 (0.294)	-0.217 (0.828)	1.400 (0.162)	-0.755 (0.450)	-0.445 (0.656)	-0.639 (0.523)
β_6	-0.904 (0.366)	0.423 (0.672)	0.882 (0.378)	-0.897 (0.370)	-2.221 (0.026)	-1.357 (0.175)	-0.234 (0.815)	-1.166 (0.244)	-0.364 (0.716)	-0.949 (0.343)
β_7	1.680 (0.093)	2.856 (0.004)	4.360 (0.000)	-0.654 (0.513)	-3.174 (0.002)	-4.335 (0.000)	-0.937 (0.349)	-1.349 (0.177)	-4.695 (0.000)	-1.680 (0.093)
β_8	0.004 (0.997)	0.551 (0.581)	0.798 (0.425)	-0.885 (0.376)	-0.461 (0.645)	-0.179 (0.858)	-0.678 (0.498)	0.705 (0.481)	1.086 (0.278)	-1.467 (0.142)

<u>Panel 4.2</u>: Anova-F test for evaluating the equality of the MEAN, between the Chain 1 and the Chain 2

Parameters	MEAN									
	Anova-F									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	2.581 (0.108)	0.817 (0.366)	22.014 (0.000)	7.564 (0.006)	4.915 (0.027)	26.211 (0.000)	8.762 (0.003)	25.560 (0.000)	30.455 (0.000)	59.934 (0.000)
β_1	0.854 (0.356)	2.527 (0.112)	14.816 (0.000)	0.027 (0.871)	11.529 (0.001)	20.373 (0.000)	0.682 (0.409)	2.531 (0.112)	25.345 (0.000)	2.453 (0.117)
β_2	0.013 (0.909)	2.145 (0.143)	0.254 (0.615)	6.984 (0.008)	2.116 (0.146)	3.348 (0.067)	0.804 (0.370)	1.023 (0.312)	0.901 (0.342)	7.315 (0.007)
β_3	3.453 (0.063)	1.226 (0.268)	19.909 (0.000)	7.535 (0.006)	9.880 (0.002)	24.364 (0.000)	9.142 (0.003)	25.747 (0.000)	28.889 (0.000)	51.183 (0.000)
β_4	0.383 (0.536)	0.267 (0.605)	0.168 (0.682)	1.350 (0.245)	1.003 (0.317)	1.669 (0.196)	5.621 (0.018)	0.636 (0.425)	0.499 (0.480)	0.091 (0.763)
β_5	0.345 (0.557)	2.339 (0.126)	6.690 (0.010)	3.538 (0.060)	1.101 (0.294)	0.047 (0.828)	1.961 (0.162)	0.571 (0.450)	0.198 (0.656)	0.408 (0.523)
β_6	0.818 (0.366)	0.179 (0.672)	0.777 (0.378)	0.804 (0.370)	4.934 (0.026)	1.840 (0.175)	0.055 (0.815)	1.360 (0.244)	0.133 (0.716)	0.901 (0.343)
β_7	2.824 (0.093)	8.160 (0.004)	19.007 (0.000)	0.427 (0.513)	10.071 (0.002)	18.796 (0.000)	0.877 (0.349)	1.821 (0.177)	22.038 (0.000)	2.823 (0.093)
β_8	0.000 (0.997)	0.304 (0.581)	0.636 (0.425)	0.784 (0.376)	0.212 (0.645)	0.032 (0.858)	0.460 (0.498)	0.496 (0.481)	1.178 (0.278)	2.153 (0.142)

Table 5.

The convergence of the chains - Diagnostics tests for the equality of the VARIANCE

Table 5 reports the tests for the equality of the **variance** between the **Chain 1** and the **Chain 2,** with the aim to evaluate the convergence of the chains. The tests for the **variance** are the following: F-test, Siegel-Tukey (**S-T**) test, Bartlett test and Brown-Forsythe test (**B-F**). The brackets report the levels of the probability that allow to reject the null hypothesis.

<u>Panel 5.1</u>: F test for evaluating the equality of the VARIANCE, between the Chain 1 and the Chain 2

Parameters	**VARIANCE**									
	F									
	Bayesian 1					**Bayesian 7**				
	ALL	**[0, 0.25)**	**[0.25, 0.50)**	**[0.50, 0.75)**	**[0.75, 1.00]**	**ALL**	**[0, 0.25)**	**[0.25, 0.50)**	**[0.50, 0.75)**	**[0.75, 1.00]**
β_0	1.008 (0.000)	1.001 (0.851)	1.018 (0.000)	1.015 (0.000)	1.031 (0.000)	1.029 (0.000)	1.072 (0.000)	1.012 (0.002)	1.052 (0.000)	1.008 (0.039)
β_1	1.007 (0.001)	1.009 (0.029)	1.030 (0.000)	1.003 (0.424)	1.014 (0.000)	1.001 (0.731)	1.003 (0.400)	1.003 (0.504)	1.005 (0.189)	1.007 (0.067)
β_2	1.007 (0.000)	1.002 (0.661)	1.012 (0.003)	1.001 (0.784)	1.021 (0.000)	1.005 (0.015)	1.002 (0.550)	1.001 (0.817)	1.007 (0.069)	1.009 (0.028)
β_3	1.007 (0.001)	1.004 (0.300)	1.014 (0.001)	1.012 (0.002)	1.031 (0.000)	1.028 (0.000)	1.069 (0.000)	1.013 (0.001)	1.050 (0.000)	1.008 (0.043)
β_4	1.001 (0.666)	1.000 (0.944)	1.003 (0.463)	1.007 (0.097)	1.001 (0.894)	1.004 (0.029)	1.003 (0.407)	1.005 (0.252)	1.007 (0.073)	1.002 (0.056)
β_5	1.004 (0.028)	1.001 (0.722)	1.018 (0.000)	1.005 (0.253)	1.006 (0.141)	1.003 (0.103)	1.007 (0.095)	1.005 (0.190)	1.005 (0.193)	1.009 (0.020)
β_6	1.003 (0.127)	1.004 (0.298)	1.001 (0.862)	1.013 (0.002)	1.005 (0.253)	1.001 (0.616)	1.004 (0.344)	1.002 (0.562)	1.001 (0.863)	1.009 (0.019)
β_7	1.006 (0.004)	1.009 (0.031)	1.017 (0.000)	1.010 (0.012)	1.012 (0.004)	1.001 (0.564)	1.003 (0.530)	1.001 (0.780)	1.003 (0.531)	1.006 (0.158)
β_8	1.004 (0.084)	1.002 (0.694)	1.005 (0.205)	1.009 (0.021)	1.001 (0.777)	1.003 (0.167)	1.003 (0.417)	1.000 (0.949)	1.004 (0.300)	1.003 (0.397)

Panel 5.2: **SIEGEL-TUKEY test for evaluating the equality of the VARIANCE, between the Chain 1 and the Chain 2**

Parameters	VARIANCE									
	SIEGEL-TUKEY									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	1.003 (0.316)	0.169 (0.866)	0.663 (0.508)	3.082 (0.002)	0.591 (0.554)	15.971 (0.000)	11.963 (0.000)	2.992 (0.003)	11.690 (0.000)	5.447 (0.000)
β_1	1.785 (0.074)	1.855 (0.064)	1.815 (0.070)	0.182 (0.856)	0.053 (0.958)	0.978 (0.328)	0.496 (0.620)	1.878 (0.060)	1.666 (0.096)	1.227 (0.220)
β_2	1.483 (0.138)	0.755 (0.451)	0.114 (0.909)	0.024 (0.981)	3.574 (0.000)	1.613 (0.107)	1.241 (0.215)	0.013 (0.989)	0.128 (0.898)	2.099 (0.036)
β_3	0.761 (0.447)	0.188 (0.851)	0.092 (0.927)	3.267 (0.001)	2.029 (0.042)	14.862 (0.000)	11.238 (0.000)	2.569 (0.010)	10.713 (0.000)	5.292 (0.000)
β_4	0.429 (0.668)	0.011 (0.991)	1.542 (0.123)	1.731 (0.083)	2.427 (0.015)	0.077 (0.939)	0.245 (0.806)	0.281 (0.779)	0.853 (0.394)	0.963 (0.336)
β_5	0.235 (0.815)	3.277 (0.001)	1.708 (0.088)	1.644 (0.100)	0.546 (0.585)	0.391 (0.696)	0.624 (0.533)	0.538 (0.591)	0.490 (0.624)	0.109 (0.913)
β_6	0.120 (0.904)	0.110 (0.912)	0.109 (0.913)	0.905 (0.366)	0.674 (0.500)	0.511 (0.609)	0.450 (0.652)	0.050 (0.960)	0.039 (0.969)	1.466 (0.143)
β_7	2.563 (0.010)	0.842 (0.400)	3.495 (0.001)	1.711 (0.087)	0.938 (0.349)	0.810 (0.418)	0.276 (0.783)	2.237 (0.025)	0.816 (0.415)	0.480 (0.632)
β_8	1.678 (0.093)	0.777 (0.437)	3.014 (0.003)	1.902 (0.057)	0.782 (0.435)	1.476 (0.140)	1.428 (0.153)	0.992 (0.321)	0.187 (0.851)	0.714 (0.475)

Panel 5.3: BARTLETT test for evaluating the equality of the VARIANCE, between the Chain 1 and the Chain 2

Parameters	VARIANCE									
	BARTLETT									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	16.791 (0.000)	0.035 (0.851)	20.276 (0.000)	14.063 (0.000)	59.849 (0.000)	205.442 (0.000)	300.824 (0.000)	9.554 (0.002)	158.608 (0.000)	4.254 (0.039)
β_1	11.114 (0.001)	4.746 (0.029)	54.505 (0.000)	0.639 (0.424)	12.895 (0.000)	0.118 (0.731)	0.710 (0.400)	0.448 (0.504)	1.725 (0.189)	3.346 (0.067)
β_2	13.306 (0.000)	0.193 (0.661)	8.724 (0.003)	0.075 (0.784)	25.838 (0.000)	5.894 (0.015)	0.357 (0.550)	0.053 (0.817)	3.320 (0.068)	4.825 (0.028)
β_3	11.783 (0.001)	1.073 (0.300)	11.601 (0.001)	9.637 (0.002)	58.972 (0.000)	189.236 (0.000)	280.240 (0.000)	10.416 (0.001)	148.774 (0.000)	4.100 (0.043)
β_4	0.187 (0.666)	0.005 (0.944)	0.540 (0.463)	2.763 (0.097)	0.018 (0.894)	4.749 (0.029)	0.689 (0.407)	1.313 (0.252)	3.208 (0.073)	0.347 (0.556)
β_5	4.848 (0.028)	0.127 (0.722)	19.142 (0.000)	1.309 (0.253)	2.167 (0.141)	2.661 (0.103)	2.791 (0.095)	1.720 (0.190)	1.696 (0.193)	5.433 (0.020)
β_6	2.328 (0.127)	1.084 (0.298)	0.030 (0.862)	9.740 (0.002)	1.308 (0.253)	0.251 (0.616)	0.894 (0.344)	0.336 (0.562)	0.030 (0.863)	5.520 (0.019)
β_7	8.446 (0.004)	4.643 (0.031)	16.559 (0.000)	6.312 (0.012)	8.201 (0.004)	0.333 (0.564)	0.395 (0.530)	0.066 (0.798)	0.393 (0.531)	1.997 (0.158)
β_8	2.985 (0.084)	0.154 (0.694)	1.608 (0.205)	5.358 (0.021)	0.080 (0.777)	1.906 (0.167)	0.660 (0.417)	0.004 (0.949)	1.077 (0.300)	0.718 (0.397)

<u>Panel 5.4</u>: BROWN-FORSYTHE test on the equality of the VARIANCE, between the Chain 1 and the Chain 2

Parameters	VARIANCE									
	BROWN-FORSYTHE									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	1.507 (0.220)	0.002 (0.962)	6.190 (0.013)	14.219 (0.000)	14.836 (0.000)	256.825 (0.000)	226.184 (0.000)	0.084 (0.771)	169.760 (0.000)	14.445 (0.000)
β_1	8.029 (0.005)	5.213 (0.022)	22.421 (0.000)	0.090 (0.764)	2.686 (0.101)	0.422 (0.516)	0.858 (0.354)	0.527 (0.468)	3.363 (0.067)	2.161 (0.142)
β_2	7.350 (0.007)	0.349 (0.555)	2.431 (0.119)	0.016 (0.898)	21.147 (0.000)	4.572 (0.033)	1.082 (0.298)	0.017 (0.895)	0.650 (0.420)	5.276 (0.022)
β_3	1.001 (0.317)	0.160 (0.689)	2.474 (0.116)	13.275 (0.000)	20.659 (0.000)	227.414 (0.000)	203.402 (0.000)	0.000 (0.998)	150.092 (0.000)	13.851 (0.000)
β_4	0.159 (0.690)	0.000 (0.991)	0.232 (0.630)	3.384 (0.066)	2.361 (0.124)	1.096 (0.295)	0.391 (0.532)	0.173 (0.678)	1.752 (0.186)	0.073 (0.788)
β_5	0.834 (0.361)	2.703 (0.100)	9.931 (0.002)	2.019 (0.155)	1.212 (0.271)	0.304 (0.582)	1.403 (0.236)	0.089 (0.766)	0.855 (0.355)	1.135 (0.287)
β_6	0.607 (0.436)	0.241 (0.623)	0.013 (0.908)	3.978 (0.046)	0.025 (0.875)	0.365 (0.546)	0.445 (0.505)	0.041 (0.840)	0.007 (0.935)	3.977 (0.046)
β_7	8.475 (0.004)	3.107 (0.078)	17.122 (0.000)	4.042 (0.044)	4.195 (0.041)	0.352 (0.579)	0.096 (0.757)	1.224 (0.269)	1.096 (0.295)	0.673 (0.412)
β_8	3.702 (0.054)	0.224 (0.636)	5.664 (0.017)	5.033 (0.025)	0.079 (0.778)	2.068 (0.151)	1.200 (0.273)	0.480 (0.488)	0.206 (0.650)	0.400 (0.527)

Figure 1.

Estimated Probability of Resolution for Mergers and Acquisitions (M&As)

The figure shows two **Relative Operating Characteristic (ROC) curves**: (i) **Figure 1.1** shows the percentage of correctly classified *successful* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. (ii) **Figure 1.2** shows the percentage of correctly classified *withdrawn* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. **The figures show in-sample results, based on a sample of 1090 U.S. stock swap M&As.**

Figure 1.1:

Successful Mergers and Acquisitions (M&As)

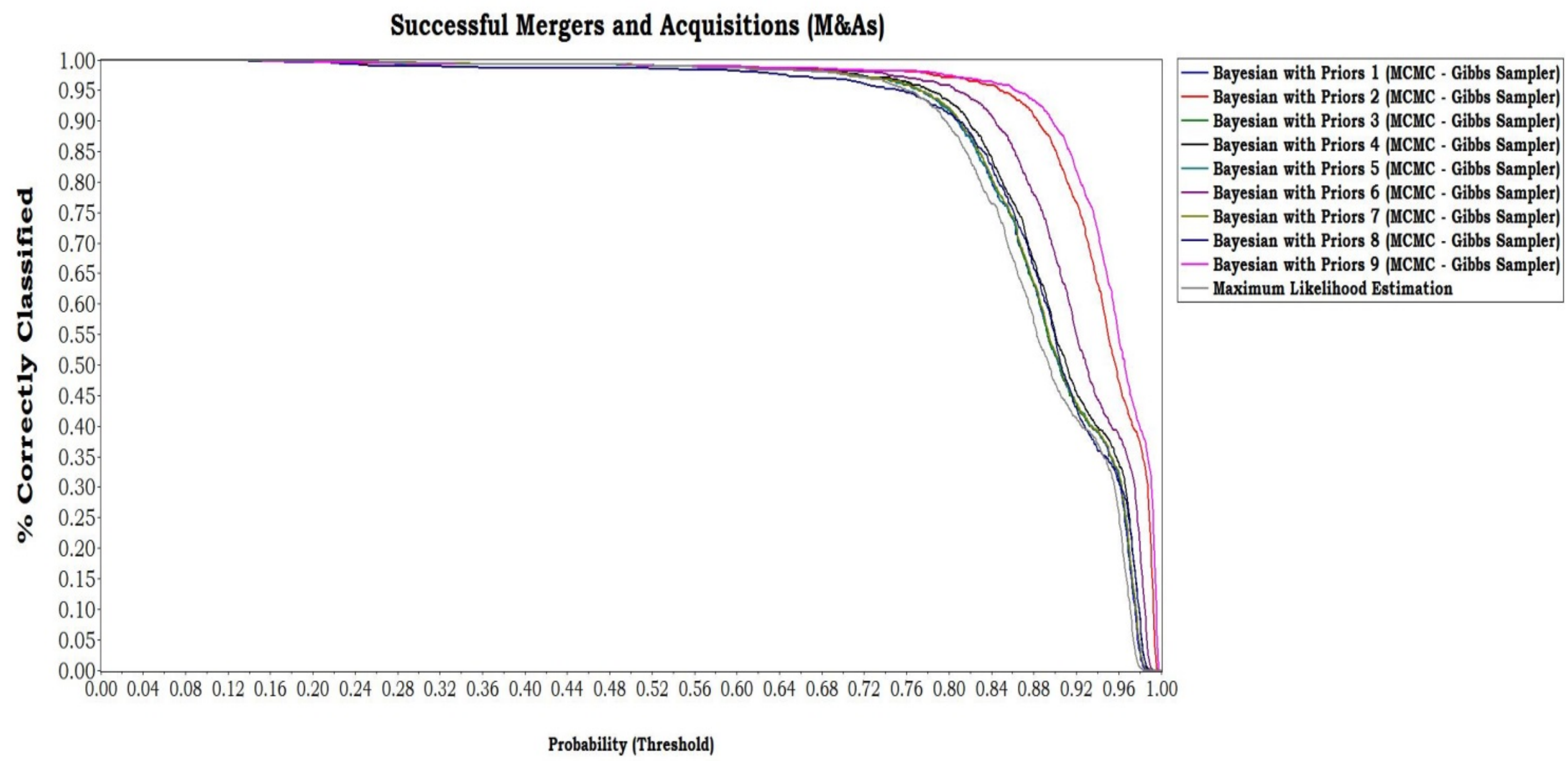


Figure 1.2:
Withdrawn Mergers and Acquisitions (M&As)

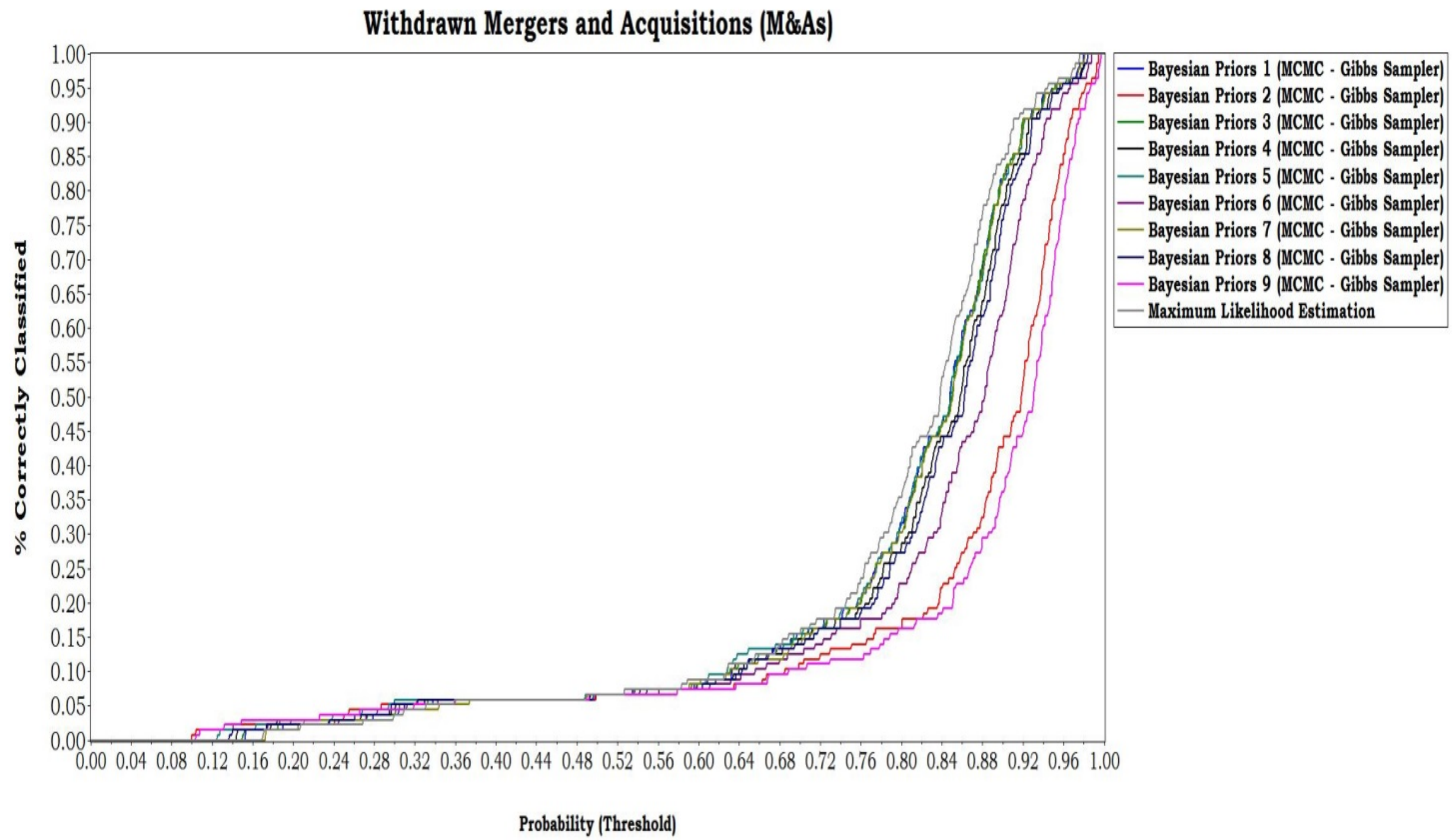


Figure 2.
Pseudo Out-of-Sample validation for Mergers and Acquisitions (M&As), based on 50% of the sample

The figure shows two **Relative Operating Characteristic (ROC) curves**: (i) **Figure 2.1** shows the percentage of correctly classified *successful* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. (ii) **Figure 2.2** shows the percentage of correctly classified *withdrawn* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. The Figures show the pseudo out-of-sample results, **based on 50% of M&As**.

Figure 2.1:
ROC - Pseudo out-of-sample validation for successful M&As (50% of M&As)

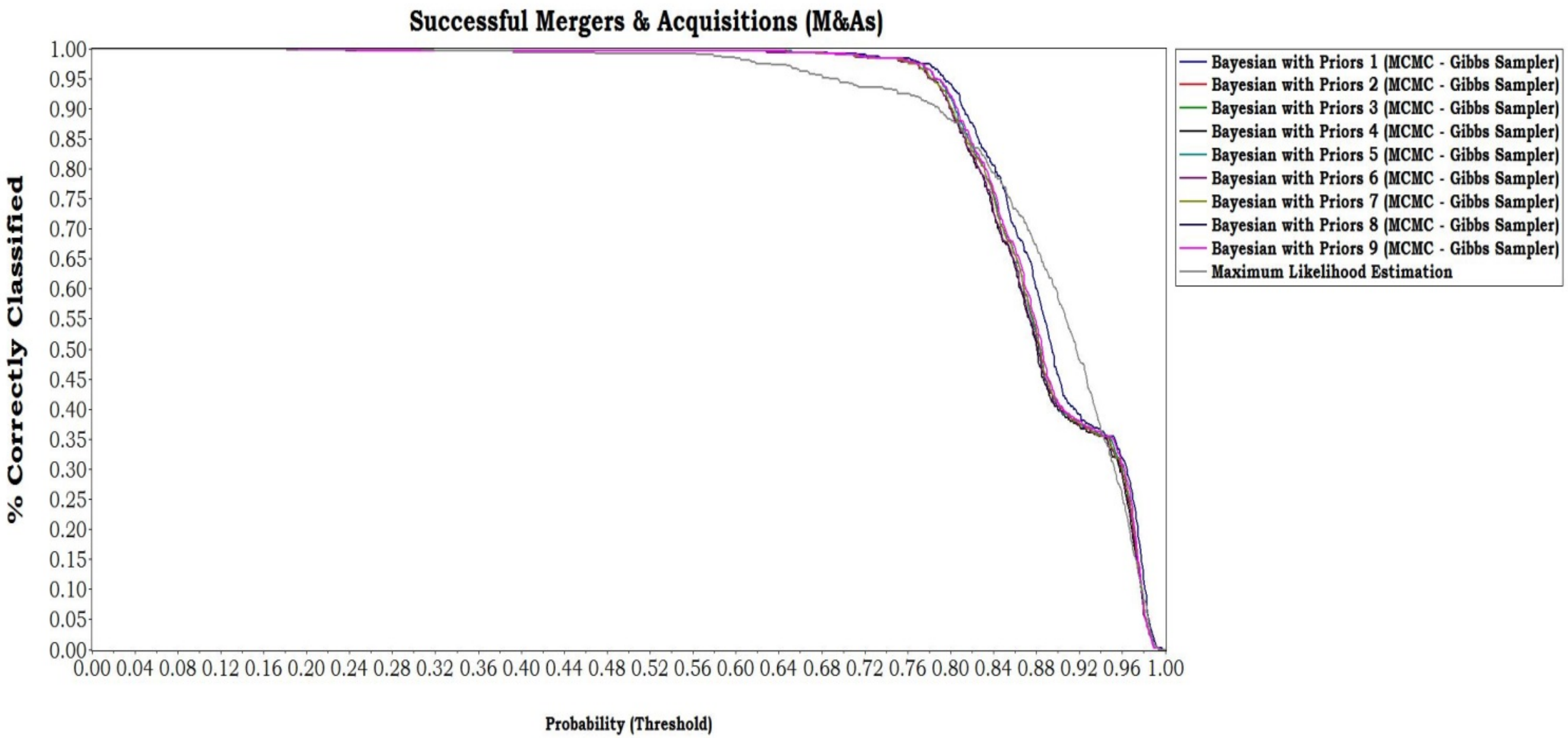


Figure 2.2:
ROC - Pseudo out-of-sample validation for withdrawn M&As (50% of M&As)

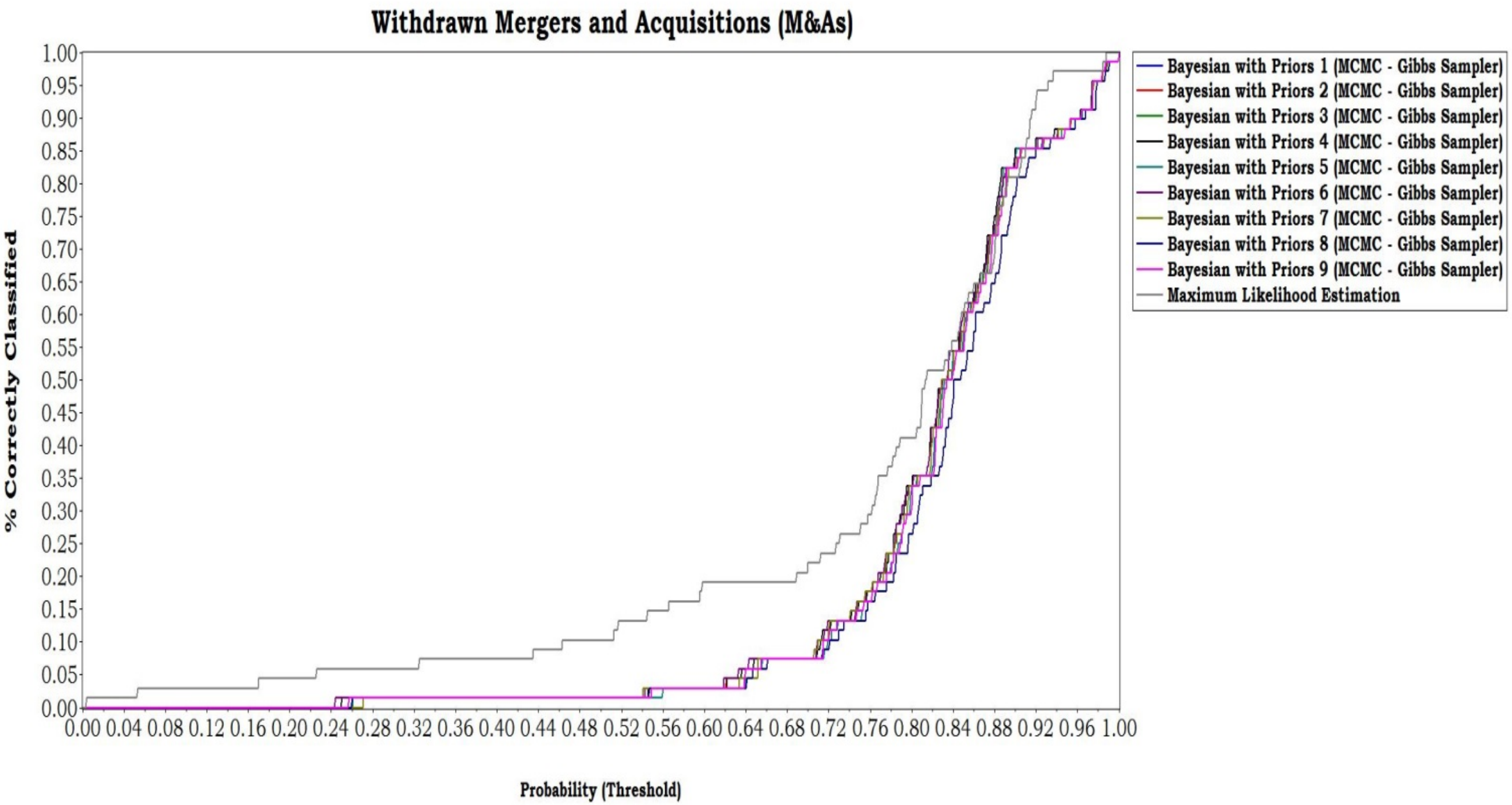


Figure 3.

<u>Pseudo Out-of-Sample validation for Mergers and Acquisitions (M&As), based on 30% of the sample</u>

The figure shows two **Relative Operating Characteristic (ROC) curves**: (i) **Figure 3.1** shows the percentage of correctly classified *<u>successful</u>* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. (ii) **Figure 3.2** shows the percentage of correctly classified *<u>withdrawn</u>* Mergers and Acquisitions (M&As), as a function of the probability from 0 to 1. The Figures show the pseudo out-of-sample results, **based on 30% of M&As**.

Figure 3.1:

<u>ROC - Pseudo out-of-sample validation for successful M&As (30% of M&As)</u>

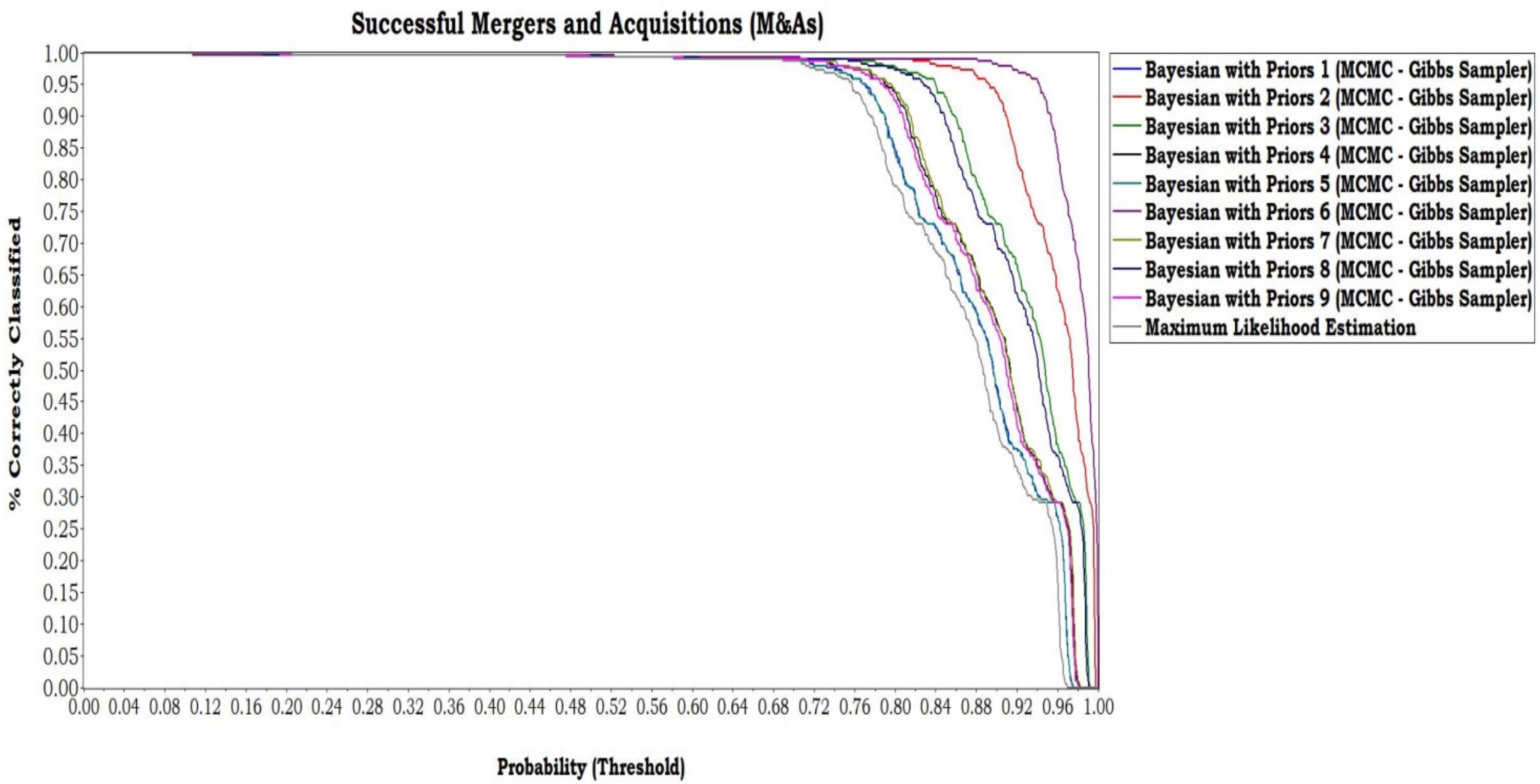
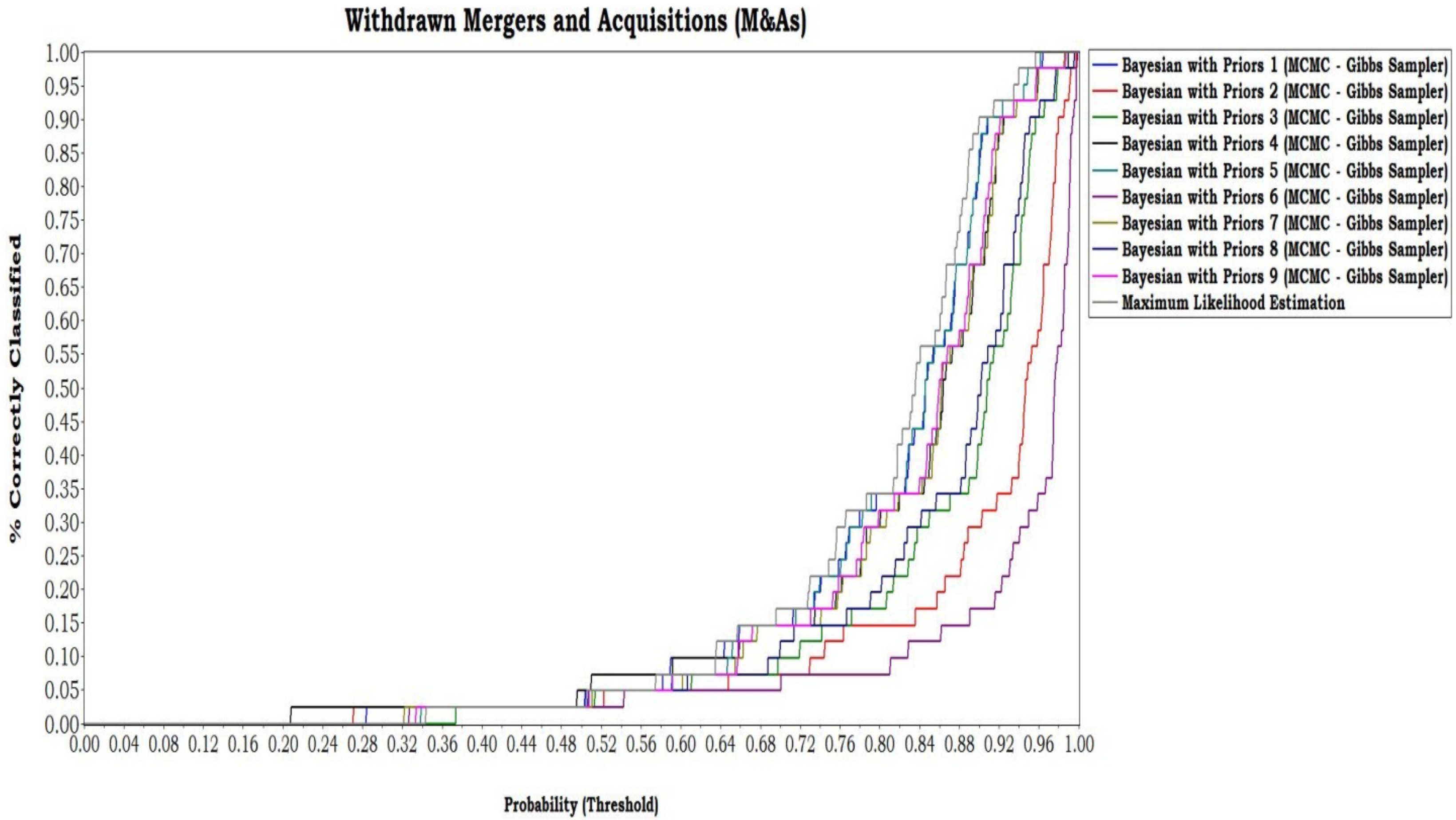


Figure 3.2:
ROC - Pseudo out-of-sample validation for withdrawn M&As (30% of M&As)



Appendix A.
The list of covariates

Appendix A reports the list of dependent (SUCCESS) and independent (FINANCIAL DUMMY, DIVERS, FRIENDLY, RET_TARGET_PRE, RET_ACQUIRER_PRE, RELSIZE, LEVE_FIN, LEVE_NO_FIN) variables, with the aim to estimate the probability of a deal completion.

Variables	Definition
SUCCESS	**Dependent Variable** Dummy variable that takes a value equals to 1, in case of a deal completion and 0 otherwise
	Independent Variables
FINANCIAL DUMY	Dummy variable that takes a value equals to 1, if the target company is a financial institution
DIVERS	Dummy variable that takes a value equals to 1, if the target and the bidder companies belong to the same primary business
FRIENDLY	Dummy variable that takes a value equals to 1, if the deal is friendly and a value equals to 0, if it is hostile
RET_TARGET_PRE	Percentage variation of the target stock price, 1 week before the announcement date (TARGET RUN-UP)
RET_ACQUIRER_PRE	Percentage variation of the acquirer stock price, 1 week before the announcement date (ACQUIRER RUN-UP)
RELSIZE	Ratio between the target market value and the acquirer market value
LEVE_FIN	Leverage computed as the ratio between the total amount of assets and the total amount of equity, for target financial companies
LEVE_NO_FIN	Leverage computed as the ratio between the total amount of assets and the total amount of equity, for target non-financial companies

Appendix B.

Diagnostics for the convergence of the chains

Appendix B reports the convergence rates for the estimated parameters of the logistic model, based on a different set of (informative) priors. The estimation procedure relies on the **MCMC simulations with a Gibbs sampler**. **Appendix B** shows the convergence rate, computed as a ratio between the MC error and the standard deviation of the posterior distribution, the root median square deviation (**RMSD**) and the median absolute percentage deviation (**MAPD**), between the chains for each parameter. The number of iterations is equal to 1000000 and the procedure relies on 2 chains. The "**burn-in**" procedure for computing the descriptive statistics for each posterior distribution is equal to 10% of the observations. The results are provided for the Bayesian estimation procedure, based on the **first** and the **seventh** set of priors.

Bayesian with Priors 1

Parameters	Covariates	Convergence Rate (x100)	RMSD	MAPD
β_0		0.2034	0.674	0.581
β_1	**FINANCIAL DUMMY**	0.1622	0.547	0.224
β_2	**DIVERS**	0.1389	0.195	0.418
β_3	**FRIENDLY**	0.2050	0.651	0.242
β_4	**RET_ACQUIRER_PRE**	0.1197	0.777	0.312
β_5	**RET_TARGET_PRE**	0.1264	0.434	1.380
β_6	**RELSIZE**	0.1227	0.126	0.274
β_7	**LEVE_FIN**	0.1528	0.033	0.408
β_8	**LEVE_NO_FIN**	0.1557	0.035	0.713

Bayesian with Priors 7

Parameters	Covariates	Convergence Rate (x100)	RMSD	MAPD
β_0		0.6710	0.656	0.628
β_1	**FINANCIAL DUMMY**	0.2308	0.546	0.224
β_2	**DIVERS**	0.1495	0.195	0.419
β_3	**FRIENDLY**	0.6532	0.632	0.245
β_4	**RET_ACQUIRER_PRE**	0.0911	0.780	0.314
β_5	**RET_TARGET_PRE**	0.1041	0.434	1.381
β_6	**RELSIZE**	0.0819	0.126	0.274
β_7	**LEVE_FIN**	0.2082	0.033	0.408
β_8	**LEVE_NO_FIN**	0.0986	0.035	0.715

Appendix C.
Diagnostics tests, based on the equality of the MEAN, for evaluating the convergence of the chains

Appendix C reports alternative tests for the equality of the **mean** between the **Chain 1** and the **Chain 2,** with the aim to test the convergence of the chains. The alternative tests on the **mean** are the following: Satterwhaite-Welch test, Welch test. The brackets report the levels of the probability that allow to reject the null hypothesis.

Appendix C.1: Satterwhaite-Welch test for the equality of the MEAN, between the Chain 1 and the Chain 2

Parameters	MEAN									
	Satterwhaite-Welch									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	-1.607 (0.108)	0.904 (0.366)	-4.692 (0.000)	2.750 (0.006)	-2.217 (0.027)	5.120 (0.000)	2.960 (0.003)	5.056 (0.000)	-5.519 (0.000)	7.742 (0.000)
β_1	-0.924 (0.356)	-1.590 (0.112)	-3.849 (0.000)	0.163 (0.871)	3.396 (0.001)	4.514 (0.000)	0.826 (0.409)	1.591 (0.112)	5.034 (0.000)	1.566 (0.117)
β_2	0.114 (0.909)	-1.465 (0.143)	0.504 (0.615)	2.643 (0.008)	-1.455 (0.146)	-1.830 (0.067)	-0.897 (0.370)	-1.012 (0.312)	0.949 (0.342)	-2.705 (0.007)
β_3	1.858 (0.063)	-1.107 (0.268)	4.462 (0.000)	-2.745 (0.006)	3.143 (0.002)	-4.936 (0.000)	-3.024 (0.003)	-5.074 (0.000)	5.375 (0.000)	-7.154 (0.000)
β_4	0.618 (0.536)	-0.517 (0.605)	-0.409 (0.682)	1.162 (0.245)	1.001 (0.317)	-1.292 (0.196)	-2.371 (0.018)	0.797 (0.425)	-0.707 (0.480)	-0.301 (0.763)
β_5	0.587 (0.557)	1.529 (0.126)	2.587 (0.010)	-1.881 (0.060)	-1.049 (0.294)	-0.217 (0.828)	1.400 (0.162)	-0.755 (0.450)	-0.445 (0.656)	-0.639 (0.523)
β_6	-0.904 (0.366)	0.423 (0.672)	0.882 (0.378)	-0.897 (0.370)	-2.221 (0.026)	-1.357 (0.175)	-0.234 (0.815)	-1.166 (0.244)	0.133 (0.716)	-0.949 (0.343)
β_7	1.680 (0.093)	2.856 (0.004)	4.360 (0.000)	-0.654 (0.513)	-3.173 (0.002)	-4.335 (0.000)	-0.937 (0.349)	-1.349 (0.177)	-4.695 (0.000)	-1.680 (0.093)
β_8	0.004 (0.997)	0.551 (0.581)	0.798 (0.425)	-0.885 (0.376)	-0.461 (0.645)	-0.179 (0.858)	-0.678 (0.498)	0.705 (0.481)	1.086 (0.278)	-1.467 (0.142)

Appendix C.2: Welch test for the equality of the MEAN, between the Chain 1 and the Chain 2

Parameters	MEAN									
	Welch									
	Bayesian 1					Bayesian 7				
	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]	ALL	[0, 0.25)	[0.25, 0.50)	[0.50, 0.75)	[0.75, 1.00]
β_0	2.581 (0.108)	0.817 (0.366)	22.014 (0.000)	7.564 (0.006)	4.915 (0.027)	26.211 (0.000)	8.762 (0.003)	25.560 (0.000)	30.455 (0.000)	59.934 (0.000)
β_1	0.854 (0.356)	2.527 (0.112)	14.816 (0.000)	0.027 (0.871)	11.529 (0.001)	20.373 (0.000)	0.682 (0.409)	2.531 (0.112)	25.345 (0.000)	2.453 (0.117)
β_2	0.013 (0.909)	2.145 (0.143)	0.254 (0.615)	6.984 (0.008)	2.116 (0.146)	3.348 (0.067)	0.804 (0.370)	1.023 (0.312)	0.901 (0.342)	7.315 (0.007)
β_3	3.453 (0.063)	1.226 (0.268)	19.909 (0.000)	7.535 (0.006)	9.880 (0.002)	24.364 (0.000)	9.142 (0.003)	25.747 (0.000)	28.889 (0.000)	51.183 (0.000)
β_4	0.383 (0.536)	0.267 (0.605)	0.168 (0.682)	1.350 (0.245)	1.003 (0.317)	1.669 (0.196)	5.621 (0.018)	0.636 (0.425)	0.499 (0.480)	0.091 (0.763)
β_5	0.345 (0.557)	2.339 (0.126)	6.690 (0.010)	3.538 (0.060)	1.001 (0.294)	0.047 (0.828)	1.961 (0.162)	0.571 (0.450)	0.198 (0.656)	0.408 (0.523)
β_6	0.818 (0.366)	0.179 (0.672)	0.777 (0.378)	0.804 (0.370)	4.934 (0.026)	1.840 (0.175)	0.055 (0.815)	1.360 (0.244)	0.133 (0.716)	0.901 (0.343)
β_7	2.824 (0.093)	8.160 (0.004)	19.007 (0.000)	0.427 (0.513)	10.072 (0.002)	18.796 (0.000)	0.877 (0.349)	1.821 (0.177)	22.038 (0.000)	2.823 (0.093)
β_8	0.000 (0.997)	0.304 (0.581)	0.636 (0.425)	0.784 (0.376)	0.212 (0.645)	0.032 (0.858)	0.460 (0.498)	0.496 (0.481)	1.178 (0.278)	2.153 (0.142)