

Charles University
Faculty of Social Sciences
Institute of Economic Studies



MASTER'S THESIS

**Determinants of the Mode of Payment in
Mergers & Acquisitions in the European
Union**

Author: **Bc. Adam Maryniok**

Supervisor: **prof. Ing. Evžen Kočenda, M.A., Ph.D., DSc.**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, December 29, 2018

Signature

Acknowledgments

The author is grateful especially to his parents who were of enormous support during his entire university studies. Furthermore, the author would like to thank to the thesis's supervisor, Mr. Evžen Kočenda, who provided guidance and valuable commentary. Many thanks are extended also to Mr. Tomáš Havránek, Michal Mejstřík and Josef Baruník who were always willing to answer various questions regarding the thesis. Furthermore, this work would not be possible without Mr. Petr Ščučinský of Reuters, who provided the author with an initial access to the Reuters Eikon database. Last but foremost, the author thanks his girlfriend for never-ending support.

Abstract

Topic of mergers and acquisitions (M&A) is popular both in academia and financial circles and press. A great deal of research has been focused on the value creation side of M&A deals, nonetheless factors influencing the particular method of payment used in M&A transactions are equally interesting. This thesis focuses on number of factors influencing the choice of medium of exchange in M&A deals with European Union domiciled bidders. Using Bayesian model averaging and a relatively new dataset of transactions announced between 2010 and 2018, the analysis finds several bidder, target and deal specific characteristics to be of a provable effect on the choice of payment. Finally, several enhancements and research questions for a further research are identified.

JEL Classification G24, G30, G32, G34

Keywords mergers and acquisitions, M&A, method of payment, Bayesian model averaging, BMA

Author's e-mail adam.maryniok@gmail.com

Supervisor's e-mail kocenda@fsv.cuni.cz

Abstrakt

Oblast fúzí a akvizic je populárním tématem jak v akademické obci, tak ve finančních kruzích a tisku. Množství výzkumných prací se zaměřovalo a zaměřuje na oblast tvorby, respektive destrukce akcionářské hodnoty v souvislosti s M&A. Zkoumání problematiky výběru platební metody v M&A je však neméně zajímavým tématem. Tato práce studuje různé faktory ovlivňující výběr platební metody v transakcích, v nichž byla kupujícími společnost se sídlem v Evropské unii. Práce používá metodu bayesovského průměrování modelů na relativně novém vzorku transakcí oznámených v průběhu let 2010 až 2018. Výsledky analýzy potvrzují prokazatelný vliv některých faktorů, specifických jak na úrovni kupce, tak akvizičního cíle a transakcí jako takových. Závěrem práce nabízí několik nápadů, kam by se mohl ubírat budoucí výzkum na toto téma.

Klasifikace JEL	G24, G30, G32, G34
Klíčová slova	fúze a akvizice, M&A, typ kompenzace, Bayesovské průměrování modelů, BMA
E-mail autora	adam.maryniok@gmail.com
E-mail vedoucího práce	kocenda@fsv.cuni.cz

Contents

List of Tables	viii
List of Figures	ix
Acronyms	x
Thesis Proposal	xi
1 Introduction	1
2 Literature Review	3
2.1 Theoretical Models	3
2.2 Empirical Studies	4
3 Theoretical Framework	9
3.1 Logit Model	10
3.1.1 Logit Model Estimation	11
3.1.2 Bayesian Model Averaging	12
4 Methodology	14
4.1 Financing Decision Hypotheses	14
4.1.1 Financial Factors	14
4.1.2 Corporate Governance Factors	15
4.1.3 Information Asymmetry Factors	15
4.1.4 Other Factors	17
4.2 Full model specification	17
5 Data	19
5.1 Data Description	19

6	Empirical results	25
6.1	Full Model Results	25
6.2	Model Selection	26
6.2.1	Automatic Methods	27
6.2.2	All Possible Regressions Approach	30
6.2.3	Bayesian Model Averaging	34
6.2.4	Model Selection - Summary	39
6.3	Regression Analysis - Results	43
6.3.1	Evidence for an Effect of Independent Variables	43
6.3.2	Coefficients Interpretation and Comparison with Previous Research	43
6.4	Binary Logistic Regression Analysis	47
7	Conclusion	51
	Bibliography	59
A	Histograms of Independent Variables Categorized by the Payment Method	I
B	Regression Tables of Models Selected by Automatic Selection Methods	V
C	BMA Posterior Distribution of Independent Variables	VII

List of Tables

5.1	Distribution of Transactions by Year and Payment	20
5.2	Country Distribution of Transactions	22
5.3	Descriptive Statistics of Independent Variables	24
6.1	Comparison of the Full Model & the Backward Elimination Model	29
6.2	Best 10 Models Selected by the Best Subset Regression	32
6.3	Best 5 Models Selected by BMA	37
6.4	BMA estimates	38
6.5	Summary of BMA regression results	42
6.6	Summary of BMA Regression Results	50
B.1	Models Selected by the Automatic Selection Methods	VI

List of Figures

5.1	Visualization of Missing Data Points	20
6.1	AIC Path of the Automatic Algorithms	30
6.2	AIC Profile	31
6.3	Model-averaged Importance of Terms	33
6.4	Model Inclusion in Bayesian Model Averaging	36
6.5	Posterior Probabilities from BMA versus P-values from the Full Model and Stepwise model	41
6.6	Model Inclusion in Bayesian Model Averaging - Binomial Logit .	49
A.1	Conditional histogram for the dependent variable <i>Cash</i>	II
A.2	Conditional histogram for the dependent variable <i>Stock</i>	III
A.3	Conditional histogram for the dependent variable <i>Mix</i>	IV
C.1	Posterior Distribution of Independent Variables	VIII

Acronyms

MA Mergers and Acquisitions

UK United Kingdom

BMA Bayesian Model Averaging

AIC Akaike Information Criterion

BIC Bayesian Information Criterion

USA United States of America

LSE London Stock Exchange

Master's Thesis Proposal

Author	Bc. Adam Maryniok
Supervisor	prof. Ing. Evžen Kočenda, M.A., Ph.D., DSc.
Proposed topic	Determinants of the Mode of Payment in Mergers & Acquisitions in the European Union

Motivation The market for corporate control or simply put mergers & acquisitions (M&A) are a common place in developed economies. The worldwide M&A activity for 2017 alone is projected to be at 45,000 deals or USD 3,154 billion in value terms (compare it to the value of Czech GDP which stood at USD 192 billion in 2016). M&A are a frequent topic of both press and academics due the size of the market, the general importance of M&A in terms of improving corporate efficiency and somewhat popular allure in both financial and public circles.

Most of the research on M&A has focused on the value creation for shareholders stemming from the transactions as it is the most essential element of each deal (at least it ought to be, nonetheless it is not always the case). However, another important feature of each is the medium of exchange used, be it cash, payment in stocks or mix of both. What factors determine the type of payment in M&A? How do stock markets react to the announcement of the payment method? And what is the underlying logic? These questions are the primary motivation for the paper.

Hypotheses The academic literature on the topic provides us with a range of testable hypotheses.

First, how the corporate governance of a company (both bidder and the target) influences the medium of exchange. Both the theory and empirics suggests that for instance targets with a high ownership concentration prefer rather a consideration in shares in order not to lose the control influence and at the same time to acquire sort of an employment insurance for its managers.

Financial characteristics come into play as well. The relative size of acquirer and target has an influence on the exchange medium, as well as for instance indebtedness.

Is the target privately held or a subsidiary? Then liquidity concerns ought to influence the transaction as it is reasonable to assume that cash will be the preferred mode of payment.

Another key factor is information asymmetry. This problem is mostly peculiar to transactions where the target is a privately held company. The stock offer ought to be preferred to the bidder, primarily if the target is relatively large, as it aligns interest of the target with the acquirer.

Last but not least, stock market valuations play role as well. Theory suggests that overvalued acquirers offer a stock consideration and vice versa.

Hypothesis #1: Cash financing is preferable to the acquirer when the target has concentrated ownership and the deal size is relatively big compared to acquirer's equity.

Hypothesis #2: Stock consideration ought to be less attractive for the target if there is an information asymmetry issue (vice versa for the acquirer).

Hypothesis #3: Overvalued bidders tend to pay for targets in stocks.

Hypothesis #4: Companies with a higher leverage are inclined to use stocks for financing of acquisitions.

Methodology I will utilize data on M&A where i) the bidder's domicile is in the EU, ii) the bidder is publicly listed company and iii) the deal closed between 2010 – 2016. Next, a probit/logit model will be estimated with the dependent variable being the transaction's exchange medium (cash, stock or mix of both) and several explanatory variables concerning the corporate governance, financial characteristic, information asymmetry and market valuations. The results will be then compared with previous studies on the topic, primarily with Faccio and Masulis (2005).

Furthermore, the model's predictive power will be then tested on out-of-sample transactions and a sensitivity analysis (by employing more specific regressors) will be conducted in order to calibrate the model further.

Expected Contribution The last study on the topic was done in 2005 by Faccio and Masulis who used data on M&A transactions which took place between 1997 – 2000. However, the situation on the market for corporate control and generally on capital markets back then was a far cry from the current situation as Europe has become far more integrated since then. Hence, the paper ought to bring a fresh perspective and determine whether the factors concerning M&A transactions have stayed the same or changed. A more practical contribution ought to be the predictive power of the econometric model and its use in identifying M&A deals characteristics.

These in turn have empirically proved influence on the target and bidder's stock valuations and consequently might be exploited by investors in the M&A arbitrages.

Outline

1. Introduction
2. Literature Overview and Hypotheses
3. Methodology
4. Data Description
5. Model estimation and Hypotheses Testing
6. Model Testing and Sensitivity Analysis
7. Outcomes discussion and Comparison of the Outcomes with Previous Studies
8. Conclusion

Core bibliography

Hansen, R. (1987). A Theory for the Choice of Exchange Medium in Mergers and Acquisitions. *The Journal of Business*, 60(1), 75-95. Retrieved from <http://www.jstor.org/stable/2352948>

Faccio, M., & Masulis, R. (2005). The Choice of Payment Method in European Mergers and Acquisitions. *The Journal of Finance*, 60(3), 1345-1388. Retrieved from <http://www.jstor.org/stable/3694929>

Sundaram, A. K. (2004). Mergers and acquisitions and corporate governance. In *Advances in mergers and acquisitions* (pp. 193-219). Emerald Group Publishing Limited.

Rappaport, Alfred and Sirower, Mark L. (1999, Nov-Dec). Stock or Cash? The trade-offs for Buyers and Sellers in Mergers and Acquisitions. *Harvard Business Review*.

Ghosh, A. and Ruland, W. (1998), Managerial Ownership, the Method of Payment for Acquisitions, and Executive Job Retention. *The Journal of Finance*, 53: 785-798. doi:10.1111/0022-1082.325125

Author

Supervisor

Chapter 1

Introduction

The market for corporate control or simply put mergers & acquisitions (M&A) is a common place in developed economies. The worldwide M&A activity in 2017 exceeded USD 3 trillion for the fourth consecutive year (Massoudi *et al.* (2017)) and in 2018 is expected to reach the USD 4 billion mark should the final quarter of 2018 follow in the tracks of the previous 9 months (Platt (2018)). M&A is a frequent topic of both press and academics due the size of the market, the general importance of M&A in terms of improving corporate efficiency and somewhat popular allure in both financial and public circles. Most of the research on M&A has focused on the value creation for shareholders stemming from the transactions as it is the most essential element of each deal. However, another important feature of each transaction is the medium of exchange used, be it cash, payment in stocks or mix of both. A fair amount of research has been done on this topic since the first theoretical model of the choice of exchange medium in M&A by Hansen (1987). Since then, researchers have been investigating various factors affecting the choice of payment, both theoretically and empirically. Nonetheless, majority of the research have been focused on deals originated in the USA. To our knowledge, the only one research paper investigating European transactions was written by Faccio & Masulis (2005).

This thesis aims to build on the research knowledge which has been gathered on the topic over the years and investigate factors influencing the choice of payment in M&A transactions with a focus on bidder companies domiciled in the European Union countries and, along the way, to contribute with its up-to-date findings to the already vast body of research done on corporate finance topics. We believe that the main difference of this thesis in comparison with

the previous research is, first, in using Bayesian model averaging for statistical inference rather than the classical frequentist approach based on P-values, which tends to overestimate effects as shown in this thesis as well as elsewhere (see Hoeting *et al.* (1999) or Kjérulff (2017)). A scientific journal *Basic and Applied Social Psychology* even stopped publishing research papers based on the null hypothesis significance testing procedure, i.e. using P-values for the (non)-rejection of null hypotheses. Second, as previously mentioned, the last research on the topic with a focus on Europe was conducted as back as 2005. Moreover, the authors used a dataset with transactions announced during the years 1997-2000, i.e. during the very bullish market of the dot-com bubble with its sky-high stock market valuations.

The thesis is structured as follows: Chapter 2 provides chronologically ordered overview of the academic research done on the topic of choice of payment methods in M&A transactions. It is divided to two parts. The first covers theoretical frameworks, the second empirical research which has been so far done in the field. Chapter 3 covers the theoretical framework on which the thesis bases its analyses - the logit model and Bayesian model averaging. Chapter 4 describes the various hypotheses which are believed to govern the choice of payment in M&A transactions. Chapter 5 describes our data and finally, Chapter 6 presents core of this thesis - its empirical results. Chapter 7 summarizes thesis's findings and lays out ideas for a prospective research.

Chapter 2

Literature Review

2.1 Theoretical Models

Probably the very first who, from a theoretical perspective, touched upon the topic of how corporations decide what mode of payment to use in mergers & acquisitions (M&A) was Hansen (1987). Hansen constructed the very first theoretical model of the choice of exchange medium in M&A. The model is a two-agent bargaining game under imperfect information. In such a model, the well-known lemons problem arise when both the acquirer and target dispose of inside information regarding their company's value. This asymmetry explains why some companies might use their shares as a payment method - the stock trade offers a contingent-pricing effect because the risk of unsuccessful deal is bore by both acquirer's and target's shareholders. The model proposes several more testable hypotheses. First, the probability of stock payment should increase as the target company increases in size relative to the buyer. Next, by considering a debt financing, Hansen states that the probability of a stock trade increases with the acquirer's debt and decreases with the size of the target's debt. The choice of payment signals the beliefs about the companies' values, too. If the acquirer believes his company stock to be overvalued, he ought to use stock financing as this stock "currency" has a bigger firepower and *vice versa*.

Independently on Hansen, Fishman (1989), too, created a model describing the role of medium of exchange in M&A. Fishman analogically studies the choice of payment method under asymmetric information, nevertheless he focuses on the role of the exchange medium in preempting competitive bids.

In what the two authors differ is the assumed benefit of cash payments. In Hansen, bidder's offer cash if they perceive their equity to be undervalued. In Fishman, cash offers signal a high valuation for the target to deter competition.

Eckbo *et al.* (1990) propose another theoretical model of takeovers under asymmetric information. They expand on the Hansen's model and introduce a cash-security mix payment. The authors identify a separating equilibrium in which the value of the bidder is revealed by the mix of cash and stock in the payment for the target. The paper concludes that higher valued bidders use more cash financing in acquisitions.

2.2 Empirical Studies

The very first empirical studies dealing with M&A transactions (for instance Dietrich, Sorenson (1980), Harris *et al.* (1982), Stevens (1973)) did not distinguish between different payment methods but rather studied differences between acquired and non-acquired firms. This approach was challenged by Carleton *et al.* (1983), who estimated a logit model on M&A data from 1976 and 1977 and focused primarily on financial characteristics of targets and to some extent personal tax considerations. The study concluded that companies with high dividend payouts are likelier to be acquired by a stock payment; the opposite holds for market-to-book ratio. Nevertheless, the most important contribution of the paper was the revelation that cash and stock financed takeover have different determinants.

Niden (1986) studied tax issues in the US acquisitions. She investigated the choice between taxable (all-cash) and nontaxable (all-equity) forms of payments. She estimated a logit model with variables proxying for the tax position of acquirer and acquire. Contrary to Carleton *et al.* (1983), she concluded that there is no relationship between the tax status of the target's shareholders and the method of payment.

In a similar fashion to Niden (1986), Julian R. Franks (1988) examined the determinants of method of payments in M&A on a set of deals from UK and US over the years 1955-1985. They find no evidence that tax issues would have implications on the choice of the payment method. Alan J. Auerbach (1987)

too conclude that the tax issues have no statistically significant impact on the payment method choice.

Amihud *et al.* (1990) brought forward yet another determinant trying to explain the choice of payment method – they focused on the issue of maintaining a control over the bidding company. Managers and shareholders in general with a significant ownership stake in their company willing to maintain control over it will lean towards a cash financing which does not dilute their ownership. The results of their study confirm the hypothesis, i.e. managers with relatively higher equity stake in the bidder tend to prefer cash financed acquisitions.

Chaney *et al.* (1991) studied the relation between the medium of exchange and the characteristics of the acquiring firm, thus complementing the work of Carleton *et al.* (1983). Presumably, they found out that firms which pursue cash acquisitions differ in their financial characteristics to the ones who use stock financing. Acquirers using shares tend to be larger companies with low debt/equity ratio, low return on assets and high price-to-earnings multiple. Opposite holds for cash acquirers. Quite interesting is the resulting sign of the debt/equity variable. Chaney *et al.* (1991) results are contrary to the theory proposed by Hansen who argue that highly leveraged firms ought to finance their investments with equity. He argues that high leverage might be due to a specific preference of management for a debt financing and the cash method of payment is used to preserve such a capital structure.

Martin (1996) tested several hypotheses including the effect of investment opportunities, asymmetric information, the issue of control of voting rights by managers, the availability of cash for investments and bidder's leverage, the influence of institutional investors, whether the acquisition is in a form of a tender offer or a merger (in a tender offer the target's shareholders are approached directly whereas in a merger the acquirer first approaches the target's board of directors) and finally effects of a business cycles. The results are that acquirers with high growth opportunities, proxied by Tobin's Q-ratio, prefer to pay for investments in equity. The risk-sharing hypotheses, as presented by Hansen, was not proved. Martin offers an explanation that the firm size may be a poor indicator whether the acquirer wants to share the deal's risk with the target. Probably the biggest contribution of Martin is the examination of the control issue. Logically, managers with a low ownership stakes ought to be indifferent

to the dilution of control, i.e. using a stock payment, whereas those with higher stakes ought to take dilution into account. The cash availability hypothesis is confirmed, i.e. companies with high cash holdings tend to use this cash as a payment method. Interestingly, the leverage variable is always insignificant. Regarding outside monitoring hypotheses, the results have expected signs – companies with higher institutional shareholdings prefer using cash instead of stock so not to dilute the ownership. Finally, variables proxying for a business cycle effects were examined. The study suggest that growing stock market valuations pre-merger increase the use of equity financing due to reduced information asymmetry between managers and markets as explained in Choe *et al.* (1993). All in all, the author concluded that the two key factors are the mode of the acquisition (tender offer or merger) and the growth opportunities of the acquirer.

Ghosh & Ruland (1998) focus similarly as Amihud *et al.* (1990) on the issue of control rights, however from the point of view of target firm managers. Particularly, they examine the influence of the size of target's managers equity stake on the method of payment used by acquirer and on the job retention of target's managers after an acquisition. The results show that stock acquisitions are strongly related to high managerial ownership of target firms. Furthermore, the study suggest that the target's managerial ownership has even a larger influence than of the acquirer. Consequently, the bigger the stake of target's managers and the relative size of target to the deal size, the bigger the bargaining power of target's manager regarding the method of payment.

Heron & Lie (2002) investigated the relationship between the method of payment in acquisitions and, among other things, operating performance of the acquirers. On a large sample of US deals between the years 1985-1997, they conclude that method of payment has no influence on target's post-merger operational performance.

Officer (2004) narrowed his study on the effect of a "collar" in stock mergers. Collar is an instrument which allows for a change of the exchange ratio in stock transactions conditional on the change of the bidder's stock value. The collar protects the target's shareholders from the bidder's stock volatility between the merger announcement and its closing. In complex deals, the negotiations might drag on for several months during which the bidder's stock price

might significantly change. Officer concludes that the stronger the correlation between bidder's and target's stock returns, the smaller the probability of implementing collar into an M&A agreement.

Faccio & Masulis (2005) were the first to study the determinants of method of M&A payments on a sample of European bidders. They take into account a vast range of predictors representing the effects of corporate control, financial characteristics such as collateral, financial leverage and debt capacity of the bidder, asymmetric information as proposed in Hansen (relative deal size, over/under valuation of bidder/target), the type of the target (private or subsidiary being divested), the effect of cross/intra-industry deals and bidder's investment opportunities. They use data on acquisitions announced during the period of January 1997-December 2000 by bidders from 13 European developed economies. The authors found, like Ghosh & Ruland (1998), that the bidder has a strong preference for a cash financing if its controlling shareholder has a stake in the range of 20-60%. Contrary to Martin (1996), the authors conclude that the deal's relative size and bidder's leverage are important determinants of the method of payment used. Regarding the corporate control issue, authors found out that the managerial ownership relationship is linear if bidder was from continental Europe, but cubic, i.e. bidders reluctant to use equity over intermediate levels of voting control, if domiciled in the UK or Ireland.

Chemmanur *et al.* (2009) consider a situation where both the target and bidder have private information about their own intrinsic values. The authors use predictors regarding both the target's and bidder's under/over-valuation and proxies for asymmetric information about the parties such as analysts' coverage of their stocks. The authors conclude that acquirers using stock payment tend to be significantly overvalued and cash acquirers are correctly valued. Next, using a logit and ordered logit model, authors state that the probability of using a stock payment is linearly increasing with the extent of the overvaluation. Lastly, the greater the information asymmetry faced by the bidder in the valuation of the target, the likelier a cash offer. This is once again in conflict with the theory of Hansen (1987), however in line with Fishman (1989), i.e. the bidder, by the cash offer, signals a high private valuation of the target to a prospective competing bidder.

Harford *et al.* (2009) investigate relation of firm's deviation from its target cap-

ital structure and the exchange mediums it uses in pursuing large acquisitions (relative size of the target to bidder at least 20%). Since majority of cash payments in acquisitions is sourced from debt, highly leveraged firms are reluctant to use cash in acquisitions. In addition, firms ex-ante decide on the payment type based on the acquisition's expected effect on the bidder's post-acquisition capital structure. Uysal (2011) again examines the relation of deviation from a target capital structure and, among other things, the choice of payment method in acquisitions. His models suggest that over-leveraged firms are on average less likely to pursue an all-cash acquisition and tend to decrease the cash portion of their acquisition payment. On the other hand, under-leveraged firms do not necessarily offer a higher share of cash in acquisitions payments. In line with previous studies, the study concludes that larger firms are more likely to do all-cash acquisitions due to more stable cash-flows and higher debt capacity (Titman & Wessels (1988)). If more bidders are present in a deal, the likelihood of a cash offer is higher – this confirms the theoretical predictions of Fishman (1989). In addition, the paper supports the results of Martin (1996), i.e. firms with high growth opportunities are likelier to issue equity to finance acquisitions.

Boone *et al.* (2014) conducted a thorough study of M&A deals spanning four decades. They observed that the popularity of a particular method of payment in M&A (cash, stock and mixed offers) changes over time. Stock offers peaked during the 1990, probably due to the dot-com boom of the late 1990's, whereas cash payments have surged since and dominate the M&A landscape nowadays. The authors try to explain these trends with predictors linked again to tax effects, information asymmetry and contracting costs. Furthermore, they argue that mixed payments deserve a closer examination as a separate payment category. Regarding the factors influencing the method of payment, they find a supporting evidence for the adverse selection theory of Hansen (1987). In line with the previous studies, the authors find only a weak evidence of effects of taxation.

Chapter 3

Theoretical Framework

Given our research question - how companies decide on the type of payment in acquisitions - the most suitable econometric framework for our topic are so-called discrete choice models. Based on Greene (2002), page 663, we can divide discrete choice models into four categories. First, in binary choice models the dependent variable can take on two forms - usually zero and one. The numbers do not have any real meaning here, they are just labels for real world decisions. For instance, investigating whether individual is eligible for a loan or not, we would code the outcome as 0 if not and 1 if yes. Multinomial choice models are used when individuals face more than just two choices. Let's say we would try to model and predict results of a football team. The dependent variable could then take on three forms - 1, 2 and 3, labeling a win, draw and loss. Ordered choice models are employed when individual has preferences about the outcome variable. For example, modeling an outcome of a wine tasting competition might be a good example. Lastly, event counts are situations in which we are trying to analyze the count of number of occurrences.

In our case, i.e. the multinomial choice framework, two models, probit and logit, are often considered. The logit model has gained a wide popularity in the field of social sciences as the probit model is rather complicated Greene (2002). A third option is a so-called Tobit (Tobin (1958)) which uses a truncated dependent variable. A survey of previous literature concerned with the issue of type of payments used in M&A shows that researchers indeed prefer a logit model (Martin (1996); Chemmanur *et al.* (2009); Carleton *et al.* (1983)). In line with the previous research, we will use the method of multivariate logistic regression as well.

3.1 Logit Model

This section introduces the underlying logic behind the logistic regression model as described in Hosmer & Lemeshow (2000).

The key quantity investigated in every regression is the conditional mean, i.e. the mean of an outcome variable given some value of a predictor(s), noted as $E(Y|x)$. In a classic linear regression, the conditional mean can be expressed as an equation linear in x :

$$E(Y|x) = \beta_0 + \beta_1 * x \quad (3.1)$$

In such a case, the outcome variable can take on any value from $-\infty$ to $+\infty$. However, with dichotomous data where the dependent variable can take on only discrete values, the conditional mean $E(Y|x)$ must lie in the interval $[0,1]$. Thus, a linear regression approach ought not to be used as we risk obtaining nonsensical results such as a probability outside the interval $[0,1]$.¹ In order to overcome this issue, a logistic distribution for the conditional mean $E(Y|x)$ is used in a case of a dichotomous dependent variable. The form of a logistic regression model is then:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 * x}}{1 + e^{\beta_0 + \beta_1 * x}} \quad (3.2)$$

The quantity $\pi(x) = E(Y|x)$ is used to simplify the notation.

Furthermore, in a linear regression setup of $y = E(Y|x) + \varepsilon$ we assume the error term ε to be normally distributed with a zero mean and constant variance. However, in a logistic regression setup of $y = \pi(x) + \varepsilon$, the error term ε may take on two values - if $y = 1$ then $\varepsilon = 1 - \pi(x)$ with a probability of $\pi(x)$ and if $y = 0$ then $\varepsilon = -\pi(x)$ with a probability of $1 - \pi(x)$. Hence, the error term ε has a variance of $\pi(x)[1 - \pi(x)]$ and is no longer homoscedastic.

¹However, there are proponents of linear regression models in the framework of dichotomous dependent variable; for example see Hellevik (2007)

3.1.1 Logit Model Estimation

In a linear regression setting the most suitable estimation method is the well-known least squares technique. If all the required assumptions about the true regression model and the data generating process are met, so-called Gauss-Markov assumptions, then the ordinary least squares estimator (OLS) is the best linear unbiased estimator (BLUE). However, estimating a model with a dichotomous dependent variable causes certain issues, among them that the OLS estimator is no longer BLUE. Therefore, a maximum likelihood estimation is used in a framework of discrete choice models.

Maximum likelihood estimation (MLE) is a method which tries to estimate model's parameters in such a way that the parameters' values maximize the likelihood of observing the original data. MLE function is then a function which expresses the probability of the observed data as a function of the unknown parameters. The resulting MLE estimators of the model's parameters are such that they fit the originally observed values the most. The likelihood function, having a parameters vector of $\beta = (\beta_0, \beta_1)$, has the following form:

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (3.3)$$

The underlying logic of MLE is to take an estimate of β that maximizes equation (3.3). Usually, a log transformation of the equation (3.3) is used, i.e. the log-likelihood:

$$L(\beta) = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \quad (3.4)$$

We differentiate $L(\beta)$ w.r.t. β_0 and β_1 and obtain following equations:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (3.5)$$

and

$$\sum_{i=1}^n x_i [y_i - \pi(x_i)] = 0 \quad (3.6)$$

Since the equations (3.5) and (3.6) are not linear in β_0 and β_1 they are not

easily solvable. Nevertheless, major statistical software packages do the underlying heavy-lifting computations for the user.

Finally, the value of β resulting from solution of equations (3.5) and (3.6) is the sought maximum likelihood estimate.

3.1.2 Bayesian Model Averaging

Research on the M&A payment decisions tends to build models with quite a lot of covariates. For instance, Faccio & Masulis (2005) used 16 predictors, Martin (1996) incorporates 14 predictors into his model. All in all, our full model specification, as is presented in Chapter 4, includes 13 independent variables. Although the variables used have a clear economic rationale behind them, such a high number of explanatory variables without a doubt leads to a significant model uncertainty - with K predictors we have 2^K models to estimate. Several variable selection methods exist such as forward selection, backward elimination, stepwise regression or best subset regression method. For a discussion of this methods see for example Hastie *et al.* (2009). Nevertheless, a popular method called Bayesian model averaging (BMA) addresses the problem of model uncertainty well. A brief description of the method based on Zeugner & Feldkircher (2015) follows.

Let us assume a linear model structure:

$$Y = \alpha_\gamma + X_\gamma \beta_\gamma + \varepsilon; \quad \varepsilon \sim N(0, \sigma^2 I) \quad (3.7)$$

where Y is dependent variable, α_γ constant, β_γ coefficients and X_γ predictors.

As previously mentioned, if the model includes many potential predictors, the question is which of these variables should be included in the model. BMA estimates models with all possible combinations of X and computes a weighted average of all the models. The model weights for the averaging come from posterior model probabilities from Bayes theorem:

$$p(M_\gamma | y, X) = \frac{p(y|M_\gamma, X) p(M_\gamma)}{p(y|X)} = \frac{p(y|M_\gamma, X) p(M_\gamma)}{\sum_{s=1}^{2^K} p(y|M_s, X) p(M_s)} \quad (3.8)$$

The term $p(y|X)$ is constant over all models. Thus, the posterior model probability $p(M_\gamma|y, X)$ is proportional to the term $p(y|M_\gamma, X)$, i.e. the probability of the model given the model M_γ , times prior model probability $p(M_\gamma)$, that is how probable is the model M_γ given the researcher's beliefs. The rule of thumb is to set a uniform prior probability for each model.

Furthermore, as already noted in the thesis introduction, statistical inference grounded solely on P-values might be misleading due to their effect overestimation. In the setting of BMA, there is no single cut-off value on which researcher would decide whether to consider predictor to be statistically significant or insignificant. The algorithm of model averaging inherently takes care of coefficient sizes and standard errors and does not force the researcher to drop or include variables solely on variable's P-value.

Chapter 4

Methodology

This chapter discusses various hypotheses proposed by previous research done on the topic of choice of payment in M&A and sets out motivation for independent variables employed in our model. Factors influencing the choice of payment method can be summarized in several groups. First, we discuss financial factors, i.e. financial characteristics of the acquiring company. Next, we discuss factors related to corporate governance and ownership. Third subsection deals with information asymmetry factors. Last but not least, we discuss other factors such as legal frameworks and a type of ownership of target company.

4.1 Financing Decision Hypotheses

The choose of the mode of payment in M&A transactions is not arbitrary but is influenced by numerous factors such as financing constraints, corporate governance issues or information asymmetry. Furthermore, most of the factors are related to the bidder rather than the target company since should the deal terms be unacceptable to the bidder, the deal will be likely canceled or the bidder can make a hostile offer on its own terms (Faccio & Masulis (2005)).

4.1.1 Financial Factors

Financial factors are primarily related to the bidder's capacity to use cash respectively debt financing which in turn provides cash available for acquisitions. The pecking order theory of Myers (1984) states that companies prioritize the sources of investment financing where the internal resources have the largest priority followed by a debt financing and the least favorable option is a new

equity issuance. In addition, it has been proved that cash rich companies tend to use cash in investment decisions (Jensen & Posner (1986)). Last, but not least, it is a common sense that the higher the company's leverage, the smaller the probability of paying in cash and *vice versa*.

Our model assumes three variables reflecting the bidder's ability to pay in cash. First, *COLLATERAL* is proportion of the bidder's net property, plant and equipment to its total assets and is a proxy for the bidder's ability to take on debt. Second, *CASH* is the ratio of the bidder's cash and cash equivalents position to the particular transaction value. Finally, *LEVERAGE* shows the bidder's financial leverage measured as the ratio of its book value of debt to its market value of equity. All variables are as of the particular deal's announcement date.

4.1.2 Corporate Governance Factors

Corporate governance factors are at least equally important as the financing ones. It is reasonable to assume that control over a company is beneficial to its owners. If this holds, then controlling shareholders face an issue when their company finances acquisitions by stock, since in such a case their ownership is diluted and they should be reluctant to use stock as the payment method. Our model accounts for the control issue with the variable *CONTROL* – the percentage ownership of the bidder's largest shareholder.

Next, in stock financed acquisitions where the target's ownership structure is rather concentrated, and at the same time target's size is quite large relative to the acquirer's size, a new controlling blockholder can be created. This is reflected by the variable *CONTROLLOSS* which is the product between the target's largest ownership stake and the deal's relative size which is computed as a ratio of the deal's value to the sum of the bidder's market capitalization and the deal's value.

4.1.3 Information Asymmetry Factors

Information asymmetry factors play their role too. Hansen (1987) in his theoretical model states that stock financing is favorable due to its contingent-pricing effect when the information asymmetry about the target's value is large. This information asymmetry increases with the target's size and consequently

stock financing is preferable. We control for this by introducing the variable *RELSIZE* which is the ratio of the deal's value and the bidder's size proxied by its market capitalization.

Furthermore, Hansen (1987) predicts that, *ceteris paribus*, overvalued companies are inclined to use stock for financing acquisitions rather than cash as its stock has larger "firing power" and *vice versa*, i.e. undervalued companies should use cash to finance investments. However, the question is how to proxy for over- and undervaluation. The Q-ratio, originally devised by Kaldor (1966) and later popularized by Tobin & Brainard (1976) is a popular proxy for company misvaluation. Price-to-book ratio is yet another proxy. Nevertheless, at the same time both are often used as proxies for company's growth opportunities (Dong *et al.* (2003)). Given this ambivalence, we disregard both measures and use *RUNUP* – a percentage return of the bidder's share during the year prior to the deal's announcement date as a proxy for over- and undervaluation Faccio & Masulis (2005).

Above mentioned growth opportunities are yet another factor related to the information asymmetry hypothesis. Logically, a company facing high growth opportunities will prefer to rather finance itself by equity which gives it a larger headroom for investment decisions in contrast to debt financing which usually bears covenants and limits company's investment decisions. Indeed, Jung *et al.* (1996) prove this hypothesis. Once again, we face the problem of how to proxy for such unobservable growth opportunities. The most commonly used proxies for growth opportunities are the market-to-book assets ratio, market-to-book equity ratio, the earnings-price ratio and the ratio of capital expenditures over the book value of net property, plant and equipment (Adam & Goyal (2008)). They conclude that the market-to-book assets ratio has the highest explanatory power. We will hence use this ratio, in our model coded as *QRATIO*, as a proxy for the bidder's growth opportunities. Specifically, *QRATIO* is the ratio of the bidder's enterprise value to the book value of its assets. Enterprise value is defined as a sum of its market capitalization and market value of debt (proxied by its book value in line with the corporate finance practice) net of the firm's cash and cash equivalents. Moreover, similarly to Martin (1996), we use another proxy *REV_GROWTH* – a three year compounded annual growth rate (CAGR) of the bidder's revenues as of the acquisition announcement date. Where it was not possible to compute a three year CAGR, we computed a two

or one year revenue CAGR, respectively.

Next, information asymmetry is higher in diversification deals, where the bidder and target are from different industries. The dummy variable *INDUSTRY* takes on value 1 if the bidder and target are from the same industry as defined by Reuters Eikon and 0 otherwise.

Finally, it is correct to assume that cross-border deals bear higher information asymmetry than deals where both bidder and target are from the same country. The dummy variable *DOMESTIC* accounts for this phenomenon.

4.1.4 Other Factors

We consider two further factors. First, we introduce a dummy variable *COMMON_LAW*, which reflects the possible difference between deals with UK or Ireland domiciled bidders and bidders domiciled in continental Europe. Transactions with the former were most likely governed under a common law which prevails in the Anglo-Saxon world, whereas deals with continental Europe bidders were likely governed by a civil law which is applied in the continental Europe.

Second, we employ a dummy variable *PRIVATE_TARGET* to account for the fact that owners of privately held companies might to prefer immediate cash payments over stock considerations (Faccio & Masulis (2005)). This factor also accounts for a case where the target company is a subsidiary.

4.2 Full model specification

Based on the aforementioned factors, the full model used for our analysis has the following specification:

$$\begin{aligned} Y_i = & \beta_0 \\ & + \beta_1 CONTROL + \beta_2 CONTROLLOSS + \beta_3 LEVERAGE + \beta_4 CASH \\ & + \beta_5 COLLATERAL + \beta_6 QRATIO + \beta_7 REV_GROWTH \\ & + \beta_8 RELSIZE + \beta_9 PRIVATE_TARGET + \beta_{10} INDUSTRY \\ & + \beta_{11} COMMON_LAW + \beta_{12} RUNUP + \beta_{13} DOMESTIC \end{aligned} \tag{4.1}$$

where Y_i represents the payment method and is coded as 1 for cash, 2 for stock and 3 for a mixed payment type of both cash and stock.

Chapter 5

Data

5.1 Data Description

Our full sample includes M&A transactions sourced from Thomson Reuters Eikon database and its Deal Screener app. The selection criteria were following:

- i) the deal was either completed or withdrawn between 1/1/2010 and 12/12/2018 or was still pending as of 12/12/2018
- ii) the deal was either a merger or an acquisition
- iii) the deal value was over USD 10 million
- iv) the acquirer sought at least 5 % ownership stake in the target
- v) the deal was financed entirely in cash (including earnouts and assumption of liabilities), shares or mix of both
- vi) the bidder was domiciled in one of the 28 European Union countries (including the United Kingdom)
- vii) the bidder was publicly traded at the time of the transaction

No limitations were imposed on the target company. All in all, we retrieved 2 240 transactions. However, since some of the transactions are dating back several years, we were not able to retrieve all information regarding the bidders. Figure (5.1) visualizes N/As for each of our explanatory variables.

Our raw data set has in total 8 % of missing values, majority of them in connection with the variables *REV_GROWTH*, *LEVERAGE*, *COLLATERAL*,

Figure 5.1: Visualization of Missing Data Points



Table 5.1: Distribution of Transactions by Year and Payment

	N	Cash	Stock	Mix
<i>Total</i>	<i>1492</i>	<i>1099</i>	<i>175</i>	<i>218</i>
2018	184	136	15	33
2017	212	165	23	24
2016	191	152	17	22
2015	177	132	23	22
2014	161	113	23	25
2013	129	86	20	23
2012	135	104	16	15
2011	155	111	17	27
2010	148	100	21	27

CASH and *QRATIO*. The high proportion of missing values for these variables is caused by flaws in the Reuters database. We estimate that a rather large portion of this missing data points could be restored on a case by case basis by looking up the particular companies manually in the Reuters database, however this task would be excessively time consuming with such a large number of missing rows (e.g. *REV_GROWTH* has in total over 550 N/As). After cleaning the data, we end up with 1 492 transactions.

Table (5.1) presents transaction distribution by the year of announcement and by the particular method of payment used. Presumably, majority of transactions (1 099, 74 %) were financed solely in cash, followed by mixed financing (218, 15 %) and stock financing (185, 12 %). Regarding the announcement date, the distribution is quite even with a slight increase in annual number of transactions starting in 2014.

Table (5.2) displays the transaction distribution by country. The distribution of our M&A deals is heavily skewed towards the United Kingdom (UK), both on the bidder and the target side – 39 % of transactions include a bidder from the UK, whilst 24% include a target company from the UK. The rest is relatively well distributed among the rest of the countries. The distribution of a bidder's country more or less represents the degree to which the particular country's economy and financial markets are developed. From this point of view, Poland is quite a pleasant surprise with a number of deals exceeding, for instance, that of the Netherlands or Austria. The high proportion of UK deals can be likely explained by the substantial number of listed companies at the London Stock Exchange (LSE). The LSE, with over 3000 listings, is the second largest stock exchange in Europe only after the pan-European Euronext. Nevertheless, this phenomenon is nothing unprecedented as Faccio & Masulis (2005) report even higher proportion of UK transactions at 65.3 %.

Table (5.2) presents descriptive statistics of payment methods classified by country of the bidder, too. Worth mentioning is the high proportion (almost 40 %) of transactions including at least some sort of stock payment in deals with Finnish bidders (see Table (5.2)). Again, Faccio & Masulis (2005) report similar results. Nevertheless, they do not give any explanation for this phenomenon. On the other hand, transactions including bidders from Denmark and Luxembourg were predominantly cash financed as were transactions with

Table 5.2: Country Distribution of Transactions

Country	Bidders					Targets	
	N	%	Cash	Stock	Mix	N	%
United Kingdom	583	39%	70%	11%	19%	356	24%
France	197	13%	72%	15%	39%	125	8%
Sweden	171	11%	70%	9%	22%	78	5%
Germany	104	7%	82%	11%	8%	80	5%
Italy	80	5%	80%	14%	6%	70	5%
Spain	65	4%	78%	14%	8%	53	4%
Poland	63	4%	83%	17%	0%	54	4%
Republic of Ireland	51	3%	82%	10%	8%	12	<1%
Finland	36	2%	61%	22%	17%	31	2%
Netherlands	36	2%	75%	19%	6%	41	3%
Belgium	32	2%	81%	6%	13%	24	2%
Luxembourg	15	1%	87%	7%	7%	9	<1%
Austria	15	1%	80%	7%	13%	14	<1%
Denmark	14	<1%	83%	0%	0%	20	1%
Greece	6	<1%	33%	33%	33%	4	<1%
Malta	6	<1%	50%	0%	50%	9	<1%
Portugal	6	<1%	83%	17%	0%	10	<1%
Cyprus	6	<1%	83%	0%	17%	1	<1%
Czech Republic	2	<1%	100%	0%	0%	7	<1%
Croatia	1	<1%	100%	0%	0%	0	<1%
Latvia	1	<1%	100%	0%	0%	2	<1%
Lithuania	1	<1%	100%	0%	0%	4	<1%
Hungary	1	<1%	100%	0%	0%	2	<1%
USA						206	14%
Norway						23	2%
Canada						28	2%
Australia						29	2%
India						20	1%
Brazil						17	1%
Switzerland						17	1%
South Africa						14	<1%
Russia						13	<1%
China						13	<1%
Other						106	7%
<i>Total</i>	<i>1492</i>					<i>1492</i>	

bidders from Poland, Germany, Italy, Ireland, Belgium and Austria, although to a smaller extent.

Table (5.3) presents averages and medians of our explanatory variables together with transaction deal value categorized the payment method. It is obvious that bidder and deal specific attributes differ in some cases quite substantially. Transactions financed by stocks or mixed payment are, on average, distinctively larger than cash financed transactions. Again, transactions analyzed by Faccio & Masulis (2005) display similar characteristics. At the same time, median of the deal value is well below the average for each payment method, in other words the deal value distribution is heavily skewed to the right with a number of large transactions. The variable *RELSIZE* is dramatically larger for both stock and mix financed transactions, too. This might be due to financial constraints on the bidder's side as well as higher information asymmetry as the size of the deal value increases relative to the bidder's size. Bidders using stock, be it partially or entirely, to finance transactions exhibit on average higher revenue growth than bidders using cash, although the median is same to a greater or lesser degree. This, again, means that the distribution is skewed to the right side, i.e. there are some outliers with a very high revenue growth using stock and mix financing. This would support the hypotheses that high growth companies tend to use stock in the M&A financing. On the other hand, the average and median value for *QRATIO* is quite close for the cash and stock financed deals. The very large average value of *QRATIO* in mix transactions is caused by a single one outlier value of 1 818. The average value of *CONTROL_LOSS* is larger for stock and mix financed transactions. This might be resulting from the relatively larger stock and mix transactions. The statistics for *LEVERAGE* would agree with the assumption that companies with a high leverage and, consequently, used debt capacity are prone to issue equity when financing investments. Last but not least and in line with a common sense and previous research, both the average and median value of *CASH* is drastically higher for cash financed deals.

A quick glance at the binary independent variables reveals that the statistics agree with the theory of information asymmetry that domestic transactions and transactions where both the bidder and the target are from the same industry are more likely to be financed at least in part by the bidder's stocks. On the other hand, transactions where the target company is privately held

are predominantly financed by cash. Faccio & Masulis (2005) report similar results. Finally, transactions where the bidder was domiciled either in the UK or Ireland, i.e. countries with a common law, exhibit higher share of mixed payments. Histograms of independent variables classified by the payment method are available in Appendix A.

Table 5.3: Descriptive Statistics of Independent Variables

	Cash	Stock	Mix
Averages and medians (<i>italics</i>) for continuous variables			
Deal Value (mUSD)	514	1 835	1 065
	<i>57</i>	<i>184</i>	<i>45</i>
RELSIZE	0.10	0.35	0.26
	<i>0.05</i>	<i>0.36</i>	<i>0.20</i>
REV_GROWTH	0.11	0.81	0.29
	<i>0.04</i>	<i>0.03</i>	<i>0.08</i>
RUNUP	0.26	0.29	0.34
	<i>0.18</i>	<i>0.16</i>	<i>0.19</i>
QRATIO	1.50	1.64	10.70
	<i>1.07</i>	<i>0.98</i>	<i>1.25</i>
CONTROL_LOSS	0.08	0.21	0.21
	<i>0.04</i>	<i>0.07</i>	<i>0.13</i>
CONTROL	0.24	0.24	0.21
	<i>0.16</i>	<i>0.19</i>	<i>0.16</i>
LEVERAGE	0.40	0.70	0.45
	<i>0.20</i>	<i>0.24</i>	<i>0.10</i>
COLLATERAL	0.18	0.20	0.15
	<i>0.12</i>	<i>0.10</i>	<i>0.06</i>
CASH	12.82	1.31	1.33
	<i>1.73</i>	<i>0.22</i>	<i>0.30</i>
Averages for binary variables			
DOMESTIC	0.42	0.55	0.54
INDUSTRY	0.63	0.72	0.74
PRIVATE_TARGET	0.79	0.42	0.78
COMMON_LAW	0.41	0.38	0.53

Chapter 6

Empirical results

In this section, we analyze the actual inferences our data provide us with. The section starts with the specification of a full logit model including all 13 independent variables. Next, we employ several model selection techniques and a Bayesian Model Averaging (BMA) in order to arrive at an optimal model specification given our relatively high number of independent variables. We continue with a comparison of our regression results with those of the previous research on the topic. We conclude our empirical research by simplifying the multinomial logit to a binary one and compare its results with that of the multinomial logit.

6.1 Full Model Results

First, we specify the full logit model as stated in the Equation (4.1) with the reference dependent variable being *cash*.

The results of the model are shown in Table (6.1). First, we ought to note that multinomial logit estimates $k - 1$ equations, where k is the number of dependent variables. Therefore, Table (6.1) has two columns with results – one for the *mix* equation, one for the *stock* equation.

First, the regressors *LEVERAGE*, *RUNUP* and *COMMON_LAW* are not statistically significant. Variables *QRATIO*, *REV_GROWTH*, *REL_SIZE*, *INDUSTRY*, *DOMESTIC* and the constant are statistically significant regardless whether we analyze *stock* or *mix* transactions. The rest of the variables is partially statistically significant depending whether we analyze the *stock* or

the *mix* equation. The full model's Akaike Information Criterion (AIC) is 1 778. AIC tells us how well regression model fits data relative to other models, i.e. it is a measure of relative goodness of fit. As mentioned previously, our model includes 13 independent variables meaning we have a universe of 13^2 , i.e. 8 192, models to estimate and the full model including all variables does not necessarily fit our data the best. In the next section, we will hence employ several model selection techniques and model averaging to arrive at the "best" model.

6.2 Model Selection

All models are wrong, but some are useful.

George Box

Determination of the final "best" model used for an analysis involves two opposing objectives. One is to find a model with as many predictors which would reliably describe and fit our data. The goal is basically to find all variables related to our research problem. On the other hand, in the sake of parsimony (think of Occam's razor), we should aim to include as few regressors as possible. Therefore, we need to find the right balance between simplicity and fit. Once we have selected the possible set of relevant independent variables, we need to define a criterion on which basis we will compare competing models. Some of the most well-known criteria are R-squared, AIC and Mallows's C_p , to name a few. However, a new issue arises – if we have p regressors, we end up with 2^p possible models to choose from, as stated previously. Fortunately, statisticians have at their disposal several methods which automatically allow to search through all possible model specifications and on a basis of some criterion select a model deemed to be the "best" one.

A wide array of model selection methods exists, including the well-known forward/backward/stepwise strategies, which are generally known as automatic variable selection methods. These algorithms are one side quite simple to comprehend, on the other hand their usefulness have been questioned (Thompson (1995), Harrell (2013), Tutz *et al.* (2015)). Further and more sophisticated methods are available such as best subset regression, LASSO and ridge regressions. Last but not last, a modern approach to model selection are model

averaging methods including BMA.

We proceed in a following manner. First, we will use the basic automatic methods. Then, we will employ the best subset regression method. We conclude the optimal model search with the BMA.

6.2.1 Automatic Methods

Automatic methods encompass three specific model selection techniques – forward selection, backward elimination and stepwise regression. These methods are among the most popular variable selection techniques due to their simplicity and straightforwardness. As commonly employed, these methods allow the entry of predictors one step at a time and at each step the removal of previously entered variables is also considered. On the other hand, a vast array of research challenges their accuracy (Thompson (1989), Hastie *et al.* (2009)).

Backward elimination starts with a fully specified model, i.e. a model with all predictors which are either chosen according to a theory, analyst’s intuition or data. First, we need to set a significance level according to which we decide whether to drop or keep a predictor in our model, a so-called Alpha-to-Remove significance level denoted α_R which is typically higher than the usual 0.05 and is set at 0.15 (Bendel & Afifi (1977)). Then we examine the predictors with a t-test P-value above 0.15 and drop the one with the highest P-value. Again, we refit our model and repeat the previous step until all predictors’ P-values are below α_R .

Forward selection is, basically, a reversed backward elimination. The forward selection starts only with an intercept and gradually adds predictors with the lowest t-test P-values until none of the remaining variables shows a P-value lower than the pre defined cut-off value (again, typically 0.15).

Stepwise regression combines both aforementioned approaches. In the stepwise regression, we again specify significance levels used in the decision whether to include (remove) a predictor to (from) a stepwise model. These values are so-called Alpha-to-Enter and once again Alpha-to-Remove significance levels denoted α_E and α_R . As is the case with backward elimination, typically these are greater than the usual cutoff P-values and often are set at the level of 0.15

(Bendel & Afifi (1977)). Once we have set up α_E and α_R , we regress each of our predictors one by one on the dependent variable. The predictor with the lowest t-test P-value is selected as the first variable entering our model, in our case let it be x_1 . Then we analogically regress y on x_1 and x_2 , x_1 and x_3 and so forth. Again, predictor with the lowest t-test P-value is included in the model. If no predictor has a P-value lower than our cutoff value of 0.15, we stop the process and use the one-variable model. Nevertheless, let's assume x_2 was chosen. Now we need to evaluate whether entering x_2 into our model somehow alters the significance of x_1 . If the t-test P-value for testing $\beta_1 = 0$ is greater than 0.15, we disregard x_1 from the model. We repeat above described process until an additional predictor results in a t-test P-value higher than 0.15. All this being said, it is important to note that the stepwise method does not necessary lead to a selection of an optimal model, as it depends on selecting a specific criterion used in the model assessment.

Let us now proceed to the actual regression analysis. First, we start with backward elimination method followed by forward selection and conclude with stepwise regression. Unfortunately, R does not allow user to use a function which would perform the stepwise methods based on the α_E and α_R parameters, instead an R function called `step` selects the appropriate model based on the AIC. Results of the backward elimination applied to our data and model are in Table (6.1) which compares the full model to that resulting from the backward elimination. The method has eliminated 3 variables, namely *LEVERAGE*, *RUNUP* and *COMMON_LAW*, i.e. those variables which are statistically insignificant according to the full model. Removal of these 3 variables leads to a drop in the AIC to 1 769 compared to the full model's AIC of 1 778, i.e. this model fits our data better than the full one. Moreover, we see that the backward elimination has not led to any alteration of the coefficient signs and the coefficient values have changed only slightly.

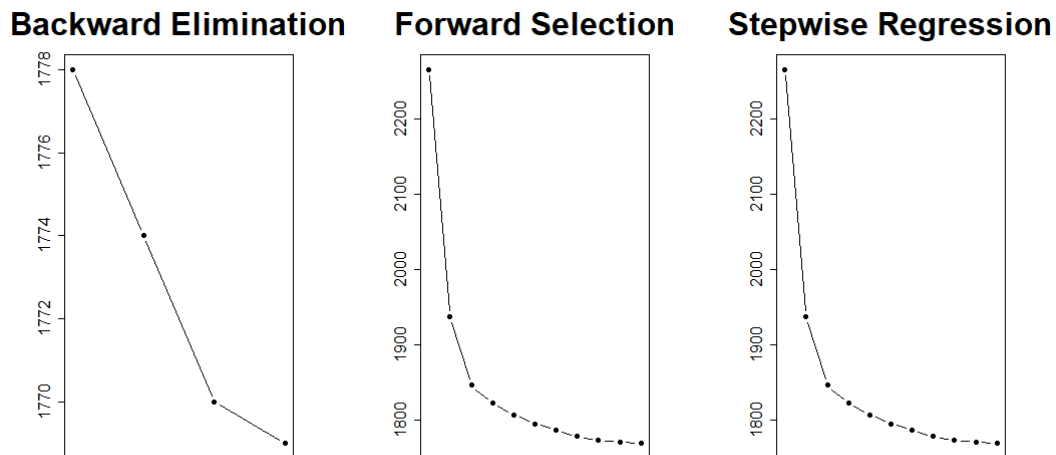
Table 6.1: Comparison of the Full Model & the Backward Elimination Model

	<i>Dependent variable:</i>			
	Full model		Backward elimination	
	mix	stock	mix	stock
COLLATERAL	-0.771* (0.436)	0.498 (0.459)	-0.746* (0.434)	0.448 (0.451)
CASH	-0.068*** (0.025)	-0.022 (0.018)	-0.069*** (0.025)	-0.022 (0.018)
LEVERAGE	0.033 (0.079)	0.013 (0.076)		
RUNUP	0.005 (0.152)	0.063 (0.174)		
QRATIO	0.110*** (0.041)	0.104** (0.046)	0.111*** (0.036)	0.107*** (0.037)
REV_GROWTH	0.191** (0.083)	0.256*** (0.088)	0.193** (0.082)	0.254*** (0.086)
CONTROL	-1.078** (0.493)	-0.096 (0.540)	-1.171** (0.479)	0.020 (0.524)
CONTROL_LOSS	-3.355** (1.442)	-2.016 (1.326)	-3.446** (1.432)	-1.909 (1.315)
RELSIZE	7.280*** (1.355)	8.460*** (1.189)	7.387*** (1.349)	8.357*** (1.180)
INDUSTRY	0.636*** (0.185)	0.437** (0.221)	0.645*** (0.185)	0.423* (0.220)
DOMESTIC	0.480*** (0.166)	0.637*** (0.202)	0.496*** (0.165)	0.631*** (0.201)
PRIVATE_TARGET	0.429 (0.316)	-1.667*** (0.314)	0.465 (0.311)	-1.710*** (0.309)
COMMON_LAW	0.176 (0.170)	-0.211 (0.213)		
Constant	-3.129*** (0.379)	-3.021*** (0.362)	-3.061*** (0.370)	-3.062*** (0.354)
Akaike Inf. Crit.	1,777.968	1,777.968	1,769.197	1,769.197

Note:

*p<0.1; **p<0.05; ***p<0.01

Figure 6.1: AIC Path of the Automatic Algorithms



Next, we employ forward selection and stepwise regression methods. Even though it is not necessarily always the case, both the forward selection and stepwise regression methods give, in our case, the same results (same AIC, same independent variables, same coefficients and standard errors) as the backward elimination method (summary table of the models is available in Appendix B). Figure (6.1) presents AIC path for each of the methods. Backward elimination, starting from the full model specification, gradually eliminated the three aforementioned regressors to settle down at the AIC of 1 769. On the other hand, forward selection started with the basic model which included only the intercept and was gradually adding variables. Interestingly, the AIC path for stepwise regression is identical to that of a forward selection.

6.2.2 All Possible Regressions Approach

The superiority of the all possible regression approach (alternatively called best subset regression) over automatic methods is due to their different set-up. Whereas in stepwise methods successive models are limited by variables already in the model from previous steps of the analysis, the best subset regression approach provides analysis of certain statistical criteria (e.g., AIC, Adjusted R^2 , Mallows' C_p) computed for every possible model of every size (Weisberg (1985)).

Let us now look at the results of our analysis, which was conducted with the help of R package `glmulti` (Calcagno & de Mazancourt (2010)). This package al-

allows user to conduct the all possible regression approach with generalized linear models. The criterion used for the model selection was AIC. After running the analysis, the package selected 100 "best" models ranked by their AIC. Figure (6.2) presents AIC for the "best" 100 models. The AIC of the "best" model is, once again, 1 769 and its specification is the same as was the case with the automatic methods mentioned above, i.e. we drop the variables *LEVERAGE*, *RUNUP* and *COMMON_LAW* from the full model. Nevertheless, there are two other models which are within 2 AIC units of the very "best" model (see again Figure (6.2); the horizontal line separates models which are more than 2 AIC units away from the "best" model). This cut-off is somewhat arbitrary though and critiques of this rule exist (e.g. Anderson (2007)). A better approach is to compare the models not only on the ground of absolute values of AIC (or other information criterion for that matter), but by analyzing the relative values of a selected information criterion. In our case, we look at the so-called Akaike weights of the "best" 10 models. Akaike weight for a particular model can be regarded as the probability that the model is the "best" model in a Kullback–Leibler sense of minimizing the loss of information when approximating full reality by a fitted model (Viechtbauer (2018)). Table (6.2) presents the "best" 10 models and their AIC and AIC weights.

Figure 6.2: AIC Profile

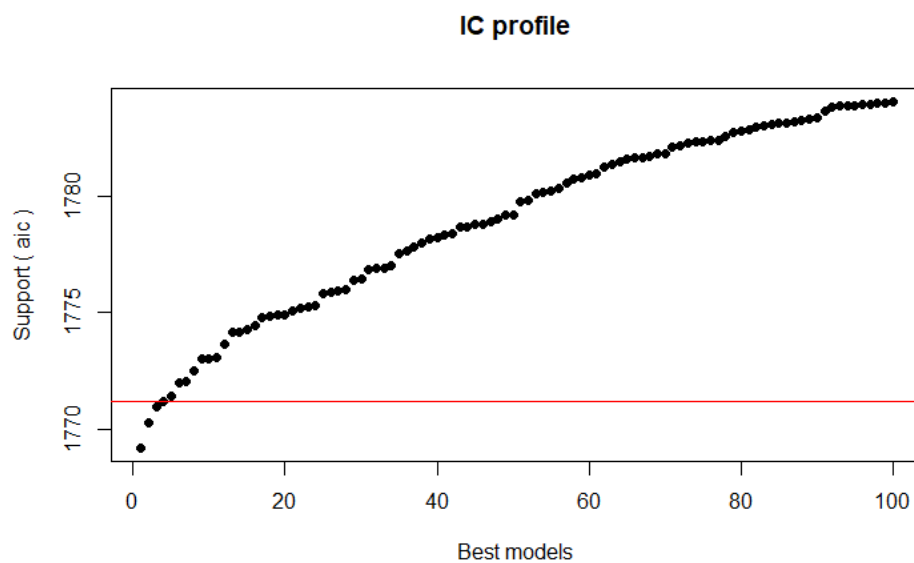


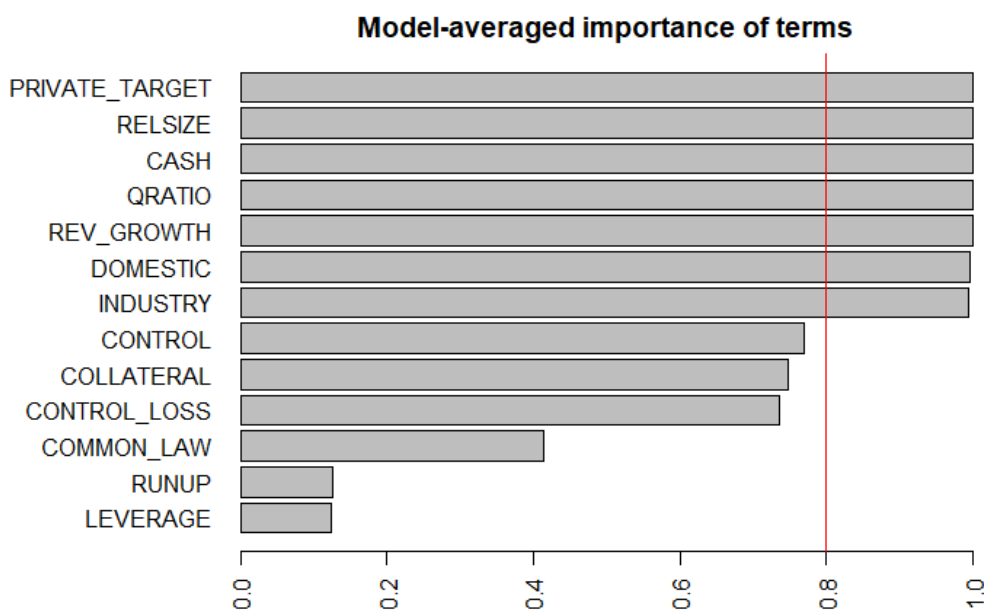
Table 6.2: Best 10 Models Selected by the Best Subset Regression

Mod	REL	REV	RUN	Q	C_L	CON	LEV	COLL	CASH	PRIV	DOM	INDS	COMM	AIC	AIC weight
1	•	•	•	•	•	•	•	•	•	•	•	•	•	1769.197	0.193
2	•	•	•	•	•	•	•	•	•	•	•	•	•	1770.312	0.110
3	•	•	•	•	•	•	•	•	•	•	•	•	•	1770.975	0.079
4	•	•	•	•	•	•	•	•	•	•	•	•	•	1771.216	0.070
5	•	•	•	•	•	•	•	•	•	•	•	•	•	1771.402	0.064
6	•	•	•	•	•	•	•	•	•	•	•	•	•	1771.987	0.048
7	•	•	•	•	•	•	•	•	•	•	•	•	•	1772.039	0.046
8	•	•	•	•	•	•	•	•	•	•	•	•	•	1772.517	0.037
9	•	•	•	•	•	•	•	•	•	•	•	•	•	1773.042	0.028
10	•	•	•	•	•	•	•	•	•	•	•	•	•	1773.045	0.028

Note: Variables' names were abbreviated to save space. REL is REL_SIZE, REV is REV_GROWTH, RUN is RUNUP, Q is QRATIO, C_L is CONTROL_LOSS, CON is CONTROL, LEV is LEVERAGE, COLL is COLLATERAL, PRIV is PRIVATE_TARGET, DOM is DOMESTIC, INDS is INDUSTRY, COMM is COMMON_LAW.

We see that models 2,3 and 4 are within the range of two AIC units from the "best" model and model 5 is very close, too. We can hence conclude that these models are basically equal as to their plausibility. However, if we closer investigate the AIC weights, we see that the first model has by far the largest AIC weight. The result thus suggests that model 1 is clearly superior above the rest in terms of AIC weight. Regarding parsimony, the first four models include 10, 11, 9 and 9 predictors, respectively. Model 2 is thus out of discussion as it simultaneously less parsimonious and has a lower AIC and AIC weight than model 1. Moreover, it includes the variable *COMMON_LAW*. *Ceteris paribus*, the fact that a company operates in a country with a common law should probably not affect its decision how to pay for an acquisition. We are then left with three "best" models and the trade-off between parsimony and fit as the model 3 and 4 have only 9 predictors, however a higher AIC and less than a half of the AIC weight compared to model 1.

Figure 6.3: Model-averaged Importance of Terms



Yet another method how to assess variables' importance is to analyze their relative importance as shown at Figure (6.3). The importance value for a particular predictor is equal to the sum of the AIC weights for the models in which the variable appears, i.e. variables which show up most often through all the investigated models receive a high importance value (Viechtbauer, 2018). Here, the cut-off value is set at 0.8, as demonstrated by the vertical line. Again,

this is somewhat arbitrary number which should be taken with a grain of salt (Viechtbauer, 2018). As the Figure (6.3) suggests, seven variables are above the relative importance cutoff value, three are slightly below and three variables are clearly out of discussion. The model-averaged importance of terms hence proposes yet another "best" model specification. Nevertheless, the variables *CONTROL*, *COLLATERAL* and *CONTROL_LOSS* are only slightly below the cut-off value of 0.8 and moreover all three are part of the two "best" models as shown in Table (6.2).

Recall, that the model proposed by automatic methods dropped the variables *RUNUP*, *LEVERAGE* and *COMMON_LAW* from the full model (Table (6.1)). This model specification is exactly the same as of the "best" model proposed by the best subset regression here. On the other hand, best subset regression gives a broader picture as to the model uncertainty – recall that four models are in the range of 2 AIC units. In the next section we will employ the BMA which addresses this model uncertainty by averaging over all possible models.

6.2.3 Bayesian Model Averaging

Statistical tests and P-values in large sample sizes can give unsatisfactory and somewhat misleading results (Raftery (1995), Lantz (2012)). The issue with P-values in large samples is that they tend to reject null-hypotheses despite the fact that the null model seems both theoretically and empirically right (Raftery (1995)). Further, in regressions with many independent variables, the standard model selection techniques (automated methods, subset regression) can give misleading results. Moreover, by selecting a single model, the analyst ignores the inherent model uncertainty (Raftery (1995)). A narrative describing the aforementioned issues follows as found in Hoeting *et al.* (1999).

Let's assume researcher has gathered data concerning a cancer of esophagus. The researcher has a large data set with a variety of possible explanatory variables at hand. Her goal is to determine which covariates are statistically significant and to define a model which would do well in future predictions of patients' survival time (i.e. Cox Regression). She then conducts a data-driven regression analysis and arrives at a "true" model M . She checks that M fits the data well and concludes the resulting parameter estimations sensible. She

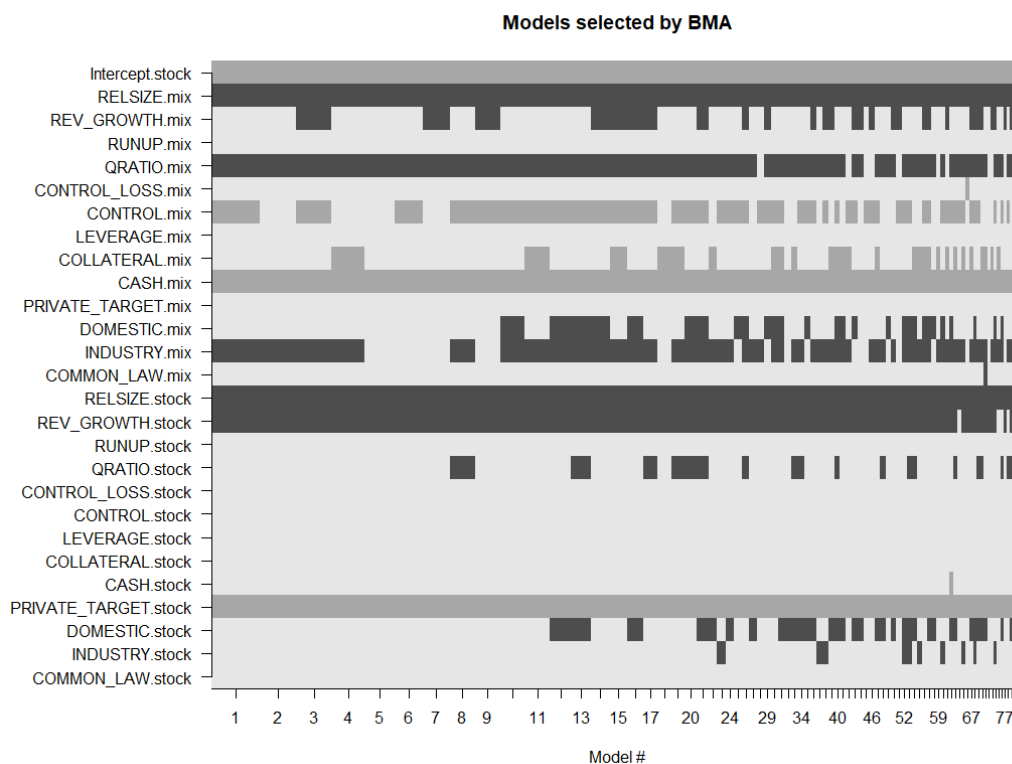
then uses the model M for further predictions. However, what if there exist a competing model M^* that fits the data well too, but has different parameter estimations, standard errors and/or makes different predictions? This is when BMA comes into play. A more interested reader can consult Raftery (1995), Raftery *et al.* (1997), Hoeting *et al.* (1999) or a more recent works of Leuven *et al.* (2008) and Eicher *et al.* (2009).

As mentioned previously, a model with K explanatory variables has in total 2^K possible combinations to be estimated. BMA takes either all such possible combinations (in this way being similar to the all possible regression approach) or, depending on the specific method and R package used, uses Markov chain Monte Carlo approximation (BMS R package) or Occam's window method of Madigan & Raftery (1994) (BMA and mlogitBMA R packages). In the next step, BMA computes weighted averages of the estimated coefficients (posterior means) across all the estimated models using posterior model probabilities (similar to information criteria in a frequentist approach) as weights (Havranek *et al.* (2018)). BMA thus does not conduct a classic variable selection as the above used automatic and all regression subset methods. On the contrary, it keeps all predictors in the model and averages over all possible sets of predictors (Hoeting *et al.* (1999)).

As mentioned, several R packages which implement BMA exist. The most popular ones are BMA(Raftery *et al.* (2018)) and BMS (Zeugner & Feldkircher (2015)). BMS can handle only linear regression models, while BMA handles generalized linear models too, however, not a multinomial logit regression. Fortunately, a package mlogitBMA (Sevcikova & Raftery (2013)) is specifically designed for a BMA analysis of multinomial logit models as it converts a multinomial response variable to a binary one according to Begg & Gray approximation. For a greater detail on the Begg & Gray approximation see Ševčíková & Raftery (2012). mlogitBMA is based on the BMA package and as such it assumes that all prospective models have same prior weights, specifically $1/2^K$ where K is the total number of covariates. With our 13 independent variables the prior weights are thus negligible. Second, unlike the BMS package, BMA and mlogitBMA do not employ Zellner's g-priors for the choice of prior distribution for the regression coefficients. Instead, BMA and mlogitBMA use a BIC approximation. Nevertheless, the BIC approximation behaves closely to a unit information prior given by the BMS. For a comparison of various R packages

focused on the BMA see M. & F. (2011).

Figure 6.4: Model Inclusion in Bayesian Model Averaging



Note: The figure depicts the results of Bayesian model averaging. On the vertical axis, the explanatory variables are grouped by the *mix* and *stock* equation. The horizontal axis shows the values of the cumulative posterior model probability. Dark color means the estimated parameter of a corresponding explanatory variable is positive. The lighter color means the estimated parameter of a corresponding explanatory variable is negative. No color means the corresponding explanatory variable is not included in the model. Note that the signs are robustly consistent across different models. Numerical results are reported in Table (6.4). Based on Havranek *et al.* (2018)

Let us now discuss the actual results of our BMA analysis. The BMA procedure selected 83 suitable models in total. Figure (6.4) demonstrates the frequency of variables as selected by the various models. Different colors represent positive and negative coefficient signs. All coefficient signs are stable among all the 83 models. Next, we see that some variables are present in all models, whilst some are not chosen by the BMA at all. Table (6.3) lists five "best" models out of these 83.

Table 6.3: Best 5 Models Selected by BMA

Model	REL	REV	RUN	Q	C_L	CON	LEV	COLL	CASH	PRIV	DOM	INDS	COMM	PMP
1	stock	•	•						•					0.060
	mix	•		•		•			•			•		
2	stock	•	•						•					0.044
	mix	•		•					•			•		
3	stock	•	•						•					0.044
	mix	•	•			•			•			•		
4	stock	•	•						•					0.041
	mix	•		•					•	•		•		
5	stock	•	•						•					0.037
	mix	•		•					•					
$Pr_{M4}[\beta_i \neq 0]$	stock	1	0.33	0.00	0.94	0.00	0.00	0.00	0.00	0.01	1	0.29	0.76	0.00
	mix	1	0.98	0.00	0.20	0.01	0.68	0.00	0.26	1	0.00	0.29	0.06	0.00

Note: Variables' names were abbreviated to save space. REL is REL_SIZE, REV is REV_GROWTH, RUN is RUNUP, Q is QRATIO, C_L is CONTROL_LOSS, CON is CONTROL, LEV is LEVERAGE, COLL is COLLATERAL, PRIV is PRIVATE_TARGET, DOM is DOMESTIC, INDS is INDUSTRY, COMM is COMMON_LAW.

Here, the ranking criterion is not an information criterion such as AIC or BIC, but posterior model probability (PMP). The PMP value of model 1 indicates that this model represents only 6 % of the total posterior probability. In total, the five best models have a total posterior probability of only 22.6 %, i.e. almost 80 % of posterior probability is spread among the other 78 models and, consequently, model uncertainty is of an issue.

Table 6.4: BMA estimates

	Mix			Stock		
	$P(\beta \neq 0 D)$	Mean βD	SD βD	$P(\beta \neq 0 D)$	Mean βD	SD βD
RELSIZE	100.00	4.37	0.47	100.00	6.97	0.44
REV_GROWTH	32.90	0.07	0.11	97.70	0.19	0.11
RUNUP	0.00	0.00	0.00	0.00	0.00	0.00
QRATIO	93.70	0.08	0.04	19.60	0.02	0.05
CONTROL_LOSS	0.80	-0.01	0.14	0.00	0.00	0.00
CONTROL	67.60	-0.84	0.70	0.00	0.00	0.00
LEVERAGE	0.00	0.00	0.00	0.00	0.00	0.00
COLLATERAL	26.20	-0.26	0.48	0.00	0.00	0.00
CASH	100.00	-0.07	0.02	0.50	0.00	0.00
PRIVATE_TARGET	0.00	0.00	0.00	100.00	-1.90	0.20
DOMESTIC	29.30	0.13	0.22	28.80	0.14	0.24
INDUSTRY	75.80	0.39	0.27	6.30	0.03	0.12
COMMON_LAW	0.40	0.00	0.03	0.00	0.00	0.00

Note: Probabilities are in percent.

Table (6.4) presents posterior means, standard deviations and posterior effect probabilities, $P(\beta \neq 0|D)$, for each variable coefficient (this is equivalent to the last row of Table (6.3)). Posterior effect probability indicates the probability that the particular variable's coefficient is not equal to zero among the 83 models selected by the BMA. Furthermore, the parameter estimates and standard deviations incorporate model uncertainty. The lower the posterior effect probability, the more BMA shrinks this parameter estimate towards zero and simultaneously tends to increase the standard deviation, thus taking into account model uncertainty (Hoeting *et al.* (1999)).

Let us now have a look at the results presented in the Table (6.4). Left part of the table presents results for mix financed transactions, right for stock financed. The column with posterior effect probabilities nicely demonstrates the

model uncertainty as some variables are highly significant for one type of transaction and at the same time having a small effect or no effect whatsoever for the other type. Moreover, this posterior effect probabilities correspond with the results presented in the Figure (6.4). For instance, the the only variable being significant for both equations is *RELSIZE*, the ratio of deal value to a sum of deal value and bidder's market capitalization, as is obvious both from the posterior effect probabilities in Table (6.4) as well as from the Figure (6.4). Next, variables *CASH*, *QRATIO* and, to a smaller degree, *INDUSTRY* and *CONTROL* have some effect in the case of the stock equation. Turning to the mix equation, apart from *RELSIZE*, only variables *REV_GROWTH* and *PRIVATE_TARGET* have posterior effect probabilities worth mentioning.

Finally, Figure (C.1) in Appendix C presents yet another way of variable's evidence for an effect. It depicts the posterior distribution of each coefficient, in other words frequency plots for each variable coefficient given by the 83 models selected by the BMA. The black vertical line is actually $1 - P(\beta \neq 0|D)$, i.e. the probability that the particular coefficient *is* zero.

6.2.4 Model Selection - Summary

Each of the automatic methods, i.e. forward selection, backward elimination and stepwise regression, yielded the same final model dropping variables *LEVERAGE*, *RUNUP* and *COMMON_LAW* from the full model. This reduction of predictors led to a more parsimonious model and a decrease in AIC to 1 769 from the full model AIC of 1 778. Next, the best subset regression, which searched through the entire model universe of 8 192 possible models, arrived at the very same model specification as the automatic methods and, hence, the same coefficients and AIC value. Nevertheless, the results of the best subset regression demonstrated that focusing on a single "best" model might be misleading since three other models were within the difference of two AIC units from the "best" model's AIC. These two methods thus gave us a somewhat blurry picture as to which final model to select. In such a situation, BMA proves to be useful. Rather than selecting a "best" model, which moreover in the case of BMA has the posterior model probability of mere 6 %, BMA averages over all models. Moreover, BMA allows us to make statistical inference based not on P-values, which are said to overestimate evidence (Edwards *et al.* (1963), Berger & Delampady (1987), Berger & Sellke (1987)), but rather

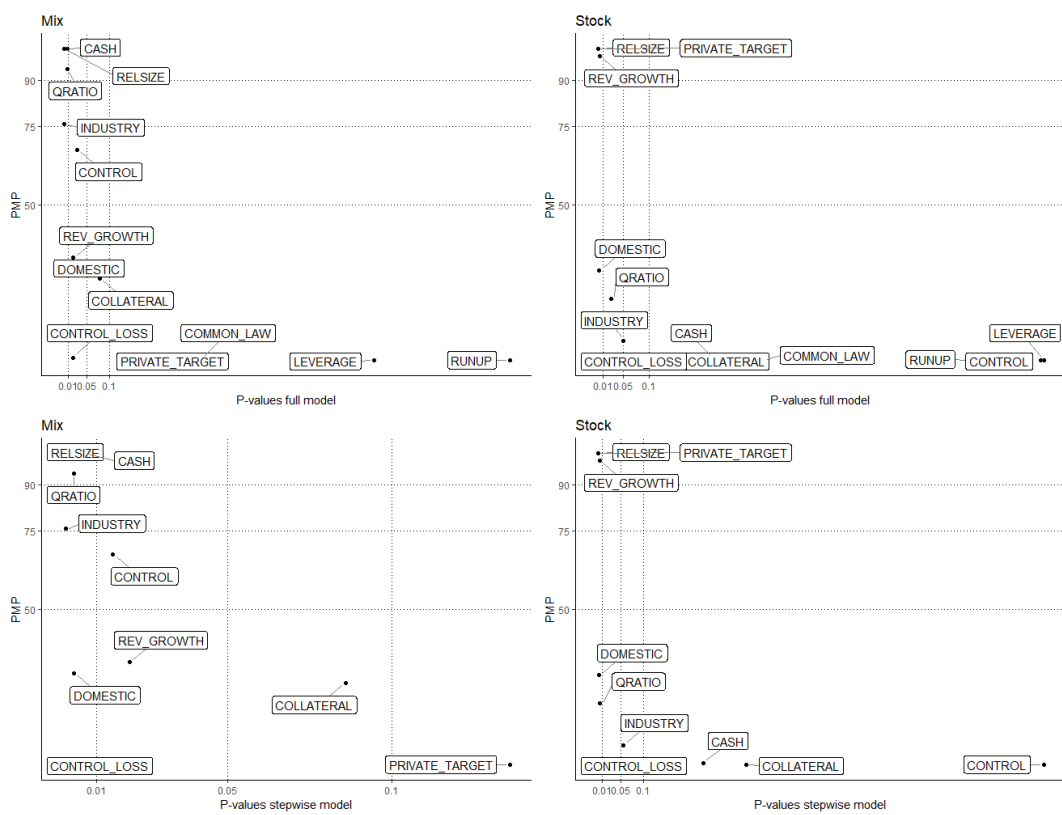
on the grounds of the posterior model probabilities.

Figure (6.5) compares P-values obtained from the full model and stepwise model¹ to the posterior effect probabilities from the BMA. The upper half of the figure compares results of the full model to those of BMA, while the lower half results of stepwise to those of BMA. The left part than presents results for the *mix* equation, right for the *stock* equation. In some cases, the P-values correspond to the posterior probabilities and agree that there is either a strong evidence for an effect or no evidence – these are the points in the top–left and bottom–right part of the plots, respectively (recall that low P-values are equivalent to high posterior probabilities). Nevertheless, there are quite a few situations where the two approaches disagree on the importance of variables. Let us focus on the lower part of the figure, that is stepwise vs. BMA. In the case of the *mix* equation, *CONTROL_LOSS* has a P-value of 0.017, thus being significant on the 0.05 significance level, yet its posterior probability is a mere 0.8 %. This contrast is even more stark in the *stock* equation with four variables having low P-values, yet low posterior probabilities. This once again demonstrates the presence of a strong model uncertainty. Nonetheless, this discrepancy between the posterior probabilities and P-values is nothing unusual and is well documented in Hoeting *et al.* (1999).

In the next chapter we proceed to an interpretation of our regression results. Nevertheless, a question remains as to which of the models selected we ought to use for our final interpretation. Having discussed the various qualitative differences between the automatic methods and BMA, we will solely interpret regression results as presented by the BMA.

¹Recall that stepwise regression led to the very same model specification as backward elimination and forward selection, i.e. the P-values are identical for all three methods.

Figure 6.5: Posterior Probabilities from BMA versus P-values from the Full Model and Stepwise model



Note: The upper half of the figure compares variables' posterior effect probabilities obtained by BMA to their P-values from the full model. The lower part then compares BMA to stepwise regression.

Table 6.5: Summary of BMA regression results

The financing category Mix includes payments made in a combination of cash and stock. The financing category Stock includes payments made entirely in bidder's shares. All independent variables are measured as of deal's announcement date and related to the acquiring firm, unless stated otherwise. *RELSIZE* is (deal value + market capitalization)/(market capitalization). *REV_GROWTH* is a three year (two or one year if data not available) compounded average growth rate of Last Twelve Months revenues prior to a deal. *RUNUP* is 12 month stock return prior to a deal. *QRATIO* is (enterprise value)/(total assets). *CONTROL_LOSS* is (*RELSIZE*)*(target company largest shareholder's stake). *CONTROL* is bidder's largest shareholder. *LEVERAGE* is (book value of debt)/(market capitalization). *COLLATERAL* is (net property, plant & equipment)/(total assets). *CASH* is (cash and short term investments)/(deal value). *PRIVATE_TARGET* is 1 if the target company is privately owned or is a subsidiary. *DOMESTIC* is 1 if bidder and target are domiciled in the same country. *INDUSTRY* is 1 if bidder and target are from the same macro industry as defined by the Reuters Deal Screener App. *COMMON_LAW* is 1 if bidder is domiciled in the UK or Ireland.

	Mix			Stock		
	Coeff.	OR	P ($\beta \neq 0 D$)	Coeff.	OR	P ($\beta \neq 0 D$)
RELSIZE	4.370	79.033	100.00	6.970	1063.875	100.00
REV_GROWTH	0.068	1.070	32.90	0.194	1.214	97.70
RUNUP	0.000	0.000	0.000	0.000	0.000	0.000
QRATIO	0.076	1.079	93.70	0.021	1.021	19.60
CONTROL_LOSS	-0.010	0.990	0.80	0.000	0.000	0.000
CONTROL	-0.840	0.432	67.60	0.000	0.000	0.000
LEVERAGE	0.000	0.000	0.000	0.000	0.000	0.000
COLLATERAL	-0.263	0.769	26.20	0.000	0.000	0.000
CASH	-0.075	0.928	100.00	0.000	0.000	0.50
PRIVATE_TARGET	0.000	0.000	0.000	-1.899	0.150	100.00
DOMESTIC	0.126	1.135	29.30	0.137	1.147	28.80
INDUSTRY	0.395	1.484	75.80	0.027	1.027	6.30
COMMON_LAW	0.002	1.002	0.40	0.000	0.000	0.000

Note: OR means odds ratio.

6.3 Regression Analysis - Results

Results of our regression analysis are available in Table (6.5). The table contains regression coefficients, odds ratios which are obtained by exponentiating coefficients and posterior effect probabilities for each regressor. As mentioned, multinomial logit in our case estimates two equations as we have three dependent variables, meaning we end up with 26 coefficients, odds ratios and posterior probabilities – 13 for the *mix* equation and 13 for the *stock* equation.

6.3.1 Evidence for an Effect of Independent Variables

In interpretation of the posterior effect probabilities, we follow Jeffreys (1998) (as cited in Havranek *et al.* (2018)). Values between 0.5 and 0.75 are categorized as weak, values between 0.75 and 0.95 as positive, values between 0.95 and 0.99 as strong and values above 0.99 as decisive evidence for an effect. Let us first focus on the factors influencing transactions fully financed by bidder's stock, i.e. the right column of the Table (6.5). According to our data and analysis, out of the total 13 independent variables only the regressors *RELSIZE* and *PRIVATE_TARGET* are of a decisive effect. Variable *REV_GROWTH* has a posterior effect probability of almost 98 %, meaning it still displays a strong evidence for an effect. Nevertheless, the rest of the variables displays a subpar evidence for an effect and some of them even none.

Turning to the *mix* equation, we see that the variable *RELSIZE* again displays a decisive evidence for an effect together with *CASH*. Variables *QRATIO* and *INDUSTRY* would be categorized as weak, although being at the different extremes of the scale. Finally, *CONTROL* belongs to the "weak" part of the effect scale.²

6.3.2 Coefficients Interpretation and Comparison with Previous Research

Let us now compare our results with those of the previous research. Faccio & Masulis (2005) used a tobit model in analysis of a percentage of cash financing

²At this point, recall that Table (6.5) displays P-values of the full and automatic methods models and the according posterior effect probabilities from the BMA. P-values clearly overestimate statistical significance of some of the variables, e.g. *DOMESTIC* or *CONTROL_LOSS*.

used in European M&A deals. Regarding the corporate governance concerns, they report that the bidder voting control, in our analysis proxied by the variable *CONTROL*, is statistically significant. (Martin 1996) used a multinomial logit and a spline variable to capture a probable non-linear relationship between the size of the bidder's largest controlling stake and the likelihood of stock financing. He too finds a statistical significance for the concern of bidder's voting control. Finally, Amihud *et al.* (1990) finds that the higher the managerial ownership of the acquirer, the less likely is it to finance acquisitions with stock. On the other hand, our analysis finds a somewhat mixed evidence as the posterior effect probability of *CONTROL* is of a weak effect in the *mix* equation and of a no effect whatsoever in the *stock* equation. Next, the variable *CONTROL_LOSS*, capturing the risk of creating a new dominant shareholder should the bidder finance an acquisition with shares, is found to be insignificant by (Faccio & Masulis 2005). Our analysis concludes that this is the case too.

Consistently with (Martin 1996) and (Faccio & Masulis 2005), we report that our results support the growth opportunities hypothesis, i.e. companies with high growth opportunities shall prefer to finance investments with equity. The variable *QRATIO* is of a positive evidence (posterior effect probability of 93.7 %) in the *mix* equation, while *REV_GROWTH* is of a strong evidence in the *stock* equation. (Faccio & Masulis 2005) propose explanation focusing on the target's incentives – owners of a target which is acquired by a high-growth company might prefer to be compensated in shares which promise high future returns.³ Another plausible explanation is that high-growth companies prefer to finance their investments with equity in order to maintain a maneuvering headroom, since debt puts various constraints and obligations on company's management Killi *et al.* (2011), Myers (1984).

As noted, the choice of payment in M&A is also influenced by financial constraints on the side of the acquiring company. Our analysis finds decisive support only for the variable *CASH* in the case of mix payment. The other two financial characteristics variables, *LEVERAGE* and *COLLATERAL*, display very weak and no effect evidence for mix and stock transactions, respectively.

³A prime real-world example is an acquisition of WhatsApp by Facebook in October 2014 with almost 80 % of the purchase price paid in stock. Facebook's market capitalization is up by almost 100 % since.

(Martin 1996) reports that the ratio of bidder's cash balance to the deals value is statistically significant and of the same sign as in our case. Contrary to our results, (Faccio & Masulis 2005) find that bidder's leverage ratio and its debt capacity proxied by its capacity to collateralize assets are statistically significant, however only at the 10 % level. Taking into consideration that P-values tend to overstate the evidence for an effect (Hoeting *et al.* (1999), recall Figure (6.5)), the actual significance of these factors is questionable.

Next, we investigate the influence of the information asymmetry factors – *RELSIZE*, *RUNUP*, *DOMESTIC* and *INDUSTRY*. Our analysis concludes that the ratio of target's size, approximated by deal value, to the bidder's market capitalization is of a decisive effect in both equations. This is in line with Hansen's prediction that stock financing is more likely as the bidder's information asymmetry increases with the target's size (Hansen (1987) as in (Faccio & Masulis 2005)). Faccio & Masulis (2005) also find this variable to be statistically significant, although only at the 10 % significance level. On the contrary, Martin (1996) finds no evidence whatsoever. Next, we find the dummy variable *INDUSTRY* to be of a weak evidence in the *mix* equation and close to none in the *stock* equation. Faccio & Masulis (2005) employ the very same variable and find a significant and negative relationship, i.e. industry-related transactions are more likely to be financed at least partially with shares. This claim again reflects decreasing information asymmetry concerns since it is reasonable to assume that companies in the same industry are aware of its trends and fair market valuations and thus more willing to accept stock as a payment. Similar logic is behind the positive sign of the variable *DOMESTIC* as companies from the same country face smaller information asymmetry risks. Nevertheless, our model suggest that this effect's evidence is rather weak. Faccio & Masulis (2005) report significance at a 10 % level.

Last but not least, we analyze situations where the bidder is subject to a common law system and when the target company is privately owned. Our results suggest that the effect of a common law legal system on the choice of payment is nonexistent. Nevertheless, a variable interaction between *COMMON_LAW* and *CONTROL* might be considered as to investigate whether the common law legal system exerts different incentives for the controlling shareholder. Finally, we find that if the target company is privately owned or is a subsidiary, the likelihood of a pure stock transaction decreases. Possible

explanation for this phenomenon are liquidity and/or consumption preferences of private owners. In the case of deals involving subsidiaries, plausible explanation might be the reluctance of the bidder to give away equity to another corporation in order not to create a new strategic shareholder which might, so to speak, get a foot in the door. This motivation might be even stronger in transactions where the selling company is a bidder's competitor.

We do not report here intercept coefficients and their posterior effect probabilities, nevertheless they are of a decisive evidence (100 %) and have a negative sign in both equations. Although this result is given simply by the fact that our dataset contains predominantly cash financed transactions, it still confirms the pecking order theory proposed by Myers (1984) which claims that companies prefer to fund investments first with internal sources, i.e. cash, then debt and finally with an issue of new equity, which should be a source of a last resort.

Let us now closer investigate the actual meaning behind coefficients on variables with at least weak effect evidence. First, we ought to note that logit coefficients are interpreted in an entirely different manner than those of a linear regression. Interested reader might for instance consult UCL (2018). Coefficients on *RELSIZE* are 4.37 and 6.97 for the mix and stock transactions, respectively. This results in odds ratios of 79.033 and 1 063.875, which are computed by exponentiating the respective coefficients. Finally, the odds ratio translated to a plain English mean that with a one-unit increase in *RELSIZE*, the odds for a mix transactions compared to a *cash* one increase by 7 803.3 %, *ceteris paribus*. Nonetheless, this is a rather hypothetical example, since we do not usually see companies acquiring targets which are larger than themselves, let alone double their size. Our data indeed confirm this assumption as the distribution of *RELSIZE* has a fat left tail and the maximum value of the variable is 0.95 (see Appendix A). An easier interpretation to grasp would be to say that with a one percentage point increase in the relative size, e.g. from 0.5 to 0.51, the odds for a mix payment increase by 4.5 % relative to a cash transaction (7.2 % for a stock transaction).

A one percentage point increase in the three year compounded annual growth rate of bidder's revenues leads approximately to a 0.2 % increase in the odds for a stock transaction, holding other variables constant. As mentioned, we report very weak evidence in the case of mix transactions. Next, it is rather difficult

to explain the economic intuition behind the coefficient of *QRATIO*. We could say that a one-unit increase in the Q-ratio⁴ of the bidder leads, *ceteris paribus*, to approximately a 8 % higher odds for a mix financed transaction as compared to a cash financed one. Finally, the higher the cash balance of acquirer relative to the deal size, the smaller the probability of using stock. This again agrees with the pecking order theory, as firms should prefer to finance investments from their internal sources. Nevertheless, the interpretation of the coefficient *signs* rather than sizes is more intuitive and meaningful both from the economic point of view and from the perspective of which factors influence the choice of payment in M&A transactions.

In the next section, we consider a simplification of our analysis – we transform our dependent variable to a binary one as well as we transform the independent variable *CONTROL* to its spline to account for its possible non-linear nature.

6.4 Binary Logistic Regression Analysis

Recall the pecking order theory proposed by Myers (1984). The theory claims that companies prefer internal financing to an external one. However, if an external financing is required, i.e. in situations when the internal sources are not sufficient to cover the funds outlay, companies prefer to draw on debt first as opposed to issue a new equity. According to Meyers, 62 % of all capital expenditures by American non-financial corporations in the period of 1973-1982 was covered by internally generated cash. The majority of the external financing was covered by a debt and a new equity issuance was no more than 6 % of the external financing. In short, the decision to fund investments with a newly issued equity is not arbitrary. On the contrary, company issuing a new equity was most likely forced to do so due to some company or market specific factors such as its financial, ownership or market valuation characteristics. Consequently, it might then be more reasonable not to distinguish between purely cash financed transactions, mixed transactions and solely stock financed transactions, but rather between the former and the case where company uses at least some equity, i.e. considering the mix and stock financed transactions to be of the same nature. Therefore, in this section we introduce a binomial logit

⁴Recall that Q-ratio measures the "over-valuation" of a company. Q-ratio of 1 means that the company's enterprise value equals its assets replacements cost, i.e. an equilibrium valuation.

where the dependent variable takes on only two values:

$$y_i = \begin{cases} 1, & \text{if the } i\text{-th transaction payment method was cash} \\ 0, & \text{otherwise} \end{cases} \quad (6.1)$$

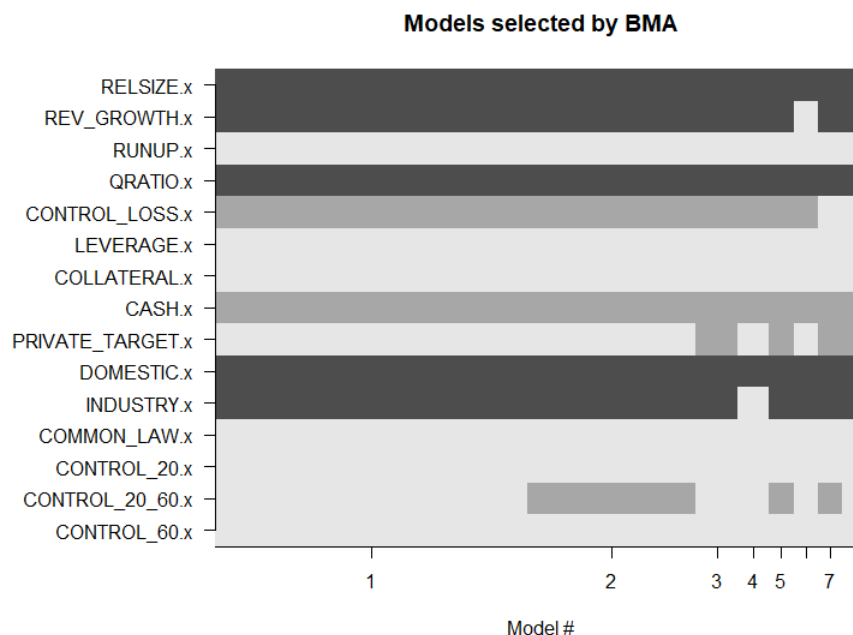
Furthermore, we replace the variable *CONTROL* with its spline specification based on Faccio & Masulis (2005). The variable is defined as follows:

$$\begin{aligned} \text{CONTROL}_{.20} &= \text{CONTROL} && \text{if } \text{CONTROL} < 20 \% \\ &= 20\% && \text{if } \text{CONTROL} \geq 20 \% \\ \text{CONTROL}_{.20_60} &= 0 && \text{if } \text{CONTROL} < 20 \% \\ &= \text{CONTROL} - 20 \% && \text{if } 20 \% \geq \text{CONTROL} \leq 60 \% \\ &= 40 \% && \text{if } \text{CONTROL} > 60 \% \\ \text{CONTROL}_{.60} &= 0 && \text{if } \text{CONTROL} < 60 \% \\ &= \text{CONTROL} - 60 \% && \text{if } \text{CONTROL} \geq 60 \% \end{aligned}$$

This spline variable ought to better capture the possible non-linear nature of the variable *CONTROL* as each coefficient measures the slope of the regression line over the stated intervals (Martin (1996), Faccio & Masulis (2005)). Otherwise than that, we keep all the previously stated variables in the model and again employ the BMA to average over all possible model specifications.

Figure (6.6) displays the variable inclusion over various models selected by the BMA. Interpretation of this figure is analogous to that of Figure (6.4), i.e. horizontal lines represent frequency of variable inclusion over all models selected by the BMA (8 in this case). We can infer from the figure that the variables *RELSIZE*, *QRATIO*, *CASH* and *DOMESTIC* most likely have a decisive evidence for an effect since they are included in all of the models. Variables *REV_GROWTH*, *CONTROL_LOSS* and *INDUSTRY* will be of a relatively strong effect, too. The rest will either be of a weak or no evidence for an effect. Table (6.6) formally presents results from the BMA regression. Indeed, the posterior effect probabilities agree with the above stated claims. Coefficient signs are in line with expectations and are stable through both multinomial logit (Table (6.5)) and binary logit. Recall, that the cumulative posterior probability of the five best *multinomial* logit models was 22.6 % with the best model having a posterior probability of only 6 %. The cumulative posterior probability of the best five *binary* logit models is 89.1 % with the best model's posterior probability of 48.1 %. Collapsing the dependent variables *stock* and *mix* to only one option resulted in a stark increase in the posterior

Figure 6.6: Model Inclusion in Bayesian Model Averaging - Binomial Logit



probabilities and at the same time led to a model which is simpler to interpret.

By far the largest effect on the payment method decision has the variable *RELSIZE*, which is proxy for the relative size of the target and the bidder. The prospect of creating a new strong shareholder in the bidder's ownership structure, represented by *CONTROL_LOSS*, strongly deters bidders to pay in stock. Again, we find a strong support for the investment opportunities hypothesis as both *QRATIO* and *REV_GROWTH* are of a very strong evidence for an effect. Regarding the information asymmetry, domestic deals and intra-industry deals are a lot more probable to be financed at least in part with stock as compared to a sole cash payment. *Ceteris paribus*, industry related and domestic deals increase the odds ratio of a stock payment relative to a cash method by 67.2 % and 70.4 %, respectively.

In regard to the newly defined spline variable, we report that intermediate levels of ownership by the bidder's largest shareholder are connected with a lower probability of stock financing. Nonetheless, the evidence for the effect is rather weak with a posterior effect probability of 33.4 %. On the other hand, Martin (1996) and Faccio & Masulis (2005) find this effect to be statistically

significant.

Table 6.6: Summary of BMA Regression Results

	Coefficient	Odds Ratio	$P(\beta \neq 0 D)$
RELSIZE	8.987	7 998.426	100 %
REV_GROWTH	0.205	1.228	96.2 %
RUNUP	0.000	1.000	0.00 %
QRATIO	0.114	1.121	100 %
CONTROL_LOSS	-4.368	0.013	92.8 %
LEVERAGE	0.000	1.000	0.00 %
COLLATERAL	0.000	1.000	0.00 %
CASH	-0.044	0.957	100 %
PRIVATE_TARGET	-0.109	0.897	17.7 %
DOMESTIC	0.533	1.704	100 %
INDUSTRY	0.514	1.672	95.3 %
COMMON_LAW	0.000	1.000	0.00 %
CONTROL_20	0.000	1.000	0.00 %
CONTROL_20_60	-0.456	0.634	33.4 %
CONTROL_60	0.000	1.000	0.00 %

Chapter 7

Conclusion

In this thesis, we attempted to analyze which factors influence choice of payment in M&A transactions, specifically focusing on transactions in which the bidder firm was domiciled in one of the 28 European Union countries. Our model included in total 13 independent variables proxying for bidder's financial and corporate governance characteristics, issues of information asymmetry faced by both the bidder and the target firm as well as other factors such as different law systems between continental Europe and United Kingdom and Ireland. Focal point of our analysis, the payment method used in a transaction, was coded as a three level categorical variable representing three possible consideration options - cash, stock and a mix of these. Given the nature of our analysis, we employed a multinomial logit model for our analysis.

With 13 independent variables we had in total 8 192 possible model specifications to analyze. First part of our empirical analysis thus focused on finding the "best" model. We used the so-called automated model selection algorithms, specifically backward elimination, forward selection, stepwise regression and best subset regression. All of these methods arrived at the very same "best" model keeping 10 variables out of the original 13. Having in mind possible flaws of these algorithms, we proceeded to a Bayesian model averaging, which does not necessarily select a "best" model, but rather averages over all models and accordingly adjusts variables' coefficients and standard errors. Inspired by the work of Hoeting *et al.* (1999), we compared P-values of the "best" model as selected by the automated selection methods to posterior effect probabilities of the Bayesian model averaging. We concluded that P-values might be misleading in assessing variable importance and, therefore, we based our statistical

inference solely on the results obtained by the Bayesian model averaging.

Based on the regression results, we found that the evidence for the corporate governance factors is somewhat mixed. Some of our findings on the issue of ownership dilution are in line with previous research, however some are not. Furthermore, our results agree with previous studies on the topic of growth opportunities faced by the bidder, that is high-growth and high-value companies tend to finance acquisitions at least partly with stock. Regarding the bidders financial characteristics such as financial leverage, amount of tangible assets for collateralization and cash balance, we conclude that the ratio of cash balance to the deal's value has a strong evidence for an effect. On the other hand, bidder's leverage and his ability to post assets as a collateral are of a no effect. Last but not least, we explore the effects of variables addressing information asymmetry issues. First, we do not find evidence for the effect of domestic transactions, but we report that there is an evidence for a more likely stock financing in same industry deals and in deals involving privately owned target. Finally, the bidders governing law system is of no influence on the choice of the payment method.

As the last step, we argued that the motivation for mix and stock financed transactions is virtually the same. Therefore, we combined the three level dependent variable into a binary one, which did not distinguish between mix and stock financed transactions, but only between cash financed deals and transactions with some ratio of stock financing. Moreover, we introduced a new variable capturing possible non-linear nature of the bidder's largest shareholder behavior. Simplifying the multinomial logit regression to a binary one led not only to a stark improvement in the cumulative posterior probability of the models selected by the Bayesian averaging, but at the same time to a more comprehensible interpretation of variables' coefficients. Regarding the newly created ownership control variable, we reported that it is of an expected sign, however the evidence for an effect is weak.

Last but not least, we would like to suggest several ideas for a further research on the topic. As discussed, transactions including privately held targets are less likely to be financed with stocks, probably due to owners' consumption and liquidity preferences. Therefore, it might be of an interest to investigate the relation between the selling shareholder's age and the choice of the payment.

Intuitively, younger sellers might be more willing to accept stock payments due to their lower liquidity preferences. The obvious issue is the difficulty of obtaining such data. Another plausible variable interaction worth investigating might be the case of transactions where the target is a subsidiary and at the same time the exiting target's shareholder is the bidder's competitor. Logic would dictate that in such cases the bidder ought to be reluctant to pay in stock. Last but not least, it is both argued in the academia (Myers (1984)) and well-known fact among market participants and pundits that companies tend to "time" their stock issues and wait for favorable market conditions. It thus might be of an interest to analyze whether an overall market valuation (proxied for instance by the aggregate Price-to-Earnings ratio of a market) influences choice of the payment in M&A. Speaking of the overall market conditions, one might also analyze influence of interest rates as low interest environment incentives the use of cash for funding investments. Moreover, high stock market valuations usually appear during monetary expansions, meaning that these two factors are of an opposite effect and the question is which one prevails. In regard to data used for the regression analysis, we would suggest to use dataset spanning more than a decade in order to control for effect of business cycles. For instance, our data set includes transactions announced between the years 2010-2018, i.e. during an unprecedented monetary expansion in most of the advanced economies. On the other hand, the previous research done on European transactions by Faccio & Masulis (2005) included deals from the boom years of 1997-2000, years of the dot-com bubble and sky-high stock valuations.

Finally, it is documented that bidders using stock as a medium of exchange in M&A transactions experience significant decrease in their stock price at the transaction announcement date (Travlos (1987)). Consequently, investors could earn excess returns by shorting bidder's stock knowing that he is about to finance an acquisition with a new equity issuance. One could thus construct a binary machine learning logit model which would try to classify transactions by payment methods based on past data. On the other hand, practical usefulness of such a model is questionable, since investors, with the exception of insiders, do not usually know about M&A transactions prior to the announcement date. Still, market rumors are a commonplace and speculation would be possible.

Bibliography

(2018): .

ADAM, T. & V. K. GOYAL (2008): “The investment opportunity set and its proxy variables.” *Journal of Financial Research* **31(1)**: pp. 41–63.

ALAN J. AUERBACH, D. R. (1987): *The Impact of Taxation on Mergers and Acquisitions*, chapter The Impact of Taxation on Mergers and Acquisitions, pp. 69–86. University of Chicago Press.

AMIHUD, Y., B. LEV, & N. G. TRAVLOS (1990): “Corporate control and the choice of investment financing: The case of corporate acquisitions.” *The Journal of Finance* **45(2)**: pp. 603–616.

ANDERSON, D. R. (2007): *Model Based Inference in the Life Sciences*. Springer New York.

BENDEL, R. B. & A. A. AFIFI (1977): “Comparison of stopping rules in forward ”stepwise” regression.” *Journal of the American Statistical Association* **72(357)**: p. 46.

BERGER, J. O. & M. DELAMPADY (1987): “Testing precise hypotheses.” *Statistical Science* **2(3)**: pp. 317–335.

BERGER, J. O. & T. SELLKE (1987): “Testing a point null hypothesis: The irreconcilability of p values and evidence.” *Journal of the American Statistical Association* **82(397)**: p. 112.

BOONE, A. L., E. LIE, & Y. LIU (2014): “Time trends and determinants of the method of payment in m&as.” *Journal of Corporate Finance* **27**: pp. 296–304.

CALCAGNO, V. & C. DE MAZANCOURT (2010): “glmulti: AnRPackage for easy automated model selection with (generalized) linear models.” *Journal of Statistical Software* **34(12)**.

- CARLETON, W. T., D. K. GUILKEY, R. S. HARRIS, & J. F. STEWART (1983): "An empirical analysis of the role of the medium of exchange in mergers." *The Journal of Finance* **38(3)**: pp. 813–826.
- CHANEY, P. K., L. M. LOVATA, & K. L. PHILIPICH (1991): "Acquiring firm characteristics and the medium of exchange." *Quarterly Journal of Business and Economics* **30(4)**: pp. 55–69.
- CHEMMANUR, T. J., I. PAEGLIS, & K. SIMONYAN (2009): "The medium of exchange in acquisitions: Does the private information of both acquirer and target matter?" *Journal of Corporate Finance* **15(5)**: pp. 523–542.
- CHOE, H., R. W. MASULIS, & V. NANDA (1993): "Common stock offerings across the business cycle." *Journal of Empirical Finance* **1(1)**: pp. 3–31.
- DONG, M. N., D. A. HIRSHLEIFER, S. A. RICHARDSON, & S. H. TEOH (2003): "Does investor misvaluation drive the takeover market?" *SSRN Electronic Journal* .
- ECKBO, B. E., R. M. GIAMMARINO, & R. L. HEINKEL (1990): "Asymmetric information and the medium of exchange in takeovers: Theory and tests." *Review of Financial Studies* **3(4)**: pp. 651–675.
- EDWARDS, W., H. LINDMAN, & L. J. SAVAGE (1963): "Bayesian statistical inference for psychological research." *Psychological Review* **70(3)**: pp. 193–242.
- EICHER, T. S., C. PAPAGEORGIOU, & A. E. RAFTERY (2009): "Default priors and predictive performance in bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.
- FACCIO, M. & R. W. MASULIS (2005): "The choice of payment method in european mergers and acquisitions." *The Journal of Finance* **60(3)**: pp. 1345–1388.
- FISHMAN, M. J. (1989): "Preemptive bidding and the role of the medium of exchange in acquisitions." *The Journal of Finance* **44(1)**: pp. 41–57.
- GHOSH, A. & W. RULAND (1998): "Managerial ownership, the method of payment for acquisitions, and executive job retention." *The Journal of Finance* **53(2)**: pp. 785–798.

- GREENE, W. H. (2002): *Econometric Analysis*. Prentice Hall.
- HANSEN, R. G. (1987): “A theory for the choice of exchange medium in mergers and acquisitions.” *The Journal of Business* **60(1)**: p. 75.
- HARFORD, J., S. KLASA, & N. WALCOTT (2009): “Do firms have leverage targets? evidence from acquisitions.” *Journal of Financial Economics* **93(1)**: pp. 1–14.
- HARRELL, F. E. (2013): *Regression Modeling Strategies*. Springer New York.
- HARRIS, R. S., J. F. STEWART, D. K. GUILKEY, & W. T. CARLETON (1982): “Characteristics of acquired firms: Fixed and random coefficients probit analyses.” *Southern Economic Journal* **49(1)**: p. 164.
- HASTIE, T., R. TIBSHIRANI, & J. FRIEDMAN (2009): *The Elements of Statistical Learning*. Springer-Verlag New York Inc.
- HAVRANEK, T., Z. IRSOVA, & O. ZEYNALOVA (2018): “Tuition fees and university enrolment: A meta-regression analysis.” *Oxford Bulletin of Economics and Statistics* **80(6)**: pp. 1145–1184.
- HELLEVIK, O. (2007): “Linear versus logistic regression when the dependent variable is a dichotomy.” *Quality & Quantity* **43(1)**: pp. 59–74.
- HERON, R. & E. LIE (2002): “Operating performance and the method of payment in takeovers.” *The Journal of Financial and Quantitative Analysis* **37(1)**: p. 137.
- HOETING, J. A., D. MADIGAN, A. E. RAFTERY, & C. T. VOLINSKY (1999): “Bayesian model averaging: A tutorial.” *Statistical Science* **14(4)**: pp. 382–401.
- HOSMER, D. W. & S. LEMESHOW (2000): *Applied Logistic Regression (Wiley Series in Probability and Statistics)*. Wiley-Interscience Publication.
- JEFFREYS, H. (1998): *Theory of Probability*. OXFORD UNIV PR.
- JENSEN, M. C. & R. A. POSNER (1986): “Agency costs of free cash flow, corporate finance, and takeovers.” In J. S. BHANDARI & L. A. WEISS (editors), “Corporate bankruptcy,” pp. 11–16. Cambridge University Press.

- JULIAN R. FRANKS, Robert S. Harris, C. M. (1988): *Corporate Takeovers: Causes and Consequences*, chapter Means of Payment in Takeovers: Results for the United Kingdom and the United States Kingdom and the United States, pp. 221–264. University of Chicago Press.
- JUNG, K., Y.-C. KIM, & R. STULZ (1996): “Timing, investment opportunities, managerial discretion, and the security issue decision.” *Journal of Financial Economics* **42(2)**: pp. 159–185.
- KALDOR, N. (1966): “Marginal productivity and the macro-economic theories of distribution: Comment on samuelson and modigliani.” *The Review of Economic Studies* **33(4)**: p. 309.
- KILLI, A., M. S. RAPP, & T. SCHMID (2011): “Can financial flexibility explain the debt conservatism puzzle? cross-country evidence from listed firms.” *SSRN Electronic Journal* .
- KJÓRULFF, O. (2017): “P-value abuse distorts scientific publication.”
- LANTZ, B. (2012): “The large sample size fallacy.” *Scandinavian Journal of Caring Sciences* **27(2)**: pp. 487–492.
- LEUVEN, G. K. U., B. CLAESKENS, & N. L. U. I OSLO) HJORT (2008): *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press.
- M., A. S. & P. C. F. (2011): “Bayesian model averaging in r.” *Journal of Economic and Social Measurement* **36(4)**: pp. 253–287.
- MADIGAN, D. & A. E. RAFTERY (1994): “Model selection and accounting for model uncertainty in graphical models using occam's window.” *Journal of the American Statistical Association* **89(428)**: pp. 1535–1546.
- MARTIN, K. J. (1996): “The method of payment in corporate acquisitions, investment opportunities, and management ownership.” *The Journal of Finance* **51(4)**: pp. 1227–1246.
- MASSOUDI, A., J. FONTANELLA-KHAN, & D. WEINLAND (2017): “Global m&a exceeds usd 3tn for fourth straight year.” *Financial Times* .
- MYERS, S. C. (1984): “The capital structure puzzle.” *The Journal of Finance* **39(3)**: pp. 574–592.

- NIDEN, C. (1986): “The role of taxes in corporate acquisitions: effects on premium and type of consideration.” University of Chicago, Working paper, 1986.
- OFFICER, M. S. (2004): “Collars and renegotiation in mergers and acquisitions.” *The Journal of Finance* **59(6)**: pp. 2719–2743.
- PLATT, E. (2018): “Global m&a activity hits new high.” *Financial Times* .
- RAFTERY, A., J. HOETING, C. VOLINSKY, I. PAINTER, & K. Y. YEUNG (2018): *BMA: Bayesian Model Averaging*. R package version 3.18.9.
- RAFTERY, A. E. (1995): “Bayesian model selection in social research.” *Sociological Methodology* **25**: p. 111.
- RAFTERY, A. E., D. MADIGAN, & J. A. HOETING (1997): “Bayesian model averaging for linear regression models.” *Journal of the American Statistical Association* **92(437)**: pp. 179–191.
- ŠEVCIKOVÁ, H. & A. RAFTERY (2012): “Estimation of multinomial logit model using the begg & gray approximation.” .
- SEVCIKOVA, H. & A. RAFTERY (2013): *mlogitBMA: Bayesian Model Averaging for Multinomial Logit Models*. R package version 0.1-6.
- STEVENS, D. L. (1973): “Financial characteristics of merged firms: A multivariate analysis.” *The Journal of Financial and Quantitative Analysis* **8(2)**: p. 149.
- THOMPSON, B. (1989): “Why won't stepwise methods die?” *Measurement and Evaluation in Counseling and Development* **21(4)**: pp. 146–148.
- THOMPSON, B. (1995): “Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial.” *Educational and Psychological Measurement* **55(4)**: pp. 525–534.
- TITMAN, S. & R. WESSELS (1988): “The determinants of capital structure choice.” *The Journal of Finance* **43(1)**: pp. 1–19.
- TOBIN, J. (1958): “Estimation of relationships for limited dependent variables.” *Econometrica* **26(1)**: p. 24.

- TOBIN, J. & W. C. BRAINARD (1976): “Asset markets and the cost of capital.” *Cowles Foundation Discussion Papers 427*, Cowles Foundation for Research in Economics, Yale University.
- TRAVLOS, N. G. (1987): “Corporate takeover bids, methods of payment, and bidding firms' stock returns.” *The Journal of Finance* **42(4)**: pp. 943–963.
- TUTZ, G., W. PSSNECKER, & L. UHLMANN (2015): “Variable selection in general multinomial logit models.” *Computational Statistics & Data Analysis* **82**: pp. 207–222.
- UYSAL, V. B. (2011): “Deviation from the target capital structure and acquisition choices.” *Journal of Financial Economics* **102(3)**: pp. 602–620.
- VIECHTBAUER, W. (2018): “Model selection using the glmulti package.”
- WEISBERG, S. (1985): *Applied Linear Regression (Wiley Series in Probability and Statistics)*. Wiley.
- ZEUGNER, S. & M. FELDKIRCHER (2015): “Bayesian model averaging employing fixed and flexible priors: The BMS package for R.” *Journal of Statistical Software* **68(4)**: pp. 1–37.

Appendix A

Histograms of Independent Variables Categorized by the Payment Method

Figure A.1: Conditional histogram for the dependent variable *Cash*

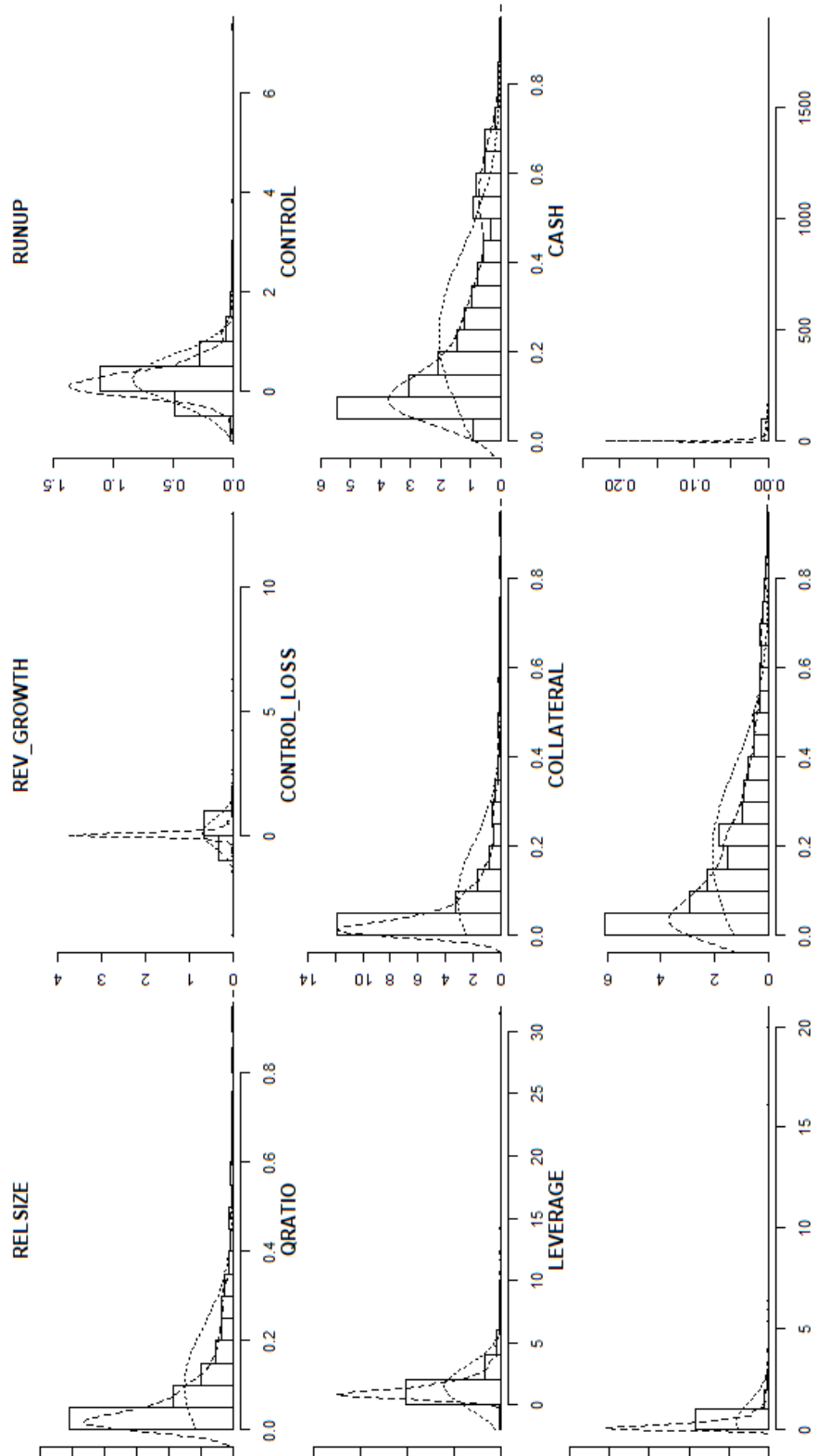


Figure A.2: Conditional histogram for the dependent variable *Stock*

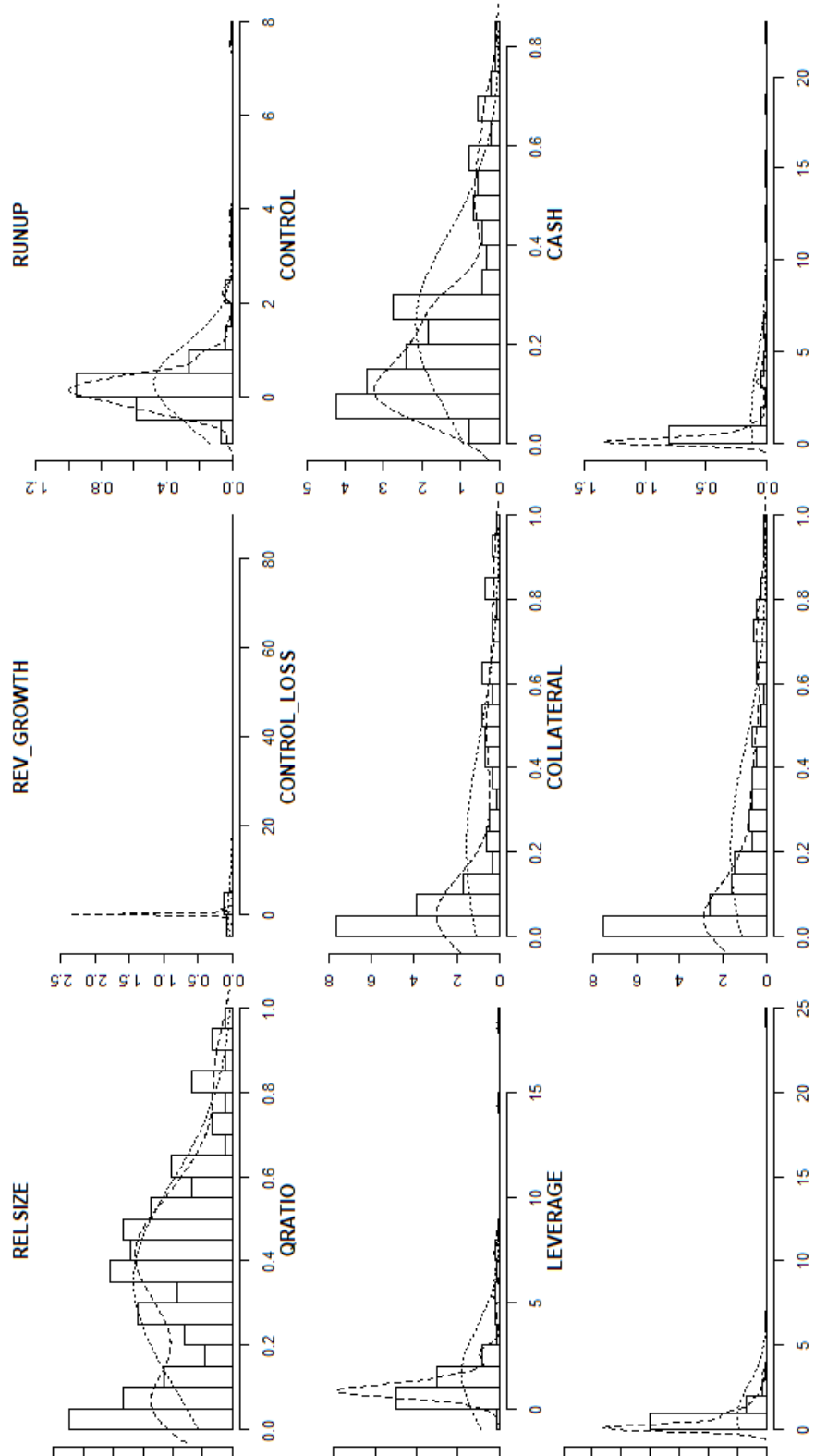
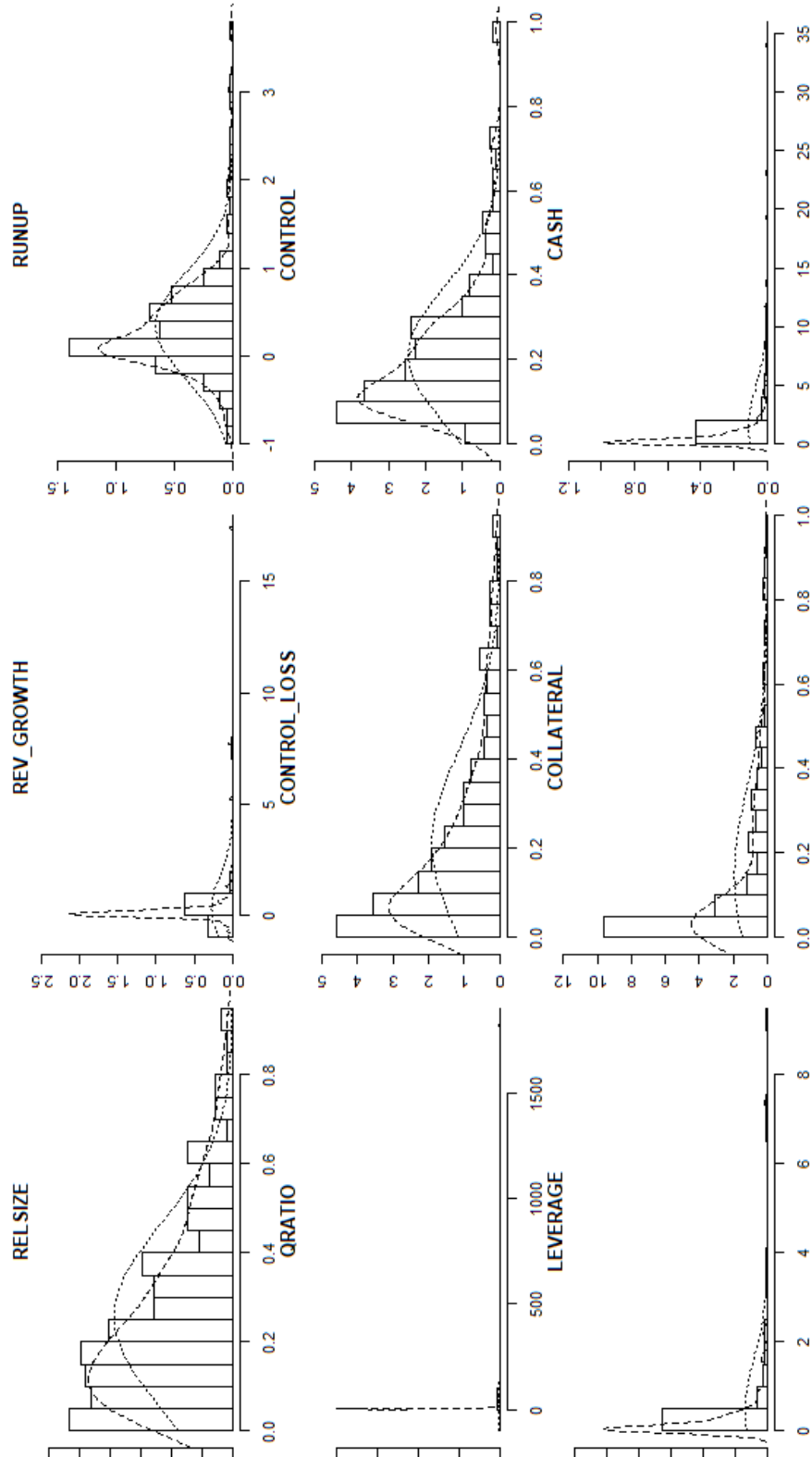


Figure A.3: Conditional histogram for the dependent variable *Mix*



Appendix B

Regression Tables of Models Selected by Automatic Selection Methods

Table B.1: Models Selected by the Automatic Selection Methods

	<i>Dependent variable:</i>					
	Backward		Forward		Stepwise	
	mix	stock	mix	stock	mix	stock
COLLATERAL	-0.746*	0.448	-0.746*	0.448	-0.746*	0.448
	(0.434)	(0.451)	(0.434)	(0.451)	(0.434)	(0.451)
CASH	-0.069***	-0.022	-0.069***	-0.022	-0.069***	-0.022
	(0.025)	(0.018)	(0.025)	(0.018)	(0.025)	(0.018)
QRATIO	0.111***	0.107***	0.111***	0.107***	0.111***	0.107***
	(0.036)	(0.037)	(0.036)	(0.037)	(0.036)	(0.037)
REV_GROWTH	0.193**	0.254***	0.193**	0.254***	0.193**	0.254***
	(0.082)	(0.086)	(0.082)	(0.086)	(0.082)	(0.086)
CONTROL	-1.171**	0.020	-1.171**	0.020	-1.171**	0.020
	(0.479)	(0.524)	(0.479)	(0.524)	(0.479)	(0.524)
CONTROL_LOSS	-3.446**	-1.909	-3.446**	-1.909	-3.446**	-1.909
	(1.432)	(1.315)	(1.432)	(1.315)	(1.432)	(1.315)
RELSIZE	7.387***	8.357***	7.387***	8.357***	7.387***	8.357***
	(1.349)	(1.180)	(1.349)	(1.180)	(1.349)	(1.180)
INDUSTRY	0.645***	0.423*	0.645***	0.423*	0.645***	0.423*
	(0.185)	(0.220)	(0.185)	(0.220)	(0.185)	(0.220)
DOMESTIC	0.496***	0.631***	0.496***	0.631***	0.496***	0.631***
	(0.165)	(0.201)	(0.165)	(0.201)	(0.165)	(0.201)
PRIVATE_TARGET	0.465	-1.710***	0.465	-1.710***	0.465	-1.710***
	(0.311)	(0.309)	(0.311)	(0.309)	(0.311)	(0.309)
Constant	-3.061***	-3.062***	-3.061***	-3.062***	-3.061***	-3.062***
	(0.370)	(0.354)	(0.370)	(0.354)	(0.370)	(0.354)
Akaike Inf. Crit.	1,769.197	1,769.197	1,769.197	1,769.197	1,769.197	1,769.197

Note:

*p<0.1; **p<0.05; ***p<0.01

Appendix C

BMA Posterior Distribution of Independent Variables

Figure C.1: Posterior Distribution of Independent Variables

