Developing a Database of Merger and Acquisition Transactions: Towards Insights into Industry Trends and Drivers

by

Terry Wayne Heathcote

Thesis presented in fulfilment of the requirements for the degree of Master of Engineering (Engineering Management) in the Faculty of Engineering at the Stellenbosch University

Supervisor: Dr. Wouter Gideon Bam
Co-supervisor: Prof. Sara Susanna Grobbelaar

March 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2020 .....

Abstract

Mergers and Acquisitions (M&A) have experienced great fluctuations in activity throughout history, with the characteristics of deal trends changing constantly throughout time. Studies in the field focus greatly on the determinants of waves in merger activity, commonly testing proposed theories by empirical means. Few studies find a consensus on appropriate proxy use in determinant analyses and as a result, often find discord with prior findings in literature. Through a systematic quantitative literature review, aimed at identifying the traditionally studied M&A activity characteristics and drivers, lists of synthesized activity and determinant variables were established. Using these outputs as information requirements for envisioned analyses, a data warehouse was developed and populated with a sample of data gathered for domestic deals in the USA during the years 1998 to 2018, between listed acquirers and targets. Using the Database Life Cycle and Data Warehouse Architecture, an information storage facility, capable of driving M&A activity and determinant analysis, was developed. M&A trends were analyzed for the sample in terms of identified activity characteristics. Recent years show a steep increase in average deal value, giving way to an era of mega-mergers. The total value of the cash and stock as well as cash only payments has increased significantly from the stock only payment dominance in the late 1990's and early 2000's. In an evaluation of traditionally studied M&A determinants, a holistic approach is taken in considering a variety of proxies, while acknowledging established theoretical classifications. By applying methods in feature selection, a refined set of relevant determinant proxies were identified and subsequently analyzed using multiple linear regressions. The resulting models for annual deal volume and value proved to support both the Neoclassical and Macroeconomic theories of M&A, with little evidence supporting the Behavioral theory. The approach to determinant analysis proved to be effective in improving predictive ability for models, while initially considering a broad variety of determinant proxies recognized in literature. However, additional proxies for Firm-Level theories could be introduced in the future, using the same or a similar approach to analyses. This could provide a more comprehensive evaluation of determinants in the field.

Opsomming

Samesmeltings en verkrygings (M&A) ervaar groot fluktuasies in aktiwiteit deur die loop van geskiedenis, met die kenmerke van transaksie tendense wat deurentyd verander. Studies in die veld fokus grootliks op die bepaalde faktore van veranderings in samesmeltingsaktiwiteit, en meestal word voorgestelde teorieë op empiriese wyse getoets. Selde vind studies ‘n ooreenstemming oor toepaslike volmaggebruik in determinant analise, en vind gevolglik dikwels onenigheid met vorige bevindings in die literatuur. Deur middel van ‘n sistematiese kwantitatiewe literatuuroorsig wat daarop gemik was om die tradisioneel bestudeerde M&A-aktiwiteitseienskappe en drywers te identifiseer, is lyste van gesintetiseerde aktiwiteit en determinantveranderlikes opgestel. Met behulp van hierdie uitsette as inligtingsvereistes vir beoogde ontledings, is ‘n datapakhuis ontwikkel en gevul met data, wat versamel is vir binnelandse ooreenkoms in die VSA gedurende die jare 1998 tot 2018, tussen genoteerde verkrygers en teikens. Met behulp van die databasis-lewensiklus en datapakhuis-argitektuur, is ‘n inligtingsbergingsfasiliteit ontwikkel wat M&A-aktiwiteit en determinant-analise kan dryf. Die data was geanaliseer vir M&A-neigings in terme van geïdentifiseerde aktiwiteitseienskappe. Die afgelope jare toon ‘n skerp toename in die gemiddelde transaksiewaarde, wat plek maak vir ‘n era van mega-samesmeltings: die totale waarde van kontant en aandele sowel as slegs kontantbetalings het aansienlik afgewyk van die hoofsaaklike slegs aandele betaaling in die laat 1990‘s en vroe 2000‘s. In ‘n evaluering van tradisioneel bestudeerde M&A-determinante word ‘n holistiese benadering gevolg in die oorweging van verskillende gevolmagtigdes, terwyl erkenning gegee word aan gevestigde teoretiese klassifikasies. Deur metodes van funksie-seleksie toe te pas, is ‘n verfynde stel relevante determinant proxy‘s geïdentifiseer en daarna met behulp van veelvuldige lineêre regressies ontleed. Die gevolglike modelle vir jaarlikse kontrak volume en waarde bewys dat dit die Neoklassieke en Makro-ekonomiese teorieë van M&A ondersteun word, met min bewyse om die gedragsteorie te ondersteun. Die benadering tot determinantanalise blyk effektief te wees in die verbetering van die voorspellingsvermoë vir modelle, terwyl daar aanvanklik ‘n wye verskeidenheid determinante-gevolmagtigdes in die literatuur erken is. Maar, bykomende gevolmagtigdes vir Firmevlak-teorieë kan in die toekoms met behulp van dieselfde of ‘n soortgelyke benadering bekendgestel word. Dit kan ‘n meer volledige evaluering van determinante in die

veld bied.

Acknowledgments

The following acknowledgments should be made of contributions to and support of this thesis:

1. Dr. Wouter Bam (supervisor), for the guidance, mentoring and time given through the execution of this thesis as well as facilitating research group activities.

2. Prof. Sara Grobbelaar (co-supervisor), for the guidance and knowledge passed on as well as facilitating research group activities.

3. Mrs. Susan Higgo, for guidance and knowledge transfer in M&A as well as for championing the research group formation.

4. Dr. Estee Willers, for the support and genuine interest shown throughout.

5. My family, for their support and encouragement through this endeavor.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Present day enterprises find it increasingly difficult to compete in the ever changing, globalized market of business. Geographical and industry borders are continuously diminishing or shifting. For this reason, companies often make efforts to stream-line operations, better realizing improved internal efficiencies. This allows entities to remain competitive players in such markets. However, this specific tactic may not always result in the attainment of strategic objectives or performance enhancements required.

Envisioned enterprises often turn to an alternative means of growth, merging and acquiring, a process perceived to allow the assertion of dominance or continued competitiveness in a market. Here, external growth is complemented by the partnering with, or acquisition of, enterprises whose assets, technologies, intellectual property or business processes may align with the strategic intent of the acquiring entity. The corporate management intentions in this process, stimulated by shareholder expectations (among other rationale), drive enterprises to pursue these external mechanisms of growth and in so doing, potentially change the configuration and performance of many facets of the former organization.

According to the Institute for Mergers, Acquisitions and Alliances, mergers and acquisitions (M&A) have seen a steady increase in global deal volume since 1985. Figure 1.1, shows that 2 676 transactions were completed in that year. 33 years later, in 2018, a total of 51 865 deals were concluded. The average value of the accumulated deals globally, per year, for the last ten years, was approximately 3.277 trillion USD, 356 billion USD more than the average of the ten year period prior to that.

The history of M&A market activity reveals that it occurred predominantly in cycles, otherwise known as waves (Cartwright *et al.*, 2012). Waves have been observed and studied as far back as the year 1895, with several peaks in activity observed since then. These could be characterized considering the

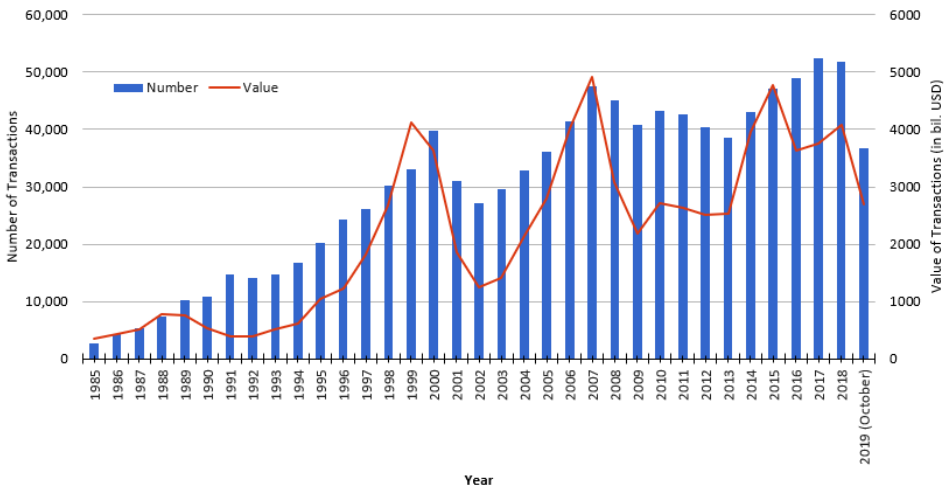


Figure 1.1: Historical Worldwide M&A number of transactions and value - Sourced from (IMAA, 2019).

frequency, value and types of M&A deals which occurred throughout a specific era or time period. As the timing and duration of these cycles has not been consistent, among other variables, many theories on the drivers of waves have been posited in literature.

During the last 20 years, the global economy has seen tremendous shake-ups. The economic effects of both the "Dot-com" market bubble and the Financial Crisis of 2008 were significant antagonists, with literature suggesting this also affected aggregate M&A activity (Yaghoubi *et al.*, 2016). Factors within the economy of the United States of America (U.S.A) were believed to be major drivers in these economic upsets. With regulation, the state of capital markets and industry shocks seen as common economic drivers of the M&A industry, the country is an appealing candidate for an analysis of the field and the influence of activity determinants, specifically.

Figure 1.2 further proves the significance of the North American market (in which the U.S.A plays the largest role) on the global stage, through a comparison of the M&A deal volume over the past 20 years. A two year moving average for the same period also uncovers similarities in activity trends between that of the North American Market and further abroad, suggesting the countries' activity may be a good indicator of global activity. The correlation coefficients for number of transactions and total value between North America and worldwide are 0.92 and 0.98 respectively.

In terms of the number of transactions and value, Figure 1.3 shows the significant contribution North America has made to the global market over the past 20 years. Interestingly, the value of North American deals annually, did not sink below the 40% mark throughout the period.

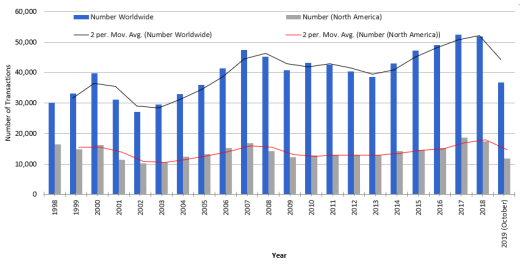


Figure 1.2: North America and Worldwide M&A for number of transactions and two year moving average - Data sourced from (IMAA, 2019).

Figure 1.3: North American deal frequency and value as a percentage of worldwide activity - Data sourced from (IMAA, 2019).

While the characterization of M&A waves has focused itself largely on frequency, value and deal types, greater depth for description lies in the stores of M&A data available for historical transactions. This is becoming increasingly accessible, with more deal attributes being recorded and due to greater transparency of information, especially in the public domain. Consultancies and financial information providers regularly produce reports and data for the industry. Usually, their aim is to convey the prior term's (being of a quarterly, bi-annual or annual frequency) trends, building towards producing an outlook for future activity. This is generally analysed in terms of frequency, size, region and industry, where, less commonly, deeper insights are declared using other market variables. Financial information providers also offer large databases of deal specific information which they use to archive, update and support analyses for interested clients. Banks, investment firms, practitioners and regular businesses then use the data and insights towards better decision making in their respective capacities, whether it be, financing, advising or embarking on M&A ventures. These sources, with their respective dispositions, all provide troves of data with which to explore the industry.

While scholars and practitioners continuously set out to understand the determinants of success within deals, a potentially meaningful contribution lies in better understanding the industry landscape and its drivers of activity. Here a valuable contribution could be made by considering a more comprehensive approach to determinants analysis, in terms of a variety of variables for the field. A thorough evaluation of drivers of past M&A activity would lead to a greater capacity for predicting future trends for the field.

Great potential for analysis in the field lies in the stores of available information and data archived for historical transactions. This may exist in the form of ready developed databases, annual company reports, industry reports and analysis institutions services etc. Further, data sets for drivers are also available through numerous sources in the form of time series and other various formats. The mining and analysis of this industry data, towards characterization and then further, understanding of driver influence on the industry, would

require intelligent data management and analysis methods, able to extract and determine meaningful insights.

Data Science and the applications in its field, have become a popular means for analysing and evaluating causal relationships across a large spectrum of features within data sets. Applications using effective data management and data analytics methods would be well suited in analyses of recent M&A activity and the subsequent drivers, using transaction data from the industry. By making use of such applications on deals data from the industry, a potentially insightful and alternative contribution to the field exists.

1.2 Problem Statement

A number of theories exist for the determinants of M&A activity with few studies finding evidence of categorical explanations. This is further hindered by a lack of consensus on appropriate determinant proxy use, when testing theories empirically. This provides potential for an improvement in the approach to analysis in the field with regard to both the determinants and subsequent characteristics of M&A activity. While stores of M&A transaction data exist in ready developed databases, other driver related data are generally stored separately and not easily consolidated. Given the limited comprehensive undertakings in analyses of determinants and resulting M&A activity, the requirement for an information consolidating and management platform capable of facilitating such an inquisition within the field, arises.

1.3 Aim and Objectives

The aim of this thesis was to develop a data management facility capable of facilitating the analysis of the M&A industry, both in terms of its variables and activity drivers. This database was to have assisted in the formulation and organization of necessary and available information in the field of M&A. More specifically, it was to have facilitated the analysis of anomalies, correlations and trends as well as allowing the evaluation of drivers on this corporate phenomenon. These analyses were to assist in developing insights into the profile of past transaction activities.

Five objectives were established to support the attainment of the research aim, namely:

1. Identify the determinants of M&A activity towards an evaluation of their influence in terms of the industry's characteristics.

2. Obtain relevant and available deal and determinant information appropriate for activity and determinants analysis of the M&A industry.

3. Develop and populate an information management facility capable of supporting the activity and determinant analysis of the M&A industry.

4. Apply relevant data analyses methods towards activity profiling of the industry.

5. Apply relevant data analyses methods towards the evaluation of M&A activity determinants.

1.4 Scope

Broadly, the scope of the research concerns completed M&A deals. Specific attention will be given to deals data for companies from the United States of America (U.S.A), which are listed on public stock exchanges. This will be the domain in which drivers will be tested. Cross border type deals will also be excluded for the additional complexities introduced through conflicts in culture and regulation. The research will be quantitatively focused in the gathering and analysis of available industry and company information for the years 1998 to June 2019.

1.5 Limitations

A major limitation to the effective execution of this thesis was identified as relating to data access and availability. If necessary information required in thesis Objective 2 was not obtained, envisioned analyses of M&A activity and determinants influence would not be supported. Determinants proxy variables gathered and analysed were quantitative or categorical, as the study takes a quantitative approach. Further, the general literature review used both Scopus and Google Scholar as literature sources, while the systematic literature review only uses Scopus as a single source of literature. Additional limitations pertaining to the inclusion criteria for the literature sample gathered were the following:

- Documents published on, or after, 1 January 2019 were excluded.
- Documents were limited to those of source type Article, Conference Paper and Review.
- Documents from the subject areas of Earth and Planetary Sciences and Physics and Astronomy were excluded from results.
- Language of search results was limited to that of English.

Chapter 2

Thesis and Data Gathering Methodologies

The following chapter conveys the methodology and data gathering methods of the thesis. It also presents the list of chapters, detailing the respective contributions to the method. The thesis methodology, in Section 2.1, sets out to convey the systematic approach and rationale for the study. Next, the data gathering methodology is documented through Section 2.2, detailing the search protocol executed for the attainment of relevant merger deal data. Finally, chapters of the thesis are broken down in terms of contributions to the main thesis methodology in Section 2.3.

2.1 Thesis Methodology

The overarching thesis methodology was built using three component methods. Firstly, the systematic quantitative literature review method, as defined by Pickering and Byrne (2014), provided the means for an analysis of literature in the field. Then, the database life cycle (summarized in Figure 2.1) was used as the grounding for the database development, documented in Rob *et al.* (2008). Finally, an analysis application, demonstrating use of the database, stands as the third major component. A schematic representation of this overarching methodology is presented in Figure 2.1, depicting the interactions of component methods. The main stages of the database life cycle are summarized here with further details on the explicit method covered in Section 2.1.3.

A general review of literature in the field was used as an introduction to the field of mergers and acquisitions. This narrative is documented through Chapter 3, Section 3.1. A more structured analysis of literature in the field was then executed, towards thesis Objective 1, utilizing the systematic quantitative literature review method. The stages and respective outcomes of contributions for each stage are documented in Section 2.1.2.

The broad aim of the database was to best facilitate the management of

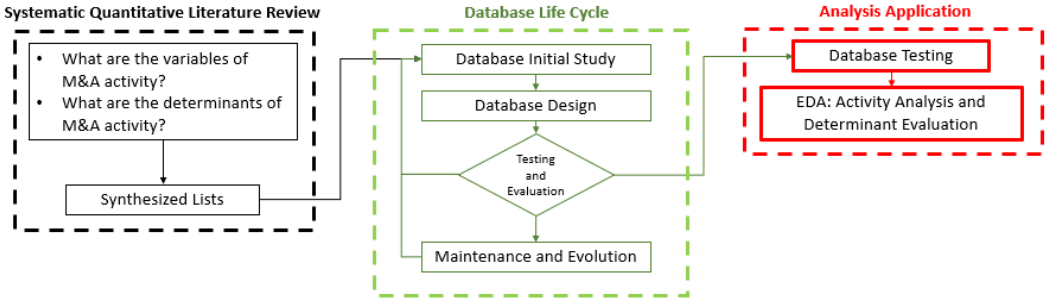


Figure 2.1: Schematic of overarching thesis methodology.

information required for activity and determinant analysis within the M&A industry, as informed by the review of literature in the field. Database development is governed by the Database Life Cycle (DBLC). Design and population was constrained by the data attainable through the deal information sources accessible. Specific stages in the database development method and the relation to the greater thesis methodology are discussed further in Section 2.1.3.

In an application of analysis methods, Exploratory Data Analysis (EDA) methods were used as a means for both testing effective database development and analysing M&A activity. The approach was chosen given its suitability to the problem and data obtained.

2.1.1 Data Warehouse Architecture

After reviewing literature on database management theory, presented in Section 3.3.1, it was decided an appropriate solution for the problem at hand would be that of a data warehouse. Therefore, the overarching thesis methodology was integrated with a typical data warehouse system architecture, explored through Chaudhuri and Dayal (2015) and Gatziu and Vavouras (1999).

This relationship of activities is demonstrated through Figure 2.1.1, where the three component methods of the overarching thesis methodology stand as contributions to the respective sections of the Data Warehouse Architecture. The main integration point is between the Database Life Cycle and Design/Storage section of the Data Warehouse architecture. Here, the product of the database development stands as a data warehouse, a type of relational database. Broadly, the systematic quantitative literature review informs the data acquisition activity, in terms of relevant data and data sources, while the analysis application is facilitated by access to the data warehouse.

Additional steps not included in the overarching thesis method, specific to the data warehousing architecture are within the Acquisition activity. Here, various data sources have their information extracted, transformed and loaded (ETL). While this activity is informed by the review of literature on M&A,

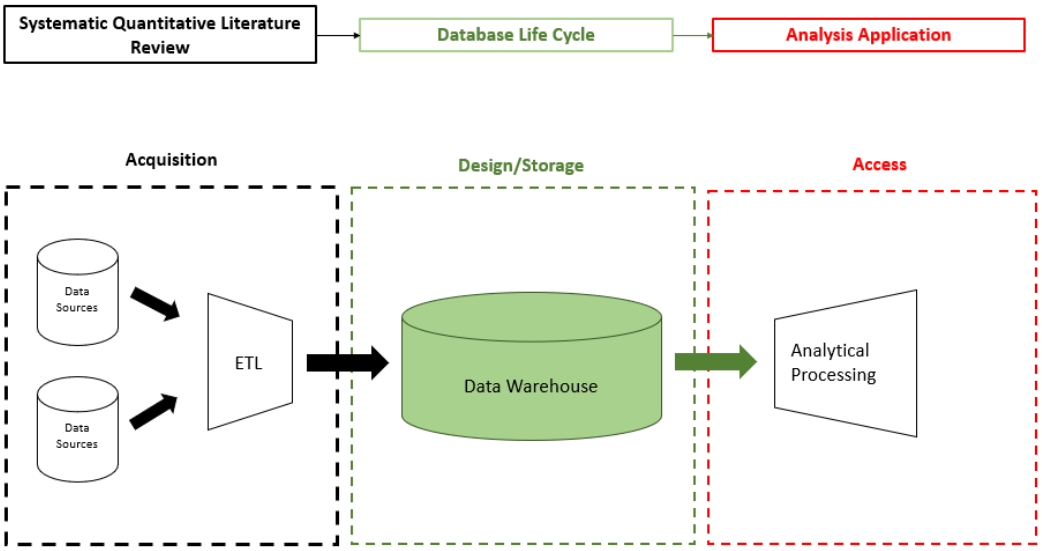


Figure 2.2: Schematic of a typical data warehouse system architecture, adapted from Chaudhuri and Dayal (2015) and Gatziu and Vavouras (1999)

it is constrained by the scope of research and accessible data. Data source information and the extraction protocol are covered in Section 2.2.

2.1.2 Literature Review Methodology

The systematic quantitative literature review process, as defined by Pickering and Byrne (2014), was followed in order to achieve the first overall thesis objective, stated in Section 1.3. Each stage of the process can be seen in Figure 2.1.2, below, while a breakdown and the development of these stages is covered through the succeeding passages.

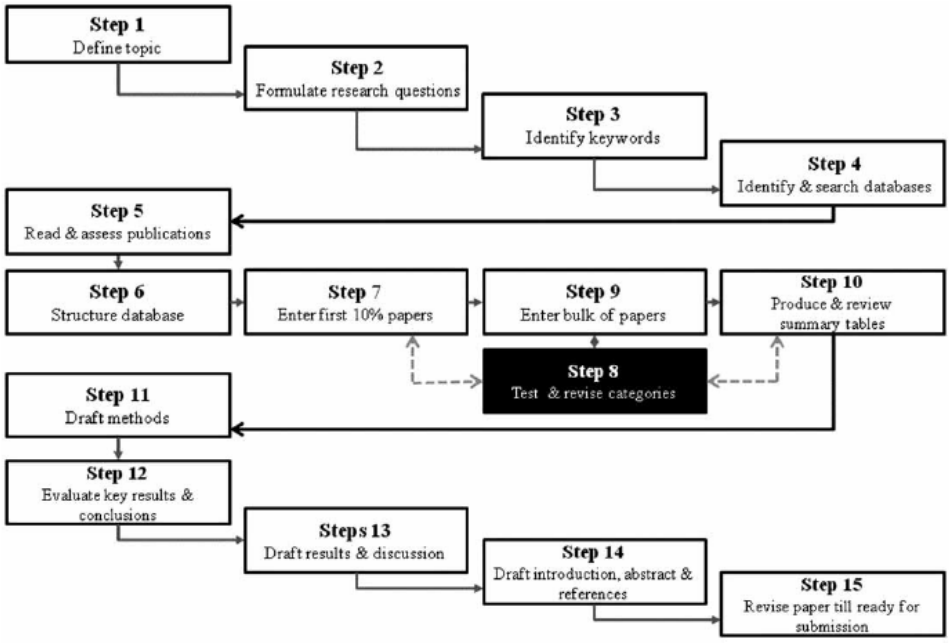


Figure 2.3: Stages of the systematic quantitative literature review - Sourced from Pickering and Byrne (2014).

Topic Definition: The subject of this research deals with the determinants of merger and acquisition activity. In other words, the drivers or forces behind the occurrence of companies purchasing one another. A definition of the research topic is as follows:

The determinants of merger and acquisition activity.

Research Question Formulation: A list of relevant questions were posed to be addressed by the study, as required in step two of the method. The list of questions is as follows:

- *What are the attributes of each research document?*
- *How has M&A activity been explained or measured?*
- *What determinants of M&A activity are identified?*
- *How are the determinants explained or measured?*
- *What methods were used in the analysis of determinants?*

Keyword Identification: Keyword identification was driven by the topic definition of step one and trialling combinations of words that delivered the most relevance in resulting article abstracts. These were identified as follows:

mergers and acquisitions, merger, determinant, driver, wave, activity, deal

Identification & Search of Databases: Scopus was identified as a suitable electronic database for literature collection. Access to the database was attained through the Stellenbosch University Library. This database was used as the sole source of literature in the systematic quantitative literature review method.

Reading and Assessment of Publications: The following subsection defines the method followed towards identifying a final sample set of documents that were to be used in the systematic quantitative literature review method. Broadly, the database was searched for predetermined Keywords using the following search query:

TITLE-ABS-Key (("mergers and acquisitions" OR merger) AND (determinant OR driver) AND (wave OR activity OR deal))

This query searched for a combination of the phrase merger and acquisition or slight variations thereof and the terms determinant or driver, in combination with the terms wave, activity or deal. By default, Scopus searches for the plurals of all search terms. The number of resulting documents was found to be 298.

A criteria was then defined in order to constrain the results to a relevant and finite sample of literature. The criteria is defined in the list below:

1. Documents published on, or after, 1 January 2019 were excluded.
2. Documents were limited to those of source type Article, Conference Paper and Review.
3. Documents from the subject areas of Earth and Planetary Sciences and Physics and Astronomy were excluded from results.
4. Language of search results was limited to that of English.

The resulting search query that satisfied the above criteria was determined to be the following:

TITLE-ABS-Key (("mergers and acquisitions" OR merger) AND (determinant OR driver) AND (wave OR activity OR deal)) AND (EXCLUDE (SUBJAREA, "PHYS") OR EXCLUDE (SUBJAREA, "EART")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "re")) AND (EXCLUDE(PUBYEAR, 2019)) AND (LIMIT-TO(LANGUAGE, "English"))

The number of resulting documents for this search query was found to be 231. After this, article titles and abstracts were scanned for their relevance to the study. This involved searching for articles which specifically pertained to the study of determinants of M&A activity. After this, the sample was found to be 56 documents. Further, documents with a specific focus on determinants of cross-border deals, alone, were then ruled out of the sample, leaving a total of 41 documents. Finally, the sample was reduced again, this time by excluding documents with industry specific studies on the determinants of M&A. The sample was found to be 22 documents.

Two of the twenty two documents were unattainable without necessary access permission while another article was ruled out after it was found to be a continuation of another article in the sample but rather covered research pertaining to performance of M&A deals (19 documents).

It was decided that the inclusion of Harford (2005), "What drives merger waves?", which was not included in the search results as the abstract did not contain the terms determinant or driver, was a worthy addition to the literature sample. This article was included for its high incidents of citing (413) at the time as well as being strongly referenced in many articles from the search sample. The author was also found to be directly involved in contributions to, or editing of, articles within the resulting article sample. The sample of literature then stood to be 20 documents.

Database Structuring: Structuring the database was driven by research questions formulated through step two. A comprehensive list of attributes was named in an effort to best capture information about documents consistently towards satisfying research questions. These attributes were grouped by the themes of document information, M&A activity variables, waves, determinants as well as a study methods.

10% Sample Analysis: Three documents out of the literature sample were entered into the database. This enabled testing of the effectiveness and relevance of the literature database according to the proposed inquisition.

Testing and Revision of Categories: Attributes were further refined and improved, after the 30 initial attributes proved to be insufficient.

90% Sample Analysis: The remaining documents from the sample were then entered into the database according to the newly defined structure.

Method Steps 10 to 15: The presentation of findings from the review can be found in Chapter 3, Section 3.2. Methods of this review were conveyed through the current section of the overall thesis methodology. Key results and conclusions can be found in a discussion of the results and findings, in Section

3.2.4. Steps 13 to 15 were found to be more specific to the development of a research article but are relevant to the methods of this thesis nonetheless. However, they are not exclusively documented throughout the thesis but rather interwoven within the main thesis introduction, methodology and literature review presentation.

2.1.3 Database Development Method

The Database Life Cycle (DBLC) was chosen as a suitable database development method. The model is an adaptation of the Systems Development Life Cycle (SDLC), documented by Rob *et al.* (2008). The method presents a well defined sequential set of steps to be followed in a database development exercise, where the output of each step becomes the input of its successor and a feedback loop allows for redesign and correction after testing and evolution of the database. The method was appropriate as it integrated effectively within the broader thesis methodology and Data Warehouse Architecture. Figure 2.4 illustrates the models steps, while execution is documented in Chapter 4. Descriptions of the respective stages follow in the passages below.

The database's initial study was driven by thesis Objective 3 as well as the main outcomes of the literature review, the synthesized list of activity and determinant variables. Figure 2.1 shows the connection within the overarching thesis methodology. This information served as the primary statement

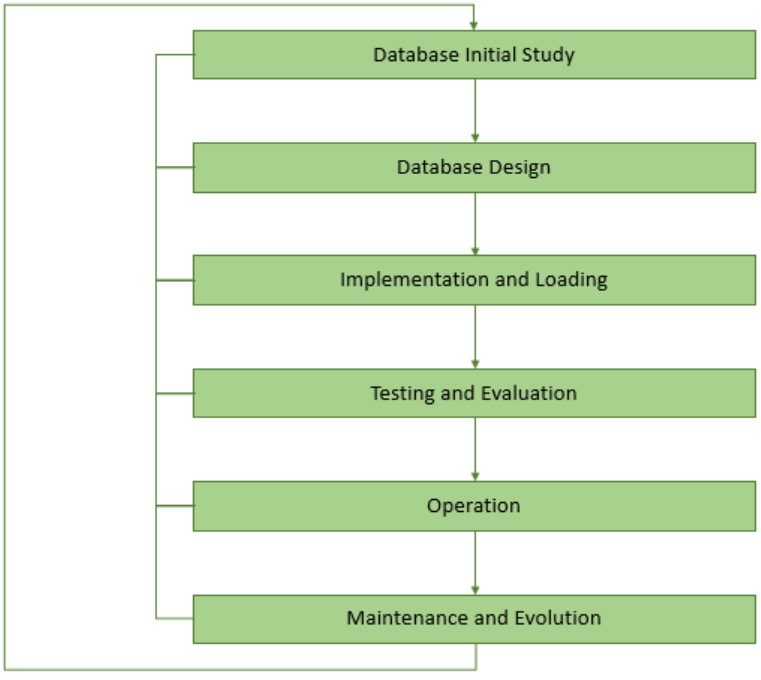


Figure 2.4: The Database Life Cycle, adapted from Rob *et al.* (2008).

of requirements for the database and was constrained by variables that were attainable, considering available information sources, helping define the envisioned solution's scope. Section 4.1 documents this analysis of requirements.

Database design was driven concurrently by normalization of available deals data, Section 4.3.1, and using entity relationship modelling, documented through Section 4.3. This, after requirements were analysed in the initial study stage. These models were then translated into a logical schema, documented in 4.3.4, that detailed all tables, data types and relationship constraints required for description of the conceptual design.

Implementation involved software selection and practical design using the chosen database management software. Practical design translated the logical schema into SQL code that defined tables and entity relationship constraints. In line with the extraction, transformation and loading stage of the Date Warehouse architecture, Figure 2.1.1, data was pre-processed, documented in Section 4.2, before data normalization continued. After this, data was loaded from the corresponding normalized tables, stored in Microsoft Excel sheets.

After effective testing and evaluation, operational status of the database was achieved by allowing for an application of analysis on the data. This also further provided the opportunity to test effective implementation, by comparing results using different software for the same desired analyses.

The database became operational after testing and evaluation had been successfully completed and facilitation of analysis capabilities achieved. This was done using Open Database Connectivity (ODBC), an application program interface used to connect to database management systems.

Explicit maintenance and evolution activities were not performed and documented for this thesis. However should improved and or additional activity or determinant variables become available, a repetition of the life cycle would be required, building from the established initial study, when considering proposed changes to the existing database.

2.1.4 Analysis Application

An analysis application allowed for the testing of effective database development and implementation. Further, an exploratory data analysis approach was utilized towards satisfying thesis Objectives 4 and 5. Graphical representations, statistical summaries and time series analysis methods were used in analysing merger activity. Determinant analysis and evaluation was supplemented by methods in time series analysis, correlations, multiple regressions and feature selection methods.

2.2 Data Gathering Method

The greatest effort was made to attain deal and determinant data that matched both the scope of study and requirements of the systematic quantitative literature review. The deal set criteria was primarily informed by the scope of study, while determinants were gathered based on the outcomes of the literature review. The following subsections detail the method and sources of data gathered, towards Objective 2.

2.2.1 Deals Data

The study scope became the primary consideration for a deal set inclusion criteria, thus an appropriate search protocol was developed towards attaining the most representative set of deals. M&A transactions were to be announced and completed between 1998 and June 2019, where both target and acquirer companies were from the U.S.A and listed on a public exchange. Further, deals were to involve a single entity target and a single entity acquirer, reducing the additional complexity inherent to multi-entity deals in terms of shares and control assumed. Table 2.1 details the resulting number of deals returned with respect to each criteria constraint. A total of 4 940 deals were returned as results for the protocol, using the Bloomberg L.P. (2019*a*) advanced deal search function, where-after, 112 multi-party acquirer or target deals were removed from the Microsoft Excel workbook of deal results.

Table 2.1: Deal set search criteria.

Variable	Constraint	Deals
Country	United States Target AND Acquirer	207 302
Deal Status	Completed	196 440
Announcement Date*	1 January 1998 - 30 June 2019	194 396
Deal Type*	M&A	147 518
Public/Private	Public Target AND Acquirer	5 473
Deal Size	Minimum 1 million USD	4 940
Acquirer & Target	Single Entity	4 828

* Announcement dates were within specified range while deal completion criteria ensured deals were completed before 30 June 2019. Deal types excluded were that of investments (mostly minority stake purchases), joint ventures, spin-offs and buybacks.

2.2.2 Determinants Data

Determinants data were gathered to meet the outcomes of the systematic quantitative literature review, aimed at identifying drivers recognized and studied

in M&A, traditionally. The obtained set of suitable proxies obtained through data gathering is presented in Table 2.2. This table provides the designated Bloomberg L.P. (2019*b*) code of each determinant variable (excepting the S&P 500 PE Ratio) and its description. All data was sourced from Bloomberg L.P. (2019*b*) except the S&P 500 PE Ratio which was obtained through Multiple.Com (2019). Data was gathered for years 1989 until June 2019.

Table 2.2: Obtained determinant variable proxies of M&A activity.

Code	**Determinant Description**
MUNRTAX_Index_(USD)_(L3)	Implied Interest Rate
USGG10YR_Index_(R2)	Ten Yr Government Bond Yields
USGG2YR_Index_(R2)	Two Yr Government Bond Yields
XAU_Curncy_(R4)	Gold Spot Price
SPX_Index_(R4)	S&P 500 Index
USRINDEX_Index_(R1)	Recession Indicator
CPI_YOY_Index_(R2)	NSA CPI YOY Index
DXY_Curncy_(L3)	US Dollar Spot Index
PRIME_Index_(R2)	Prime Interest Rate
LF98TRUU_Index_(R4)	Corporate High Yield Tot Return Index
FEDL01_Index_(USD)_(L2)	Federal Funds Effective Rate
CFNAI_Index_(USD)_(L1)2	Chicago Fed National Activity Index
M1_Index_(USD)_(R1)2	M1 Money Supply
M2_Index_(USD)_(R1)	M2 Money Supply
EPUCNUSD_Index_(USD)_(R4)	Economic Policy Uncertainty Index
EPUCTRAD_Index_(USD)_(L4)	Trade Policy Uncertainty Index
IP_CHNG_Index_(USD)_(R2)	Industrial Production MOM %
	S&P 500 PE Ratio
EHUPUS_Index_(USD)_(R2)	Unemployment Rate %
GDP_CQOQ_Index_(L2)	GDP QOQ %
SPEQPOSS_Index_(R3)	S&P EPS + Surprise
1119C01_Index_(USD)_(L3)	Real GDP IMF
1119R014_Index_(USD)_(R3)	GDP QOQ % IMF
BANBT11_Index_(R2)	Chapter 11 Bankruptcy Filings
BANBT12_Index_(R1)	Chapter 12 Bankruptcy Filings
BANBT13_Index_(L1)	Chapter 13 Bankruptcy Filings
BANKTOTL_Index_(USD)_(R1)	Total Bankruptcy Filings
EFFIUS_Index_Last_Price	Fiscal Freedom Index
HSANNHSP_Index _(R4)	Annual Housing Starts

2.3 Chapter Breakdown

The thesis is presented through six chapters. Firstly, an introduction to the domain and definition of the problem as well as consequent aims and objectives for the thesis, are detailed in Chapter 1. Methods employed in facilitating problem solving requirements, as well as the data gathering protocol, are documented in the current constituent, Chapter 2. A literature study, covering deeper exploration of the domain and results of the systematic quantitative literature review, are presented in Chapter 3. Additionally, a study of appropriate theory and methods in Data Science aid the problem solving capacity for the thesis. Chapter 4 conveys the database development stages executed in producing the information facility capable of supporting M&A activity analysis. Results for database testing as well as the analysis of merger activity and determinant evaluation are covered in Chapter 5. Finally, thesis conclusions, limitations and recommendations for further research are presented in Chapter 6.

Chapter 3

Literature Review

The following chapter sets out to present a review of literature in the field of M&A. An introduction to the field, guided by attributes of the phenomenon is delivered through Section 3.1. This section will stand as a foundation for understanding the profile of industry activity in terms of its characteristics. A general exploration of literature was executed in order to achieve this, using generic searches for articles from Scopus and Google Scholar. After this, findings of the Systematic Quantitative Literature Review are presented in Section 3.2. Here a more structured exploration of literature supported the identification of determinants of M&A activity and a better understanding of relevant theory. Additionally, Section 3.3 documents coverage of the approach and methods required to facilitate the study. This contributes to the development of a database as well as analysis methods to be used in activity and determinant analyses.

3.1 The Field of Mergers & Acquisitions

M&A activity can be well characterized through coverage of the motives, deal types and process itself. These aspects of the field, along with other important characteristics, are explored through the following subsections.

3.1.1 Definition

The terms, merger and acquisition, are often used interchangeably, nevertheless, distinctions can be made between the two words. While both mechanisms lead to corporate restructuring, a **merger** is commonly understood to indicate the bonding of two enterprises of roughly the same size, combining resources and management, with original shareholders of each party receiving a portion of the new entity. An **acquisition** on the other hand, describes the change in control and management of an enterprise after another entity seeks to take command of the entity. Controlling stake then becomes the major distinction

between a merger and an acquisition. Further, it can be assumed that a merger results from neither company taking the role of the acquired or acquirer in a deal. The stake acquired in an acquisition does not always have to be a controlling one though. A minority share of an enterprise can be acquired as well. Acquisitions are then broken into full or partial: Full being a 100% acquisition of equity and partial being above 50% but below 100%, usually (Coyle, 2000). Deeper coverage of specific M&A deal types will be covered in Section 3.1.3, where further distinctions can be identified.

3.1.2 Company Growth and Deal Rationales

An enterprise usually exists for the sole purpose of achieving a greater financial objective. This can be translated into creating best shareholder value for the entity, or an improving return on assets or investment etc. The financial objective then usually involves operational objectives or a mission statement, set out as a means for achieving it. A firm then looks to establish strategies to implement towards attaining the operational objectives (Coyle, 2000).

Strategies towards growth can be seen as a Key activity that an enterprise could implement if operational and further, financial objectives, were to be attained. A number of strategies exist for this, but firms often have to decide between two major mechanisms. Enterprises can pursue organic growth (an internal means of growth) or M&A transactions (Build vs. Buy or Exploit vs. Explore). This, as a strategy towards surviving and thriving in their respective markets. Avenues such as these should be developed and nurtured on parallel platforms and not exclusively, according to Coyle (2000). Selecting acquisition transactions appropriately, and for the right reasons, allows an entity to execute growth strategies, ensuring the attainment of major financial objectives (Coyle, 2000).

Growth should not always imply an increase in the size and value of an enterprise's operations or assets. A company could also grow/improve by shedding size. Here, a company could relieve itself of burdensome business units or assets that may be misaligned with strategic objectives or hinder bottom line performance and overall efficiency. In this case, the resulting sale of the asset, known as a divestiture, would play out as a typical M&A deal. From the perspective of a buyer, "one man's trash becomes another man's treasure". Specific rationales for undertaking deals are explored further in the following passages.

There are many rationales that enterprises use to justify undertaking an M&A deal. The categorization of which can often come down to interpretation of similar seeming motives and allows for subjectivity in the matter. No clear consensus seems to exist in literature for a defined set of categorical rationales that drive deal making.

Malik *et al.* (2014) declares four major motives for corporate deal making. These include a drive for synergy, the agency motive, management overcon-

fidence and efficiency gains. The agency motive involves a pursuit of M&A transactions by management against the intention of shareholders. A conflict of agencies arises when these managers pursue projects out of their own interest using free cash flow, instead of rewarding shareholders. While management overconfidence may be transparent in interpretation, it is important to note that this motive is attributed to managers who may be arrogant and regard entity success as a direct result of their own ability, prompting bold bids on ultimately under-performing deals. It may seem from the outset that synergy and efficiency gains are strikingly similar motives (highlighting interpretation subjectivity of motives). An important distinction is made between the two though. Synergy is seen as the increase in value and returns of two entities through coming together, over and above the summed value of the each entity independently. An analogy of the phenomenon is the "1 + 1 = 3" effect (Coyle, 2000). Synergy prospects are seen as a major motive for M&As, coming in the form of both operational and financial gains in business functionality. Efficiency on the other hand leans towards a firm's technical capabilities or operational capacities, where maximizing of outputs using minimum combined resources is the sought after outcome. This can, however, be achieved using measures other than M&A (Malik *et al.*, 2014).

Golubov (2012) also groups motives into four main categories. These being synergy motives, agency motives, managerial overconfidence and other M&A motives. Importantly, resulting efficiencies are grouped within the synergy motives, where they are broadly, operationally or financially implicated. Synergies can be attributed to outcomes such as reducing redundant business units, effecting economies of scale and scope, reduction in cost of operation and of distributing management capabilities. Other sources of synergy may come from revenue enhancements, diversification of product/service offering as well as cross over of management capabilities and practices or through strategic acquisitions of technological assets and intellectual property. Financial synergies are driven mostly by tax shield opportunities and utilizing remaining debt capacity, while operational synergies are driven by a reduction in investment as opposed to enhanced profits through larger operations (Golubov, 2012).

Agency motives are said to be driven by compensation and incentive hungry management of an enterprise. Usually investment activity and firm growth drive compensation and therefore, management embark on acquisition activity over and above dividend payout to shareholders. The drive for managing a larger enterprise for pride and fulfillment can lead executives into making blind, overconfident investments in diversifying operations. These motives can often lead to the abjection of enterprise value and therefore shareholder value. Managerial overconfidence follows a similar theme and can result in the same outcome, ultimately: decreased shareholder value. This is discussed similarly by Golubov (2012).

Other M&A motives identified by Golubov (2012) move towards more externally influencing factors. These come in the form of industry shocks and

market (miss)valuations. The implications of miss-valuations can be beneficial to an acquirer, both in terms of financing option strength as well as the actual value of a deal. Investment banks are also seen to play a role in stimulating transaction activity for the incentive they receive in financing and advising on deals (Golubov, 2012).

Motis (2007) categorizes deal rationales by two major groups. The distinction is created between deals aimed at increasing value and profits for shareholders and the deals that are pursued through the interests of a manager. Within shareholder gains, efficiency gains, synergy gains cost savings, financial cost savings, enhancement or strengthening of market power, pre-emptive and defensive and disciplinary takeovers are separated as several sub-categories for increasing enterprise value. Empire building, hubris and risk spreading or diversification are sub-categories of motives for managerial gains. Within these sub-categories, multiple, more descriptive motives, are given (Motis, 2007).

Motis (2007) also moves on to cover empirical evidence to substantiate the measurement or attempted measurement of the above mentioned motives and reiterates the difficulty of understanding the complex interrelation of factors affecting performance measurement of the outcome of these motives.

Beyond categorization of common M&A deal motives, firms may often disguise the real rationale for deals with more pleasing ones aimed at maintaining the trust and confidence of shareholders. Managers will take advantage of the subjectivity in interpretation of the true nature of a deal, inflating desired outcomes, for greater shareholder approval. This can be particularly beneficial in deals involving share-swops as a form of payment as the inflated or undervalued share prices will significantly affect the deal value. This could be a strong symptom of management hubris.

3.1.3 Deal Types

Deal types as defined by Bloomberg L.P. (2017) and other sources will be discussed in the passages to follow. This particular variable stands to be an important characterization of M&A deals as it eludes to more specific information about the deal. Deal types can also be attributed by specific characteristics of target and acquirer entities or economic factors. Investigating such relationships could lead to the identification of determinants.

In a **Company Takeover**, full ownership and control is acquired through the purchase of 100% or a majority share (taking its share up to 100%) of the outstanding shares in a target company (Bloomberg L.P., 2017). The target becomes fully acquired.

An **Institutional Buyout (IBO)** is another deal type phenomenon. These deals have institutional investors who set out to acquire a majority share in a company and in so doing, assume a controlling interest within the enterprise. Typical parties would be the likes of investment banks, pension and insurance fund management.

Another buyout termed, **Management Buyout** or **(MBO)**, is an active attempt to take ownership of an enterprise, driven by organizations management body (Bloomberg L.P., 2017). The investment returns, through ownership of the firm, are sought to be more beneficial than regular compensation packages.

In a more aggressive situation, **hostile takeovers**, come as a result of tender offers or proxy fights. The acquiring entity seeks to bypass management as well as the company's board of directors and approach shareholders of the target entity directly. This, or seeking to remove management outright, in a bid to have the acquisition approved without any resistance. A hostile takeover occurs as a consequence of the resistance from the target's existing management, after a deal is proposed and executed (Bloomberg L.P., 2017).

A **Tender Offer** occurs in a bid to purchase some or all of shareholders' shares in a corporation. The price offered is usually at a premium to the market price. This is typically offered directly to shareholders (Bloomberg L.P., 2017).

A **Cross Border** deal involves at least one party (target, acquirer, seller) of a different country of risk. A country of risk is determined by considering where, the location of primary exchange listing, management offices and reporting currency, is (Bloomberg L.P., 2017).

When a target company is de-listed and no longer traded publicly after a deal, it is termed as a **Going Private** deal. The shares of the target are held privately after this transaction. Not similarly to **Going Private**, a **Private Placement** deal involves the exclusive issue of new shares in a target company to that of an acquirer, rather than using a public offering (Bloomberg L.P., 2017).

Majority Purchase deals involve a share acquisition that increases shares in a target company to above 50%. The acquirer then attains a controlling influence over the target companies interests.

A **Squeeze out** deal is attributed to a transaction whereby an acquirer attains the last outstanding shares in a target in which it already owns a majority stake. This is done by means of a **Tender Offer**. If it were not for the original majority stake and **Tender Offer**, the deal would otherwise be a **Minority Purchase** of a company.

PE Buyout is a deal in which a a private equity firm acquires a majority, if to 100% share of equity within a target company. The firm then attains a controlling stake in the target company (Bloomberg L.P., 2017).

Reverse Mergers are an atypical deal type in the sense that the opposite of what is expected, occurs. Here, the target company becomes the acquirer as its operations become the sole activities of the original acquirer or the target becomes the majority shareholder in the acquirer as an effect of the transaction (Bloomberg L.P., 2017).

3.1.4 Integration Types

Different forms of integration occur as a result of M&A deals. These can be broadly categorized as **Vertical**, **Horizontal** and **Conglomerate Integration**. Each have different implications for the resulting combination of entities involved in a deal and are pursued to satisfy a specific deal rationale. Categorizing deals according to each type of integration depends on the existing product or service offering and relative sector placement of each entity within a transaction.

Vertical integration occurs when a company merges or purchases a controlling stake in adjacent entities within its value chain. This may be to decrease transportation expenses, improve efficiencies or reduce costs and come in the form of forward or backward integration within the value/supply chain.

Horizontal integration involves an expansion of business activity within the same sector but at the same level. Through this mechanism, the entity can increase capacity, increase market share as well as share resources.

Conglomerate integration is a diversification of product or service offering by a company. The entities involved in this type of deal have business units or operations that are completely unrelated. Conglomerate integration can either be a pure or mixed conglomerate. Here, a pure conglomerate brings firms with no similarity together while being mixed, look for product or market extensions. It should be noted that in some cases a merger can be in more than two categories. This would likely involve a conglomerate target and acquirer. (Schmidt, 2013; Motis, 2007).

Therefore, integration types depend on the orientation of merging entities to one another in terms of their position in the supply chain and economic sector in which they operate, respectively. Vertical integration occurs in the same supply chain. Horizontal integration is confined to the same sector. Conglomerate integration occurs out of the same sector and supply chain. Integration types can also stand as another variable of broad classification for deal type of a specific transaction.

3.1.5 Transactions as a Process

It seems from literature that a typical M&A venture should be executed according to a process or recipe of sorts, to be followed for intended success. Regardless of whether or not an entity achieves its desired outcomes though, it is logical that there is a clear strategy and process of some type adopted for a transaction, especially given the complexity of the task. Different types of M&A process constructs were explored in terms of their different phases and stages, in order to gain a better understanding of what literature interprets a typical transaction as.

No concrete consensus exists for definite boundaries or distinct phases in a typical M&A process. The characterization and frequency of phases varies

for many sources, with a number of models posited to represent a transaction. Most simplistically however, a fundamental distinction can be made between the time that a target remains independent and when the target surrenders ownership to the acquirer. This change in states can be regarded as the **pre-acquisition** and **post-acquisition** phases of a deal. Importantly this process defines a state after which integration can be undertaken (Gomes and Angwin, 2012).

Many models for the M&A process perspective exist, having numbers of phases that range from two to several. Similar to Gomes and Angwin (2012), Digeorgio *et al.* (2002) and Digeorgio (2003) identified two phases for success in M&A as **front-end** and **integration** success. Two of the larger components of front end success involve selecting the right target for merger or acquisition success and selecting the best transition structure base on the type of combination. Integration success is driven by achieving a successful combination of objectives. These larger components for each phase require a host of activities and checks that prepare a phase for success (Digeorgio *et al.*, 2002; Digeorgio, 2003).

Beyond two phase models, a slightly more detailed perspective, identified by Salus (1989), involved a three phase model. This included the **pre-merger**, **merger** and **post-merger** phases (Salus, 1989). Although avoiding over detailed and complex models may be wise given the inherent complexity of M&A deals, it may be beneficial to find lower level perspective representations of deal processes. A move in this direction came from Haspeslagh and Jemison (1991) in their four phase model: **idea**, **acquisition justification**, **acquisition integration** and **results** (Haspeslagh and Jemison, 1991). Carpenter and Sanders (2006) expand on this model by including necessary due diligence and deal negotiation as an addition to the justification phase (Carpenter and Sanders, 2006).

Koerner *et al.* (2014) goes on to cover phase models of incrementally increasing detail. These entailed processes of five to seven phases with varying levels of complexity. It is noted that there is great difficulty arising in the study of M&A transactions as a process. This is owing to the inconsistencies in literature on clear cut lines between start of phases and end as well as the timing of phases. Further, there is a lot of non-linearity of phases through the process (Koerner *et al.*, 2014).

3.2 Systematic Quantitative Literature Review

In an effort to identify perceived determinants of M&A activity, a review of literature in the field using the Systematic Quantitative Literature Review method, as specified by Pickering and Byrne (2014) and detailed in Chapter 2, was executed. This exploration and gathering of relevant literature not only allowed for the identification of determinants, but also an understanding of

the common methodological approaches used for studies in the field. It also allowed for a demographic breakdown of the resulting sample of papers.

Findings from the literature review are broken into three subsections that address the research questions posed in Step Two of the Systematic Quantitative Literature Review method. Firstly, attributes of the sample documents are analysed, such as where and when articles were published as well as research method specific characteristics of papers. After this, analyses of subject specific theory on M&A activity and the determinants thereof, are presented in the succeeding subsections.

3.2.1 Sample Analysis

Following the method defined through Section 2.1.2, the resulting sample of literature was found to be 20 documents. According to Scopus's cite analysis function, at the time of writing, these articles had been cited by 695 documents throughout the Scopus database.

Figure 3.1 shows a breakdown for the number of documents published per year within the literature sample. The graph suggests an increase in interest on the research topic from 2009, after an extended period of little contribution.

Figures 3.2 and 3.3 detail further demographic statistics of the sample. In terms of subject area breakdown, 17 documents are classified in the Economics, Econometrics and Finance area, while 11 documents of the sample are classified in the Business, Management and Accounting category. This shows a bias in the sample towards studies from these fields. Just two documents can be classified by an alternative category such as Social Sciences. It must be noted that a subject area classification is not exclusive and can have multiple categorizations, creating the observed overlap. Figure 3.3 shows the geographical spread of documents based on their main authors origins. The sample is clearly dominated by contributions from the U.S.A with the next most fre-

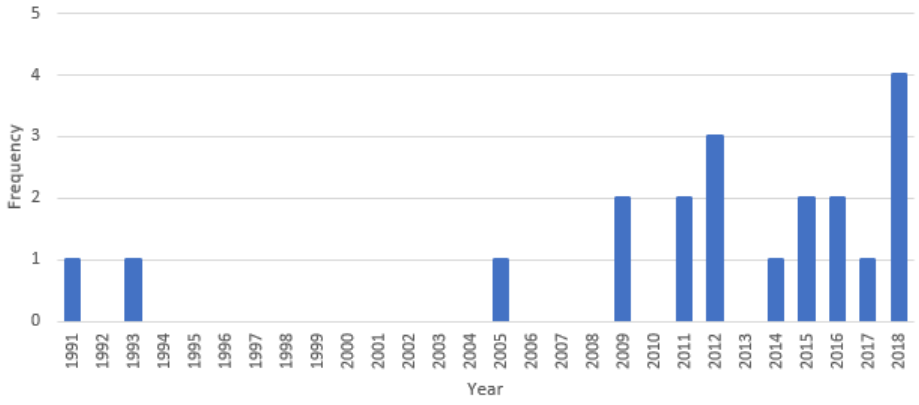


Figure 3.1: Number of documents published per year in literature sample.

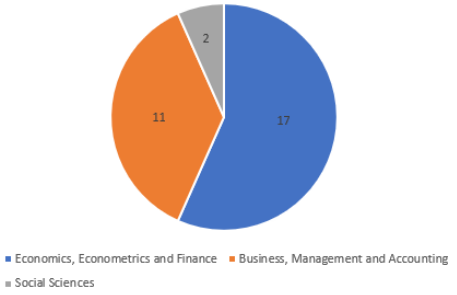


Figure 3.2: Document subject area breakdown.

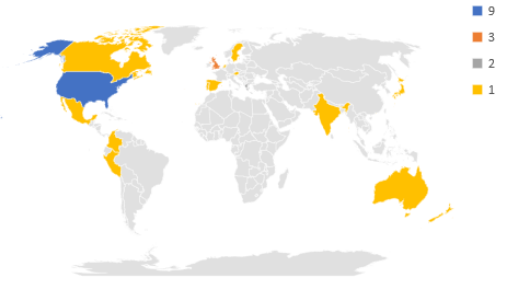


Figure 3.3: Geographical spread of document origin.

quently contributing author country being the United Kingdom, having three documents.

From a methodological perspective, the majority of papers take a quantitative approach, contributing empirically to studies in the field. Yaghoubi *et al.* (2016) is found to be the only paper to take a qualitative approach through a thorough narrative review of existing literature on determinant theory. 18 papers use various regression models in some way or another to evaluate drivers. Combinations of independent variables, as proxies for various determinants, are compared in terms of significance and relationships, by evaluating the models output coefficients. Moschieri and Campa (2018) uses descriptive statistics as a means for evaluating policy making and its effects on M&A activity characteristics across the European Union, also considering institutional attributes of companies in the various countries of its heterogeneous market place.

Thomson Reuters Corporation (previously Thomson Financial) were the popular choice for data sources used throughout the sample, specifically the Securities Data Corporation (SDC) databases. A total of nine papers use some version of SDC as their source for M&A transactions data. Another five articles use other Thomson Reuters products such as the Thomson One Banker database. Other sources are then used to supplement the primary deals database to provide more information for the studies. These are usually for financial and company information as well as other specific economic indexes. Examples of sources and companies that provide such databases include: Bloomberg, Zephyr and Amadeus from Bureau Van Dijk, the Federal Trade Commission, Federal Reserve Economic Data, Standard and Poor, Global Vantage, Chicago Fed National Activity Index, Nikkei Economic Electronic Database.

After gathering data on the window periods considered in studies throughout the literature sample, the 1990's and 2000's are seen as the most frequently observed years of merger activity. This is evident in Figure 3.4, where studies consider years as early as 1919 through to 2015. It should be noted that both studies that considered the early 1900's period were from the same article, Benzing (1991), where differing window periods were used in evaluating deter-

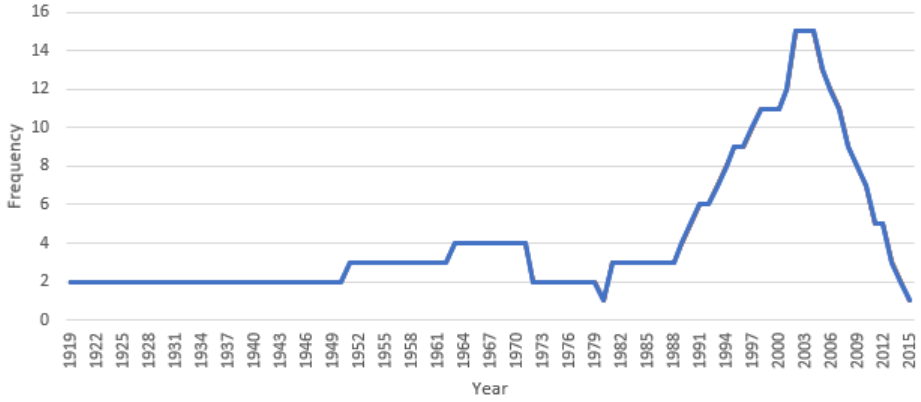


Figure 3.4: Frequency of years under examination.

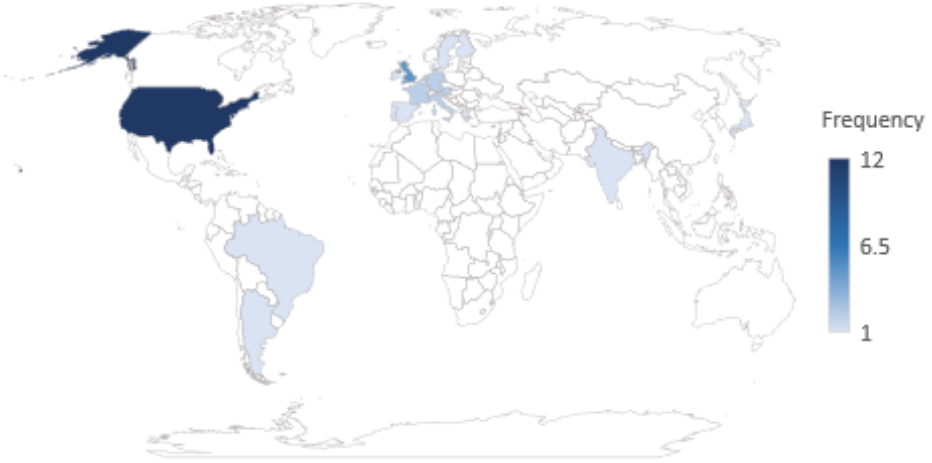


Figure 3.5: Concentration of countries under observation in sample studies.

minants using different models. A steady decline in observations for the years of 2005 and onward is also evident in Figure 3.4, while the average window period considered was 17 years in length.

In terms of countries observed in studies, the U.S.A is found to be the predominant subject of M&A determinant analysis and testing. Of the 22 countries observed in total, 12 articles observe M&A activity of the U.S.A alone, or in combination with other countries, while deal data from the United Kingdom are used five times. Figure 3.5 shows the concentration of countries under observation in the sample. Deals data from the European countries of the Netherlands, Austria, France, Germany and Italy are all used in two articles, proving the significance of developed country economies as the sample for studies.

3.2.2 Activity & Wave Theory

As mentioned in Section 1.1 of Chapter 1, trends in M&A activity, known as waves, have occurred throughout history (Cartwright *et al.*, 2012). Cycles of heightened and reduced activity have been observed as far back as the late nineteenth century. Authors typically refer to the waves of M&A activity as a consequence of determinants. While waves have usually been characterized by considering the frequency, volume and types of deals, this section aims to explore alternative attributes of activity identified through the literature sample.

14 papers make reference to, or acknowledge, waves as a description of trends in M&A activity. Six major merger waves were identified and covered by the sample. The era, characterization, reason for closing and major sectors involved in each wave can be found in Table 3.1. Interestingly, the exact starting and ending years of waves are found to differ between sources, indicating possible elements of subjectivity on the identification of waves. No information was found for a seventh wave of M&A activity and whether it has begun or not.

Methods of wave identification vary from visual interpretation to statistical means. Polemis and Paleologos (2014) identify waves graphically within a 26 year sample period of data for deals in the banking industry of the U.S.A in years 1987 to 2013. Harford (2005) and Cortés and Agudelo (2017), both use a method, defined by Harford (2005), of detecting waves by finding 24 month periods for which actual activity concentration exceeded the 95th percentile of a simulation of 1000 distributions of the same number of deals, in the same observed period, for each industry considered in the study. Gugler *et al.* (2012) used the switching method, initially used to recognize periods of economic recession, later used to test for merger waves. This time-series method determines the state of activity (either wave or non-wave) based on the maximum likelihood which is calculated using estimations of an auto-regressive model and state averages, for four lag terms.

Table 3.1: Historic M&A Waves: Information gathered from Yaghoubi *et al.* (2016); Cortés and Agudelo (2017).

Wave	Period	Characterization	Closing Reason	Major Industries
1	1897-1904	Horizontal Integration/ Monopolization	Anti Trust Laws/WWI	Steel Production, Hydraulic Power, Textiles, Telephone, Rail Road, Light and Power
2	1910/ 1916-1929	Vertical Integration/ Oligopoly Formation	Great Depression	Food, Steam Engines, Steel, Railways
3	1950/ 1965-1969	Diversification/Conglomeration	Oil Crisis Driven Recession	Electricity, Chemicals, Combustion Engines
4	1981-1989	Hostile and Leveraged Takeovers	Not Found	Oil & Gas, Textiles, Misc Manufacturing, Non-depository Credit, Food
5	1993- 2000/2001	Consolidation of Major Industries/ Deregulation/Stock Payments	Economic Recession	Metal Mining, Media & Telecom, Banking, Real Estate, Hotels
6	2003-2007	Global Scope/Cross-Border Cash Payments	Economic Recession	Banking, Media & Telecom, Utilities

Considering the information gathered on waves in Table 3.1, it is apparent that periods of economic recessions are deemed strong antagonists of merger activity. A notable observation in terms of the implicated industries, is the shift from predominately primary and secondary sector industries in the early to late 20th century, to tertiary sector industries in the 1990's and 2000's.

For studies in the literature sample, M&A activity is predominantly described in terms of value and volume, over monthly, quarterly or yearly intervals. Models mostly use these variables as a dependent variable when testing for the influence of determinants. Beyond value and volume, Table 3.2 details additional activity variables encountered within the literature sample. Most variables could be expressed in terms of total value or volume as the main metric of quantification, or a set of summary statistics in the predetermined metric of the variable, for the sample of deals in question.

Table 3.2: Firm and deal specific variables of M&A activity.

#	Firm Specific	Deal Specific
1	Firm size (number of employees)	Deal volume (Number of Deals)
2	Firm size (total assets)	Deal value
3	Economic sector	Deal type
4	Industry (SIC Codes)*	Payment type
5	Share ownership (% Change)	Completion status
6	Share ownership concentration	Premium
7	Share price	Payment currency
8	Age of a firm	Time to completion
9	Age of CEO	Domestic or cross border*
10	Domicile (country, state etc.)	Integration type*
11	Industry competitiveness	
12	Listed status	
13	Total sales	
14	Total debt	

* Standard Industrial Classification (SIC). Domestic or cross-border and integration type are variables that are directly derived by firm specific attributes of the companies involved in the deal. A cross border deal is determined by each firms respective country of domicile while integration type is determined by the respective industry and sector classifications.

The activity variables of Table 3.2 have been grouped as either being firm or deal specific. This was a logical way of organizing information based on the origin of these variables. A good way to understand this separation is to consider what information would be available if the deal between two companies was never to occur. If this was the case, then all deal specific variables would not exist.

Share Ownership (% change) is implied when the acquiring company has a minority stake in the target before an acquisition (majority stake acquirement). While share ownership concentration conveys the number of shareholders, to whom shares in the former companies were owned by. Moschieri and Campa (2018) find lower frequencies of hostile takeovers among firms within Germany, France and Spain, as the number of shareholders commonly decreases for those countries and the decision making rests on a small contingency of shareholders. Industry competitiveness is a measure for the number of firms within an industry. Chidambaran *et al.* (2018) use a normalized Herfindahl-Hirschman index, that factors in proportional market share, as a way to measure the competitiveness within industries.

3.2.3 Determinant Theory

Reasons for the occurrence of M&A activity waves are explained by two dominant theories, the Neoclassical and behavioural theories of mergers. Eight articles of the literature sample recognize these theories as an explanation for merger activity.

The Neoclassical theory declares that shocks cause M&A waves. These can be disruptions in the form of economic, technological or regulatory change. Managers are deemed to make acquisitions in the best interest of shareholder value, driven by the proposed synergies that could result from a deal (Yaghoubi *et al.*, 2016). The ensuing M&A activity pursued by enterprises, creates a wave. This theory was supported by the empirical study undertaken by Harford (2005) who noted, however, that not all shocks produce a new wave and that the necessary capital liquidity should be available to lubricate new activity. Cortés and Agudelo (2017) emphasize the market's inherent tendency toward an efficient state, complimented by the necessary reallocation of assets, through acquisitions, after an industry shock.

The second theory proposes that over-valuation of markets are exploited through M&A activity. Here, deviation in valuations allow management to leverage themselves better in the acquisition of typically under-valued or "not as over-valued shares" (real assets) of other enterprises. According to this theory, M&A activity increases during flourishing periods on stock markets. Therefore, intuitively, increased activity is driven by market over-valuation. Importantly, M&A activity correlates to stock market performance and in turn produces waves (Fuller and Pusateri, 2018; Komlenovic *et al.*, 2011; Yaghoubi *et al.*, 2016). However, Fuller and Pusateri (2018) posit that this theory is flawed in that aggregate over-valuation would imply over priced targets, suggesting that acquirers do not actually benefit from the misinformation.

Gugler *et al.* (2012) find evidence in favour of the behavioural theory, finding weak correlations between listed and non-listed company activity during stock market booms, indicating that listed firms do indeed intend to capitalize on overvaluation. This, after finding the same disparity of correlation did not

support the Neoclassical theory, in that activity trends should transcend public and private equity markets. Stock payments are expected to be the common method of payment by the behavioural theory, as acquirers aim to exchange their overvalued shares for under-valued assets. The Behavioural theory is often broken down into two sub-theories, these being the overvaluation and managerial theories (Polemis and Paleologos, 2014).

Other, less common accounts for determinants include the Macroeconomic and Firm-level theories. Six articles address Firm-level characteristics as considerable determinants while five articles cite Macroeconomic factors as drivers of merger activity.

Firm-level theories aim to address the company characteristics and financial performance attributes deemed to have a driving role in deal making activity. Perspectives are shifted to both acquirer and target attributes, away from aggregated deal information. Chidambaran *et al.* (2018) find firm characteristics of companies to be important drivers for the likelihood of acquisition pursuit of companies. Attributes such as firm size, age and cash on hand were considered in their study of M&A in India during the 2002 to 2010 period. Firm-level characteristics attributed to financial distress by Masulis and Simsir (2013), in their study of target initiated deals, can be grouped within the Firm-level theory.

Macroeconomic theories deem the state of the economy to be a major driver of M&A activity. By this theory, variables of the economy such as interest rates, GDP and unemployment rates shape the market for merger activity (Cortés and Agudelo, 2017). Fuller and Pusateri (2018) recognize the state of the economy and business cycle as the most important indicators for merger activity.

Other, notable theories found within the literature related to industry specific characteristics, market expectations, managerial decision making, business networks and learning factors. These were not explored further as they were not frequently covered throughout the sample or because they were too qualitative in nature or had elements encompassed by other theories.

A list of determinants was collected from the literature sample. These were either identified through the specific article's reviewed literature or from a study's testing model. Drivers were also chosen from the top four theories in the literature. These being, the Neoclassical, Behavioural, Macroeconomic and Firm-level theories. Additional drivers categorized by other less frequently explored theories were disregarded. Each determinant was then assessed based on whether or not it was quantitative or categorical. Determinant variables of a qualitative or non-categorical nature were disregarded. This usually included determinants such as technological or regulatory change for which an explicit index of measurement was not used as a proxy. Where determinants were redundant or not unique but the descriptions or proxies differed, they were added too the list. Therefore the collection was focused on summarizing the most unique set of determinants, avoiding redundancy evident within

the literature sample. A summary of the criteria for determinant inclusion is summarized in the list below:

1. Determinants explored within the literature sample.
2. Determinants described in terms of quantifiable or categorical variables.
3. Determinants explored or tested in the context of any of the Behavioural, Neoclassical, Firm-level or Macroeconomic theories.

Tables 3.4, 3.5 and 3.6 present the synthesized list of M&A determinants gathered through the systematic quantitative literature review. A description or proxy, by which each determinant was explained, or modelled with, is then given along side the listed determinant. The respective theory within which it was found or related to and the contributing article are also included in each table. An index for source article citations is presented in Table 3.2.3. Numbers omitted from the list are a result of corresponding articles within the literature sample not contributing to the synthesized list of determinants.

Table 3.3: Indexed list of sources for synthesized list of determinants.

Source Index	**Article**
1	(Yaghoubi *et al.*, 2016)
2	(Gugler *et al.*, 2012)
3	(Cortés and Agudelo, 2017)
4	(Takechi, 2011)
5	(Polemis and Paleologos, 2014)
6	(Fuller and Pusateri, 2018)
7	(Komlenovic *et al.*, 2011)
11	(Amel-Zadeh and Zhang, 2015)
12	(Ellwanger and Boschma, 2015)
13	(Chidambaran *et al.*, 2018)
14	(Benzing, 1991)
15	(Masulis and Simsir, 2013)
17	(Harford, 2005)
18	(Alexandridis *et al.*, 2012)

Table 3.4: Neoclassical and Behavioural determinants gathered through review of literature.

Classification	Determinant	Description/Proxy	Source
Neoclassical	Capital Liquidity	Federal fund (C&I)* loan rate spread	5
		Federal Reserve Senior Loan Officer (SLO) survey	17
	Industry capacity utilization	Utilization rate of total production capacity	6
	Economic shocks to an industry	Cash flow margin on sales	7, 17
		Asset turnover	7, 17
		Research and development	7, 17
		Capital expenditures	7, 17
		Employee growth	7, 17
		Return on assets (ROA)	7, 17
		Sales growth	7, 17
		Market to Book (M/B) ratio	7, 17
	Exchange Rates	Dollar appreciation or depreciation	6
Behavioural	Economic policy uncertainty	Baker, Bloom, and Davis index	6
	Firm Over-valuations	Relative overvaluation using Q-ratio (Acq & Tar)*	2, 18
	Mis-valuations within the market	Q-ratio, price-to-book (P/B) ratios	1, 6
		Mean monthly inflation adjusted P/E ratio for S&P500	6, 18
		M/B value, price to earnings (P/E), Q-ratio	7
		Stock prices, M/B value	2
		Stock prices	1
	Investor sentiment	Stock prices, returns	1
		Stock market booms, (P/E)	2, 5
	Largest shareholders proportion	% of Shares owned by largest shareholder	2, 5
	Cash flows	Cash flow/Total Assets	2
	Managerial decisions in overvaluation	Std.dev of M/B ratio	17
		1- and 3-year stock returns and Std.dev of returns	17
	Industry average M/B ratio	Mean M/B ratio	7
	Industry deviation of Q-ratio	Std.dev of Q-ratio	7

* Commercial and Industrial (C&I) loan rates. Chicago Fed's National Activity Index (CFNAI). National Bureau of Economic Research (NBER). Gross National Product (GNP). Standard deviation (Std.dev).

Table 3.5: Macroeconomic determinants gathered through review of literature.

Classification	Determinant	Description/Proxy	Source
Macroeconomics	Bank prime loan rate		6
	Money supply		6
	AAA bond yields		14
	Recessions	NBER recession indicator & CFNAI	6
	Business cycle	Industrial Production & GDP*	7
	Capital market conditions	Treasury bills & 10-year treasury bond yields	7
		Total monthly returns for S&P 500 Index	7
	Exchange Rate	Annual real exchange rate depreciation	3
	Economic activity	Change in GNP*	7
	Interest rates	Interest rate (Various)	3, 6, 14
	Interest rate spread	Federal and industrial loan spread	2
	Market optimism	Stock market price level	7, 14
	Market returns	S&P 500 Index Returns	7
	Macroeconomic conditions	S&P 500 Index earnings per share (EPS) & stock prices	6
	Unemployment rate		3

* Gross Domestic Product (GDP).

Table 3.6: Firm-level determinants gathered through review of literature.

Classification	Determinant	Description/Proxy	Source
Firm-level	Labour productivity	(Sales - raw material purchases)/number of employees	4
	Firm age	Firm age in years on the date of acquisition	13, 15
	Firm size	Total Assets	5, 13, 15
		Total assets/number of employees	4
	Target Size	Market value, B/M ratio, P/E ratio	11
	Size of board	Number of directors on board	13
	Cash	Cash + marketable securities scaled by total assets	13
	Cash Flows	Cash flows/total assets	5
	Competition	Herfindahl-Hirschman index	15
		Annual firm sales/total industry sales	15
	Distressed firms	Low leverage, high ownership concentration	15
		Z-score*, interest coverage, liquidity ratios	15
		S&P bond rating, stock price	15
		Relative current ratios, relative leverage ratios	15
	Economic distress	Lose of market share, sales declines	15
		Rising costs, negative operating income	15
	Target Characteristics	Size, leverage, ROA, P/E ratio, liquidity	11
	Geographical proximity	Headquarter location	12
	Industry adjusted Q-ratio	Firm Q-ratio	13
	Leverage	Debt/Assets	5, 13
	Prior M&A experience	Past M&A embarking	13
	Founder share percentage	Percentage of Shares owned by founder	13
	Relatedness of industries	Industry codes	12
	Under-performing targets	Industry adjusted ROA, Q-ratio, sales growth	15
		Abnormal returns on target stock	15

* Z-score is an index that measures a firms likelihood of bankruptcy, using 5 key financial ratios (Altman, 1968).

While determinants were arranged according to the theories within which they were posited, it is evident that there is overlap among theories, with some containing similar determinants to others. Behavioural determinants relating to M/B ratios are also seen as economic shocks to industry according to Harford (2005) and Komlenovic *et al.* (2011). Stock market related determinants in the Behavioural theory category are also seen to be Macroeconomic drivers. However, these may be calculated using aggregate calculations for the entire market and not a specific industry or company, which seems to be the focus for Behavioural and Neoclassical theories.

A distinction should be made for firm-level factors that may seem to overlap with other categories. These financial figures may be used in aggregated form for the entire market or whole industries in the Behavioural and Neoclassical theories but the firm-level factor should be regarded as specific to a single company.

Some explanations of descriptions or proxies are covered in the following passage. Price-to-book (P/B) and market-to-book (M/B) ratios are deemed to be the same thing, in that they are a valuation measure for a company. They both value a company using the current market value based on the share price and number of outstanding shares as well as its net book value, being its total assets minus its total liabilities. Similar, but different, the q-ratio describes the relationship between a company's market value of their assets and what the replacement value of those assets would be. The price earning P/E ratio is calculated as the share price divided by the earnings per share (EPS). EPS is calculated as the net profit divided by the number of outstanding shares. The P/E ratio is deemed a good indicator of companies valuation especially in relative valuations.

Harford (2005) used an index for economic shocks to an industry, based on the first principle components of the financial variables listed under the Neoclassical theory. This was done in an effort to avoid multi-co-linearity between variables. The index was calculated for each industry, using aggregated financial information.

3.2.4 Discussion

After review of the literature sample on the topic of determinants of M&A, it can be said that a general consensus exists for the respective theories of determinants. Although many exist, two hypotheses remain dominant, the Neoclassical and Behavioural Hypotheses. Papers also consider firm-level and Macroeconomic factors as alternative classifications of determinants. However, there is great variation in the choice of proxies used when testing for these and other hypotheses. While studies consistently use regression models to explain relationships, additional variation to consider in interpretation of results is introduced through the use of different sample periods, types of deals included and countries under observation. This makes summarizing of results

difficult and sometimes subjective, especially when intending to find categorical evaluations of proposed determinants. It is clear that the relationship between drivers and activity are complex, however, there is still room to find a comprehensive set of determinants capable of predicting merger activity best.

Generally, through duplication of proxies and determinants across theories a cloud is cast over the true causes for merger trends. While Fuller and Pusateri (2018) identify Macroeconomic factors as paramount determinants of M&A activity, criticism is found of categorical models that use the Neoclassical, Behavioural and Macroeconomic theories exclusively, after finding biases within models that do not consider other less conventional competing variables, particularly when not considering firm-level attributes such as firm size, industry or geographic region.

Theories try to exclusively explain M&A activity in a sort of one size fits all strategy but while there is overlap in determinant use for theory testing, efforts will remain hindered. This prompts the requirement for new approaches. It may still be important to distinguish for proxies with respect to prior theoretical grounding, but given the complex nature of determinants that is evident, great potential lies in an analysis that is cast wide in the consideration of variables capable of predicting merger activity best.

Although the sixth merger wave has been well studied and characterized, no article in the sample acknowledged the start of a seventh wave of mergers or any characterizations thereof. An analysis of more recent merger activity may lead to evidence of a seventh wave and the characterizations thereof.

3.3 Data Science Theory

The intention of this section is to explore methods capable of facilitating M&A industry activity and determinant analysis. An application such as this, falls into the realms of Data Management and Data Analysis or, more broadly, Data Science. The section is broken down into two subsections, these being, Data Management Theory and Exploratory Data Analysis, respectively. Section 3.3.1 covers database management and design concepts while the analysis methods sections cover suitable means for the exploration of M&A activity and determinant evaluation, in Sections 3.3.2.

3.3.1 Data Management Theory

If valuable insights are to be drawn from available M&A information, then a structured database should be used as a means for formally storing and managing data. Kendall and Kendall (2011) explain that a database is a formally defined and centrally controlled store of data intended for use in many different applications." An implementation of such a facility, if designed correctly, allows for the efficient storage and retrieval of data as well as data

integrity, data that is consistent and accurate (Kendall and Kendall, 2011). It is for these reasons that a database should be seen as a cornerstone of any in depth analysis facilitation, in that it allows for efficient access to information stored and that the integrity of data is not compromised.

Databases and the Data Warehouse: Database types are best characterized by the purpose of use and how regularly information is updated. These factors can vary in terms of the number of intended users, scope of information to be made available and whether data is centralized or distributed in its storage configuration. A popular and widely used type, is an Operational Database. This type is designed to reflect the up to date activities of a businesses, continuously and accurately. Data is regularly added or updated in such a database (Rob *et al.*, 2008).

Rob *et al.* (2008) defines a **data warehouse** as an alternative to an operational database. It is purpose built for decision support and analysis capabilities. In satisfying these objectives, a **data warehouse** must allow for the consolidation of relevant information from various sources that may other wise be separated. Derived information is then used to drive tactical or strategic decision making (Rob *et al.*, 2008). A data warehouse is to be used for quick and effective queries on data of a common subject theme, organized to store data sourced from multiple databases in a uniform way (Kendall and Kendall, 2011).

To summarize, data warehouses and operational databases are broadly, examples of relational database systems; distinctions are made based on their intended use. The warehouse facilitates consolidation of data sources that are otherwise separated and drive efficient queries for more complex analysis. It must allow for the periodic updating or insertion of data when bulk amounts of new information, if, and when it becomes available. The operational database facilitates the effective management of up to date transactions within a business, where queries can deliver the current state of any desired function or entity. Data is updated timeously to support this using various kinds of sensory inputs (Rob *et al.*, 2008).

The twelve rules that define a data warehouse, as developed and presented by Inmon and Kelly (1994), can be found in Appendix A. This set of guidelines provides a good characterization of a data warehouse, while considering aspects from various stages of its life cycle (Inmon and Kelly, 1994; Rob *et al.*, 2008).

From the above explanation a strong case is made for the appropriateness of a data warehouse as a tool that should be used in the facilitation of effective analyses of the M&A industry. Because, information will need to be gathered and managed from separate source data, while allowing efficient complex querying that will drive the deeper analysis of M&A field information. If and when new and relevant information becomes available, the database will allow for updating or insertion of new data. The problem requires a solution be-

yond analyses of spreadsheet data where available computing capacity would be compromised or far less efficient than queries within a formally structured database.

Design Concepts: In creating a relational database capable of facilitating the data warehouse, certain design and development considerations were established. By understanding the implementation of methods within the database development methodology, the integrity of the database and its stored information could be maintained. Consequently, concepts of relational database design are explored in the context of the design process through the following passages.

Relational databases require a combination of information **entities** and **relationships**, governed by **Keys**. These elements are all encompassed by a method of relational database design known as entity relationship modelling. Here, a conceptual design is used to model the real world scenario of required information to be stored (Rob *et al.*, 2008).

Before creating a sound conceptual model, requirements must be established and subsequently, necessary data analysed, in terms of data characteristics and where it can be sourced. This requires recognizing the real world entities, their attributes and governing relationships to be modelled from the real world scenario. After this, an entity relationship diagram (ERD) can be created detailing the various information entities and how they are related. Entities organize subject grouped information and can be characterized by things like a place, person, event or thing, reflecting the real world instance. Attributes are then used to further enhance characteristics of entities. Entities can contain characteristics known as attributes and instances or occurrences known as records of the entity. If an entity can exist on its own without requiring attributive characteristics of another entity, it is defined as a primary entity (Rob *et al.*, 2008).

Relationships associate and link entities to one another, driven by relating attributes. These cardinalities are modelled diagrammatically using Crow's Foot notation. There are three main types of entity relationships, one-to-many, many-to-many and one-to-one relationships. Keys help govern these relationships by creating an identifier for a record in a parent table that can be used to reference the parent table in a child table. Three types of Keys are primary, foreign and composite Keys. A primary Key uniquely identifies a record within a table while a foreign Key is used to refer to a specific record of corresponding instance of a parent table, within a child table. A composite Key is the combination of two primary (identifier) Keys from independent tables, combined to create a new identifier in a many-to-many relationship. These composite Keys act as unique identifiers for intersection entities (Kendall and Kendall, 2011; Rob *et al.*, 2008).

Maintaining the integrity of information stored in a database is critical to

its implementation. By maintaining integrity, data anomalies can be deterred aligning with the justification for creating a database of efficiently and effectively stored information. **Entity integrity** enforces that all primary keys have non-null values and that any composite keys contain a non-null component key. **Referential integrity** stipulates that all foreign keys in child tables have an existing corresponding primary key in the parent table of a one-to-many relationship. Conditions of such integrity imply that one cannot adjust or delete a primary key that has any corresponding child references. **Domain integrity** assists in the validation of data to be stored, in terms of its formatting. This is for all key and attributive columns, defining the ranges of data that can be stored in a field including conditions such as data formatting types accepted and whether null-values are accepted for non key attributes (Kendall and Kendall, 2011; Rob *et al.*, 2008).

Anomalies that can be deterred for databases, by enforcing the above mentioned integrity constraints include, insertion, deletion and updating anomalies. Insert anomalies occur when a non-existent primary key is used in the insertion of a child table record, essentially creating a reference to non existent information. A deletion anomaly occurs when deleting a record and subsequently, unintended related information. When attempting to update an attribute of an entity and the action results in inconsistencies within the database, or requires a change in more than one place, the result is an update anomaly (Kendall and Kendall, 2011).

Conceptual designs can be validated using normalization. Normalization is a method of data manipulation into simplified data structures that mitigate data redundancy, deterring potential instances of data anomalies. The process involves transforming complex user views or data tables through various stages of the normal form. Table creation is driven by achieving tables which are subject specific, have no unnecessary information duplication and have attributes of individual tables fully dependent on primary Keys. When done concurrently with modelling methods of entity relation design, good databases can be produced, as a correctly normalized third normal form table will avert data anomalies and when aligned with conceptual and logical designs, stand as a validation mechanism of the design (Rob *et al.*, 2008).

The normalization method is broken into three stages called the First, Second and Third Normal Forms (1NF, 2NF, 3NF). The first step of normalization requires the removal of repeating groups within the data to create new tables with independent primary Keys. The second normal form requires the removal of any partially dependent data of the primary Key, creating a new entity that is related to part of a Key another that is dependent on the full Key. 2NF must be in 1NF. The 3NF requires 2NF achievement and the removal of any transitive dependencies. These are any attributes that are dependent on both the primary Key of an entity and another non-key attribute (Rob *et al.*, 2008; Kendall and Kendall, 2011).

After a conceptual design is developed, a logical design can be created.

This stage of design enhances conceptual elements by stipulating the format of entities and attributes. A database schema can be used to represent a logical design (Rob *et al.*, 2008).

3.3.2 Exploratory Data Analysis

Coverage of theory and methods throughout this section is driven by main thesis Objectives 4 and 5. Exploratory data analysis (EDA) was used in the approach to analysis of merger activity and determinants. The following passages detail methods appropriate for EDA that are capable of facilitating the necessary inquisition. The coverage will focus on theory around the methods and models, towards leveraging existing functions within statistical software, in the analysis application.

EDA: Behrens (1997) describes EDA as a means for discovering the properties of observed data. The approach is focused on considering data that is available and using it within an experimental setting. Here, models may be adjusted until a plausible conjecture can be made about the data. An emphasis is given to the use of robust measures and residuals analysis, that may prompt the necessary remedial action in model correction. Common methods employed in EDA consider graphical representation of data, linear regression and correlation, probability distributions, analysis of variance and time series analysis (Behrens, 1997; Hoaglin and Velleman, 1949).

Time Series Modelling: Time series modelling allows the attainment of an understanding of stochastic elements within a time series. There are various types of time series and as a result, a suitable model should be used, accordingly. Then, the parameters that help define the model should be correctly estimated according to characteristics of that time series. By estimating model parameters correctly, better modelling accuracy may be achieved (Kitagawa, 2010).

Moving average and Autoregressive processes are commonly used in time series modelling. The former considers a fixed window (parameter q) of prior observations for a time series in predicting the present value. The resulting model has a smoothing affect for observations that broadly follow the trend of the series more closely. An autoregressive process relies on past observations scaled by some coefficient in predicting the present value within a series. A time lag parameter, p is specified for the number of prior observations to be considered in predicting the present time series value. A combination of the autoregressive and moving average (ARMA) processes allows for an efficient predictive and smoothing method in time series modelling. When a time series is classified as non-stationary then an appropriate model, using an additional process within ARMA, is the autoregressive integrated moving average

(ARIMA). Here the integrated term represents the preprocessing technique of differencing in order to force stationarity for the series, before applying the remaining model processes. A time series is stationary if its statistical properties such as mean and variance remain constant over time. Suitable model parameters *p, d, q* should be effectively estimated in developing an ARIMA model with a satisfactory accuracy of fit. These are, number of time lags (p), degree of differencing (d) and order of moving average (q) (Chatfield, 2004; Polasek, 2013).

A useful function, within R's Forecast package (Hyndman, 2016), towards fitting appropriate ARIMA models, is the auto.arima() function. This function automatically chooses and fits the best ARIMA parameters by using the Akaike information criterion (AIC) and Bayesian information criterion (BIC) for respective trial models. The AIC and BIC asses goodness-of-fit while considering simplicity of the model, by penalizing models with larger numbers of parameters, so as to reduce potential over-fitting (Chatfield, 2004).

Multiple Linear Regressions: Multiple linear regressions make use of multiple independent variables (predictors) in explaining a dependent variable (response). The relationship can be described through the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \epsilon \tag{3.3.1}$$

Here, k independent variables (X) can be used to explain the variance in dependent variable Y. Beta coefficients represent an expected change in dependent variable Y per unit change of an independent variable, while keeping all other independent variables constant. The constant beta and epsilon represent the intercept and error terms, respectively (Montgomery and Runger, 2003).

The least squares method can be applied in estimating regression coefficients. Here, model coefficients that produce the minimum squared value of all predicted values minus their respective observed values (residuals), are chosen as the regression of best fit (Montgomery and Runger, 2003).

Feature Selection: In statistics and machine learning, feature selection is a method of dimensionality reduction, used to remove irrelevant and redundant model variables that jeopardize model performance and increase computational complexity. The method allows for improved interpretation of models as existing variables are maintained in their original form without the need for additional transformation (Tang *et al.*, 2014).

LASSO: The Least Absolute Shrinkage and Selection Operator (LASSO) is a popular method utilized in feature selection. The analysis method introduces regularization of variables by applying a penalty to model coefficients by shrinking them or setting them to zero, while minimizing least squares error (residual square error). The application results in selection (by elimination)

of the most relevant independent variables, capable of predicting a response with the least residual error (Tibshirani, 1996).

k-Fold Cross-Validation: A suitable method to supplement overall model selection, based on predictive accuracy, is that of k-Fold cross-validation. The method is effective in avoiding the training-data/test-data trade off inherent to many machine learning problems. Sample observations are divided into k equal groups, then one group is excluded for validation, while the model is fitted to the remaining (k - 1) groups (training set). Performance in measured using an average of all k validated groups (Wong, 2015).

Through the literature review chapter, an introduction to the field of Mergers and Acquisitions was presented in Section 3.1. Here, important aspects of deal making were covered in more depth. A systematic quantitative review of literature in the field provided foundation for the creation of a synthesized list of activity and determinant variables. Coverage of exploratory data analysis methods, deemed most relevant to thesis Objectives 4 and 5, was presented in the last section of the literature review chapter.

Chapter 4

Data Warehouse Development

This chapter conveys the stages and supporting contributions towards the data warehouse development, here-in referred to as a database. As discussed in the main thesis methodology, the broad aim of the database was to best facilitate the warehousing of information required for M&A activity and driver analysis. This was informed by the literature review, through the identification of industry activity determinants, in terms of the variables which characterize it.

The database development method stipulated in Chapter 2, Section 2.1.3, is presented through the following sections. A detailed exploration of gathered data elements and requirements for the database follows in Section 4.1 where-after, necessary data pre-processing measures executed in line with the extraction and transformation activities of the Data Warehouse Architecture are presented in Section 4.2. Respective stages of the database design aspect are documented in Section 4.3 while Section 4.4 details the practical elements of implementation, including software selection, construction and data loading. Database testing, evaluation and operation as well as maintenance and evolution stages are presented through Sections 4.5 to 4.7.

4.1 Database Initial Study

Broadly, thesis Objective 3 required the development of an information management tool capable of supporting the activity and determinant analysis of the M&A industry. While this was recognized as the primary objective supporting the development of a database, an investigation into the domain and its specific information requirements was still necessary. The systematic literature review method, covered through Section 3.2 of Chapter 3, assisted in the recognition and establishment of theoretical requirements for information deemed most important by literature, for such an inquisition into the M&A field. This would ultimately be constrained by information availability. Therefore, following step one of the database development methodology, a data element analysis

was performed in an initial study, by considering outcomes of the literature review, specifically the synthesized list of activity and determinant variables and subsequent data gathered. Where, characteristics of the gathered data were explored towards supporting an effective design methodology.

While the literature review enabled the identification of what information variables and determinants that would be required, the scope, documented in Section 1.4, defined which deals would be included in the deal set. The complimenting search protocol, presented in Section 2.2 was executed towards attaining a deal set that best suited the scope of study. A deal set of 4 828 mergers resulted for domestic deals announced in the U.S.A on or after the first of January 1998 and which were completed by the end of June 2019, among the other constraints detailed in the search protocol.

4.1.1 Activity Variables

The lists of deal and company specific variables obtained through data gathering are presented in Tables 4.1 and 4.2. Both tables characterize each variable by the data format, presence of null values and whether or not the field contained repeating groups of categorical values.

Table 4.1: Obtained deal specific variables of M&A activity.

Activity Variables	**Format**	**Nulls**	**Grouping**
Announcement Date	Date	No	No
Announced Total Value (mil.)	Numeric		
Cash Terms	Numeric/Text		
Stock Terms		Yes	
Payment Type	Text	No	Yes
Attributes		Yes	
Announced Premium	Numeric		No
Current Premium			
Completion Date	Date	No	
Deal Description	Text		
Bid Nature		Yes	Yes
Percent Owned	Numeric	No	No
Percent Sought			
Net Debt		No	
Terminal Value/EBITDA		Yes	
Equity Value/Book Val			

Table 4.2: Obtained company specific variables of M&A activity.

Activity Variables	Format	Nulls	Grouping
Company Name	Text	No	Yes
Company Industry Sector			
Company Industry Group			
Company Industry Subgroup			
Company City of Domicile		Yes	
Company Primary Exchange			
Company State Code			
Company Ticker No	Alphanumeric		
Number of Employees	Numeric	Yes	No
Current Assets Reported			
FCF/Diluted Shares			
Last Pub Offer (Share Price)			
Roc/Wacc Ratio			
Asset Turnover			
Sales/Revenue/Turnover			
Assets Current Liability			
Assets/Equity			
Book Value per Share			
Cash & Equiv to Current Assets			
Cash & Equivalents			
Cash Ratio			
Cashflow to Total Liabilities			
Current Ratio			
Debt To Market Cap Ratio			
EV To Book Value			
EV To Market Cap			
Enterprise Value			
Free Cash Flow			
Gross Profit			
Number Of Shareholders			
Price to Book Ratio			
Price/Cash Flow			
Relative PE Ratio			
Change Year To Date Percent			
Earnings Per Share			

In a comparison of the literature review's identified variables and obtained activity variables, Table 4.3 highlights the resulting discrepancies after data gathering.

The list of additional and potentially useful variables obtained through data gathering can be found in Table 4.4.

Table 4.3: Firm and deal specific variables attainment.

#	Firm Specific		Deal Specific	
1	Firm Size (Employees)	Yes	Deal Volume	Yes
2	Firm Size (Total Assets)	Yes	Deal Value	Yes
3	Economic Sector	Yes	Deal Type	Yes
4	Industry (SIC Codes)	No*	Payment Type	Yes
5	Ownership (% Change)	Yes	Completion Status	Yes*
6	Ownership Concentration	No	Premium	Yes
7	Share Price	Yes	Payment Currency*	No
8	Age of a Firm	No	Time to Completion	Yes*
9	Age of CEO	No	Domestic or Cross Border	Yes*
10	Domicile (State)	Yes	Integration Type	Yes*
11	Industry Competitiveness	No		
12	Listed Status	Yes*		
13	Total Sales	Yes		
14	Total Debt	Yes*		

* Bloomberg industry classifications gathered in place of SIC Codes. All companies are listed given the search criteria stipulated for the search protocol. A net debt value for deals was gathered. All deals were completed given the search criteria stipulated for the search protocol. Time to completion derived from deal announcement and completion dates. All deals were domestic given the search criteria stipulated for the search protocol. Integration type could be derived by comparing respective industry classifications for targets and acquirers.

Table 4.4: Additional firm and deal specific variables.

#	Firm Specific	Deal Specific
1	Primary Exchange	Bid Nature
2	Other figures and ratios	Net Debt
3		Terminal Value/EBITDA
4		Equity Value/Book Val
5		Cash and Stock Terms

4.1.2 Determinant Variables

The list of M&A determinant variables obtained through data gathering is presented in Table 4.5. Each determinant's description, its theoretical grouping (according to the review of literature) along with its reporting frequency are provided. Recognition of the frequency identifies the absence of monthly values for determinants reported on a quarterly and yearly basis, if a monthly baseline time period were to be used in analysis. The Delta column indicates

if the variable measures change in a value over a specific time period (month-on-month, quarterly-on-quarterly and so on). Each variable was observed to be of numeric value and no repeating groups were evident for the gathered determinant data.

The letters B, M and N represent the Behavioural, Macro-economic and Neoclassical the theories respectively. Where parentheses are present, the allocation is deemed weak as the proxy was not explicitly identified through the literature review, but declared potentially useful as an attainable variable in data gathering and subsequently, grouped at the authors discretion.

Table 4.5: Determinant variables of M&A activity.

Determinant Description	Theory	Delta	Frequency
Implied Interest Rate	M	No	Monthly
Ten Yr Government Bond Yields	M		
Two Yr Government Bond Yields	M		
Gold Spot Price	(M)		
S&P 500 Index	B, M		
Reccession Indicator	M		
NSA CPI YOY Index	(M)	Yes	
US Dollar Spot Index	N	No	
Prime Interest Rate	M		
Corporate High Yield Tot Return Index	(M)		
Federal Funds Effective Rate	N		
Chicago Fed National Activity Index	M		
M1 Money Supply	M		
M2 Money Supply	M		
Economic Policy Uncertainty Index	B		
Trade Policy Uncertainty Index	B		
Industrial Production MOM %	M	Yes	
S&P 500 PE Ratio	B, M	No	
Unemployment Rate %	M		
GDP QOQ %	M, N	Yes	Quarterly
S&P EPS + Surprise	M, (B)	No	
Real GDP IMF	M, N		
GDP QOQ % IMF	M,N	Yes	
Chapter 11 Bankruptcy Filings	N	No	
Chapter 12 Bankruptcy Filings	N		
Chapter 13 Bankruptcy Filings	N		
Total Bankruptcy Filings	N		
Fiscal Freedom Index	(M),(N)		Yearly
Annual Housing Starts	(M)		

This section concludes the contribution to information analysis as part of

the initial database study aimed at understanding information requirements and scope of data to be stored. A foundation for effective conceptual and logical design of the database was formulated.

4.2 Data Pre-processing

Pre-processing encompassed the extraction and transformation activities to ensure that data was in line with the data warehouse system architecture. With regard to extraction, explicit details of the data gathering method were covered in Section 2.2. The resulting data were collected using Microsoft Excel (MS Excel) workbooks. Bloomberg L.P. (2017) search results were limited to 20 column variables per deal set and therefore required multiple downloads of results sets that included different variables. The following subsections document the pre-processing activities of data merging and data cleaning.

4.2.1 Data Merging

Deals data was then merged using the Bloomberg L.P. (2017) Action IDs, a uniquely assigned identification number for each deal. After the merging of columns for each deal record, a check of the merge was performed by comparing the acquirer and target names included in the new set of variables, to the original deal set company names. These proved to be successful for every deal of each merge.

Determinant variables were merged using an identifier made from a concatenation of the year and month for which a value had been recorded for. A corresponding key value for that record was then linked to each deal record, with a matching year-month combination.

4.2.2 Inconsistencies and Cleaning

An annual determinant data variable, Fiscal Freedom Index, was found to have a missing value for the year 2001. The value from 2000 was used as a substitute.

Company and deal specific variables with potential categorical repeating groups were identified and analysed. Discrepancies in spelling were found for the supposed categorical values of the City of Domicile variable where after it was found that there was repetition of city names with minor spelling differences. Examples included errors such as “Jacksonville" and “Jackonsville" as well “Minneapolis" and “Minneaplis". The cities of Dublin and Dublin 2 were found to be used correctly after Dublin was found to be a town in the U.S.A as well (apart from capital of the Republic of Ireland).

The Deal Attributes variable was found to contain multiple categorical characteristics in a single cell for each deal record. By converting the multi-

attributive data to additional columns, using MS Excel's Text-To-Columns function, new atomic value columns could be created for each of the attributes given to a deal. The maximum number of different categorical attributes was found to be six, therefore six new columns were created accordingly.

4.3 Design

Following the second rule data warehouse definition, from the Twelve Rules that Define a Data Warehouse, as described by Inmon and Kelly (1994), data required for prospective analyses were to have been integrated, regardless of its source. For this reason, a relational database was chosen as a suitable structure to facilitate the data warehouse. Contributions for the respective design stages within the greater DBLC are covered in the following sub-sections, according to the method definition of Section 2.1.3 and relevant literature coverage of Section 3.3.1. Variables identified through the initial database study may be referred to as columns or fields in the relational database development method. Aspects of the database design are covered through the following subsections. These include normalization, the entity and extended entity relationship diagrams as well as the logical design.

4.3.1 Normalization

After considering the initial study of database requirements, the obtained deals data set was normalized concurrently with ERD design, as a means for its validation. In line with the variables presented in Tables 4.1 and 4.2, the complete initial de-normalized deal set consisted of the fields presented in Table 4.6. A total of 92 variables existed for the initial table. Both duplicated columns and column groups existed for the deal variable set. The determinant variables data set was not included in the normalization process as the gathered data was found to be in the Third Normal Form. Nevertheless, it would be introduced to the relational model in the entity modelling stages of Sections 4.3.2 and 4.3.3.

Table 4.6: Initial de-normalized data table field headings.

Field Set 1	Field Set 2
Action ID	Acquirer Roc/Wacc Ratio
Announce Date	Target Roc/Wacc Ratio
Acquirer Name	Acquirer Asset Turnover
Target Name	Target Asset Turnover
Announced Total Value (mil.)	Acquirer Sales/Revenue/Turnover
Cash Terms	Target Sales/Revenue/Turnover
Stock Terms	Acquirer Assets Current Liability
Payment Type	Target Assets Current Liability
Deal Attribute_1	Acquirer Assets/Equity
Deal Attribute_2	Target Assets/Equity
Deal Attribute_3	Acquirer Book Value per Share
Deal Attribute_4	Target Book Value per Share
Deal Attribute_5	Acquirer Cash & Equiv to Current Assets
Deal Attribute_6	Target Cash & Equiv to Current Assets
Announced Premium	Acquirer Cash & Equivalents
Current Premium	Target Cash & Equivalents
Completion/Termination Date	Acquirer Cash Ratio
Deal Description	Target Cash Ratio
Nature of Bid	Acquirer Cashflow to Total Liabilities
Percent Owned	Target Cashflow to Total Liabilities
Percent Sought	Acquirer Current Ratio
Net Debt	Target Current Ratio
TV/EBITDA	Acquirer Debt To Market Cap Ratio
EqV/Book Val	Target Debt To Market Cap Ratio
Acquirer Industry Sector	Acquirer EV To Book Value
Target Industry Sector	Target EV To Book Value
Acquirer Industry Group	Acquirer EV To Market Cap
Target Industry Group	Target EV To Market Cap
Acquirer Industry Subgroup	Acquirer Enterprise Value
Target Industry Subgroup	Target Enterprise Value
Acquirer City of Domicile	Acquirer Free Cash Flow
Target City of Domicile	Target Free Cash Flow
Acquirer Ticker	Acquirer Gross Profit
Target Ticker	Target Gross Profit
Acquirer Primary Exchange	Acquirer Number Of Shareholders
Target Primary Exchange	Target Number Of Shareholders
Acquirer State Code	Acquirer Price to Book Ratio
Target State Code	Target Price to Book Ratio
Acquirer Number of Employees	Acquirer Price/Cash Flow
Target Number of Employees	Target Price/Cash Flow
Acquirer Current Assets Reported	Acquirer Relative P/E Ratio
Target Current Assets Reported	Target Relative P/E Ratio
Acquirer FCF/Dil Shr	Acquirer Change Year To Date Percent
Target FCF/Dil Shr	Target Change Year To Date Percent
Acquirer Last Pub Offer (Share Price)	Acquirer Earnings Per Share
Target Last Pub Offer (Share Price)	Target Earnings Per Share

Table 4.6 contained repeating fields for Deal Attributes as well as a duplication of the same company characterizations and company financial figures, but separated as independent columns for acquirers and targets, respectively. Repeating groups within fields were identified for certain deal and company characteristics. These repeating groups were recognized and denoted in Table 4.1 of the initial database study.

In an effort towards achieving the First Normal Form for the de-normalized deal set, the Deals attributive entity was created. Repeating groups were removed from the Deals entity, creating additional primary entities with one-to-many relationships, along side it.

A many-to-many relationship was used to eradicate the data redundancy of the repeated Deal Attribute columns of Table 4.6, resulting in the creation of the Deals Attributes entity. Repeating columns for company attributes and financial figures required a similar approach while using a new identifier variable column to distinguish for acquirer and target information, for the same fields. A Dispositions field was created in the Deals entity, to denote the disposition of a company and its subsequent characteristics and financial figures in a deal record as being either target or acquirer related. This repeating group (acquirer or target value) was then removed using a many-to-many relationship and the corresponding foreign key used as a component of the composite key created for Deals. This allowed the storage of unique records that were dependent on the designated composite key in the many-to-many relationship between Deals and Dispositions, away from the dependency on the repeating Action ID key. The resulting set of entities can be found in Table 4.7, where the primary keys of entities are denoted with the abbreviation, PK. Foreign keys for the one-to-many relationship of repeating groups are denoted with the abbreviation, FK in the main Deals attributive entity. Entities with composite primary keys contain two PK symbols.

Table 4.7: First Normal Form Fields.

	Deals		**Dispositions**
PK	ID_PK_Action ID	**PK**	ID_PK_Disposition
PK	ID_FK_Disposition		Disposition
FK	ID_FK_Payment Type		
FK	ID_FK_Nature of Bid		**Payment Types**
	Announce Date	**PK**	ID_PK_Payment Type
	Announced Total Value (mil.)		Payment Type
	Cash Terms		
	Stock Terms		**Deals Attributes**
	Announced Premium	**PK**	ID_FK_Action ID
	Current Premium	**PK**	ID_FK_Deal Attribute
	Completion/Termination Date		Deal Attribute
	Deal Description		

Table 4.7 continued from previous page

	Percent Owned		**Bid Natures**
	Percent Sought	**PK**	ID_PK_Nature of Bid
	Net Debt		Nature of Bid
	TV/EBITDA		
	EqV/Book Val		**Companies**
FK	ID_FK_Company	**PK**	ID_PK_Company Name
FK	ID_FK_Industry Sector		Company Name
FK	ID_FK_Industry Group		
FK	ID_FK_Industry Subgroup		**Industry Sector**
FK	ID_FK_City of Domicile	**PK**	ID_PK_Industry Sector
FK	ID_FK_Primary Exchange		Industry Sector
FK	ID_FK_State Code		
FK	ID_FK_Ticker		**Industry Group**
	Number of Employees	**PK**	ID_PK_Industry Group
	Current Assets Reported		Industry Group
	FCF/Dil Shr		
	Last Pub Offer (Share Price)		**Industry Sub-Group**
	Roc/Wacc Ratio	**PK**	ID_PK_Industry Subgroup
	Asset Turnover		Industry Subgroup
	Sales/Revenue/Turnover		
	Assets Current Liability		**City of Domiciles**
	Assets/Equity	**PK**	ID_PK_City of Domicile
	Book Value per Share		City of Domicile
	Cash & Equiv to Current Assets		
	Cash & Equivalents		**Primary Exchanges**
	Cash Ratio	**PK**	ID_PK_Primary Exchange
	Cashflow to Total Liabilities		Primary Exchange
	Current Ratio		
	Debt To Market Cap Ratio		**States**
	EV To Book Value	**PK**	ID_PK_State Code
	EV To Market Cap		State Code
	Enterprise Value		
	Free Cash Flow		**Tickers**
	Gross Profit	**PK**	ID_PK_Ticker
	Number Of Shareholders		Ticker
	Price to Book Ratio		
	Price/Cash Flow		
	Relative P/E Ratio		
	Change Year To Date Percent		
	Earnings Per Share		

Partial dependencies within the composite key of the 1NF Deals entity were

found for deal specific fields of Payment Type through to Equity Value/Book Value ratio. This partial dependency was removed by creating a new Deals entity and renaming the previous Deals entity, Deals Financials, the intersection entity of Deals and Deals Dispositions.

The partial dependency within the composite key for Attributes in the Deal Attributes entity was removed by creating an additional entity, Attributes, so that it relied primarily on its own new primary key. This allowed the achievement of the Second Normal Form.

Table 4.8: Second Normal Form Fields.

	Deals Financials		**Deals**
PK	ID_FK_Action ID	**PK**	ID_PK_Action ID
PK	ID_FK_Disposition	**FK**	ID_FK_Payment Type
FK	ID_FK_Company	**FK**	ID_FK_Nature of Bid
FK	ID_FK_Industry Sector		Announce Date
FK	ID_FK_Industry Group		Announced Total Value (mil.)
FK	ID_FK_Industry Subgroup		Cash Terms
FK	ID_FK_City of Domicile		Stock Terms
FK	ID_FK_Primary Exchange		Announced Premium
FK	ID_FK_State Code		Current Premium
FK	ID_FK_Ticker		Completion/Termination Date
	Number of Employees		Deal Description
	Current Assets Reported		Percent Owned
	FCF/Dil Shr		Percent Sought
	Last Pub Offer (Share Price)		Net Debt
	Roc/Wacc Ratio		TV/EBITDA
	Asset Turnover		EqV/Book Val
	Sales/Revenue/Turnover		
	Assets Current Liability		**Dispositions**
	Assets/Equity	**PK**	ID_PK_Disposition
	Book Value per Share		Disposition
	Cash & Equiv to Current Assets		
	Cash & Equivalents		**Payment Types**
	Cash Ratio	**PK**	ID_PK_Payment Type
	Cashflow to Total Liabilities		Payment Type
	Current Ratio		
	Debt To Market Cap Ratio		**Deals Attributes**
	EV To Book Value	**PK**	ID_FK_Action ID
	EV To Market Cap	**PK**	ID_FK_Deal Attribute
	Enterprise Value		
	Free Cash Flow		**Attributes**
	Gross Profit	**PK**	ID_PK_Deal Attribute
	Number Of Shareholders		Attribute

Table 4.8 continued from previous page

	Price to Book Ratio		
	Price/Cash Flow		**Bid Natures**
	Relative P/E Ratio	**PK**	ID_PK_Nature of Bid
	Change Year To Date Percent		Nature of Bid
	Earnings Per Share		
	Companies		**City of Domiciles**
PK	ID_PK_Company Name	**PK**	ID_PK_City of Domicile
	Company Name		ID_PK_City of Domicile
	Industry Sector		**Primary Exchanges**
PK	ID_PK_Industry Sector	**PK**	ID_PK_Primary Exchange
	Industry Sector		Primary Exchange
	Industry Group		**States**
PK	ID_PK_Industry Group	**PK**	ID_PK_State Code
	Industry Group		State Code
	Industry Sub-Group		**Tickers**
PK	ID_PK_Industry Subgroup	**PK**	ID_PK_Ticker
	Industry Subgroup		Ticker

After the achievement of Second Normal Form for entities, the tables were found to have achieved Third Normal Form, as no transitive dependencies remained. It initially seemed that there may have been existing transitive dependencies between the foreign Companies key and company attributes like industry sector and the other company related foreign key attributes. However, these were found to be fully determined by the Deal and Disposition foreign keys and would not repeat for companies in multiple deals as these attributes could change over time in different deals. As such, this relationship would be modelled to accommodate the potential change of these attributes of companies in different deals.

4.3.2 Entity Relationship Diagram

In an attempt to represent the conceptual database design graphically, an entity relationship diagram (ERD) was developed. This aided the visual representation of the respective data entities and their relationships within the proposed database.

Initial design attempts proposed that primary entities characterizing company attributes be modelled as attributive entities to the main companies entity, evident in Appendix B. The database was formerly modelled around two main attributive entities, Companies and Deals. However, this design

was adapted after finding that attributes of companies seemed to be repeating groups, but were found to change for certain companies over time with different deals as companies shifted in their disposition of these attributes. After normalization, it was confirmed that no transitive dependencies existed for the non-key attribute Company ID and that company attributes changed with time for companies involved in multiple deals. While this was only certain attributes such as industry subgroup, the model needed to accommodate this change categorically and therefore required a correction in modelling approach, for all company attributes. Appendix B shows the initially proposed model before normalization validated the absence of transitive dependencies within the Company entity and its related attributes. Each entity is grouped by colour depending on its subject orientation.

In line with normalization of the deals data set, the ERD, presented in Figure 4.1, was designed to reflect the discovered data entities. Colours were used to associate these entities based on their subject grouping. Primary entities were represented by thicker bordered rectangles, while the main attributive entity, Deals, was bordered otherwise. The resulting design resembled a Star Schema, a configuration commonly used for Data Warehouse design.

The determinants entity was introduced to the model by relating the Deals entity with attributes of the Determinants entity that corresponded with the year and month of deal announcement dates. This would ensure that the relevant Determinant record would be directly related to each deal record, depending on the month and year of a deal, in a one-to-many relationship. Any given Deal record was to contain one relating Determinant record while a Determinant record could relate to one or more deals.

The entities for Bid Natures and Payment Types followed the same relation in that each Deal record could contain one of each respective primary entity's record and a many Deal records could be related to a single primary entity record. Unlike these one-to-many relationships, the Deal Attributes and Dispositions entities were governed by many-to-many relations in that each Deal record could contain one or more and both Disposition and Deal Attribute records related to multiple Deal instances.

The company related primary entities of Companies, Domiciles, States, Exchanges, Tickers as well as Industry Sectors, Groups and Sub-Groups, all follow a one-to-many relationship with Deals. Here, each primary entity's record would be contained once by a Deal instance and many Deal instances could be observed for any of the respective primary entities.

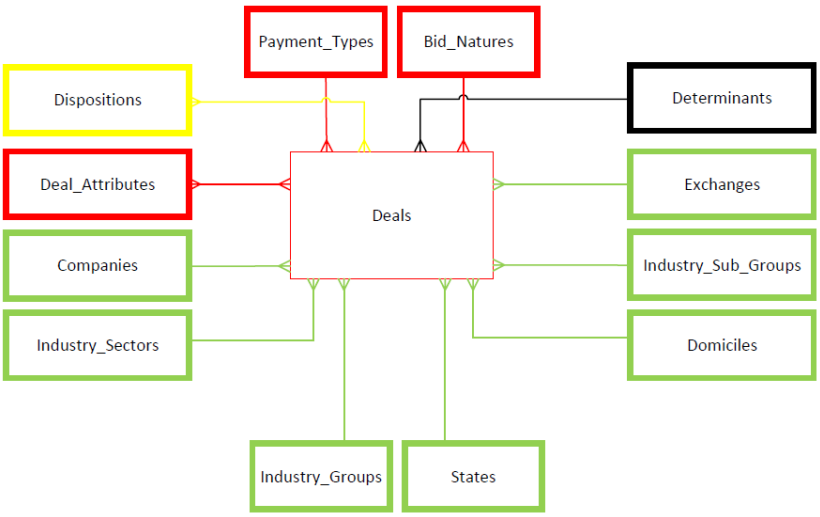


Figure 4.1: Validated Entity Relationship Diagram

4.3.3 Extended Entity Relationship Diagram

An Extended Entity Relationship Diagram (EERD) was then developed to enhance the details on relationships between entities within the database. Considering the design in Figure 4.2, the many-to-many relations of Deals and Dispositions as well as Deals and Deal Attributes received intersection entities, Deals Financials and Deals Attributes, respectively. These intersections were represented by round cornered rectangles. Once again, primary entities are depicted by thinker bordered rectangles.

The EERD introduced optionalities for relationships between entities. Excepting the Determinants entity, each one-to-many and many-to-many relationship was modelled using a "1" to identify that a record should exist in both entities before being used as an attribute of another table, enforcing referential integrity in parent and child table relationships. This was also in line with the data gathering method, in that all deals and respective attributes were endogenous to the search results, with no additional characteristics requiring storage.

The optionality of "0" for the many side of the Determinants and Deals relationship was used to allow for additional determinant records to be added to the entity without a corresponding existing Deals record. This would allow for storage and analyses of Determinants records beyond the collected deal set window period, in the event that a time lag was introduced for indicators, earlier than 1998 or, that a Deal was not recorded for any month of any year in the window, while a Determinant set was.

Intersection entities of Deals Attributes and Deals Financials were each modelled as non-optional relationships, as both required existing record keys from respective parent tables, towards creating each record's composite key.

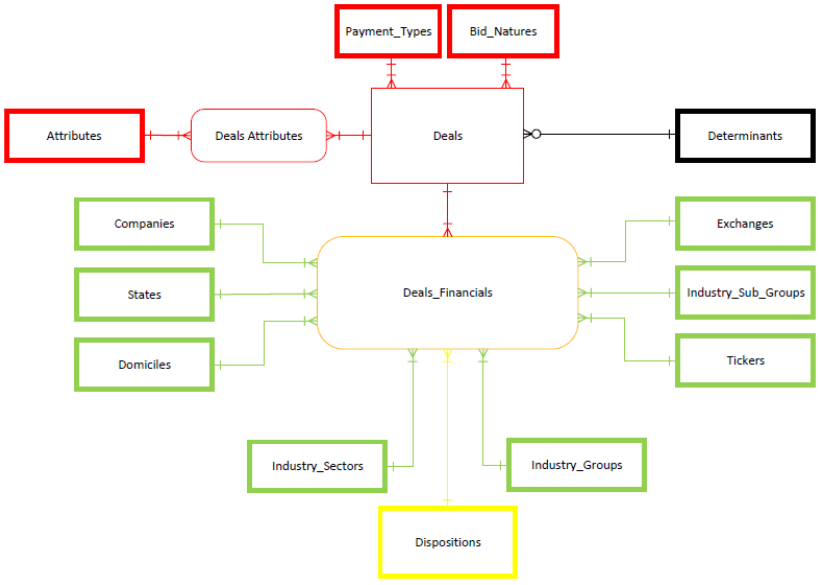


Figure 4.2: Extended Entity Relationship Diagram

4.3.4 Logical Design

The logical design produced a database schema that better reflected the physical design required for implementation of the database. The EERD was used as a structure for the schema, constituted by entities and relationships, while the initial database study helped inform the requirement of specific fields to be included within entities and their respective data types.

The database schema can be found in Appendix C, demonstrating the primary, foreign and attributive fields for each entity. The relations of which are demonstrated using the Crows's foot notation, as used in the entity relationship models. Here primary and foreign keys are indicated with the abbreviation PK and FK, respectively. Each field was allocated a data format type based on information gathered through the initial database study. A suitable data format type was allocated to fields by considering firstly, the format of gathered information and secondly, a comfortable precision factor based on the range of observed values for the field.

Suitable data types were decided upon, given the input data types. These are detailed further within the database schema. Primary Key data types were chosen based on realistic expectations for the range of records to be stored by each table.

4.4 Implementation

Owing to the author's prior experience with Microsoft Visual Studio, an integrated development environment (IDE), as well as its extensive functionality

made available for the Microsoft SQL Server DBMS, Microsoft Visual Studio Community 2017 was chosen as the database implementation platform. Phases of the database implementation are covered in the subsections that follow. These include database construction in Section 4.4.1, integrity assurance in Section 4.4.2,loading data in Section 4.4.3 and access and administration in Section 4.4.4.

4.4.1 Database Construction

Database construction required the translation of logical database design, presented in Section 4.3.4, into necessary SQL queries (to be interpreted by the DBMS) that would define the database schema. These queries stipulated the structure of entities, their data-types, the allowance of null values and the constraints that defined entity relationships, towards ensuring database integrity.

As the majority of fields for gathered data contained null values, emphasis in null value acceptance was given mainly to primary key fields. This was a measure implemented towards better ensuring entity and referential integrity, while still accommodating constraints of the gathered data.

4.4.2 Integrity Assurance

The use of normalization in parallel with the design stages of the database, better ensured the deterrence of potential data anomalies. By achieving 3NF for the gathered deal and determinant data, both entity and referential integrity was better assured. Still, the relevant SQL instruction for the database was required. Domain constraints were enforced through defining the correctly corresponding data type for fields in tables as informed by the database initial study and logical design.

Entity integrity was enforced by setting keys of primary entities to the non-null identity property and only allowing non-null values for all other entity primary keys. This ensured that all records loaded had an existing, non-repeating primary key. For the intersection entities of Deal Attributes and Deal Financials the identity property was not assigned (so as to allow for manual assignment), however, the primary key, a composite of the respective foreign keys, was defined as requiring non-null values for each. The identity property was also avoided for the Deals table primary key as the Bloomberg assigned action identifier (Action ID) was to be used and required manual insertion. The Action ID was inspected within MS Excel for duplication before loading data. In the event that a record was loaded without a corresponding primary key, the DBMS would deliver a warning message prompting the necessary remedial action.

Referential integrity was also aided by the use of the identity property for primary key values. This ensured that foreign keys used had an existing corresponding parent entity record when used as an attribute of another table.

The ON DELETE CASCADE command was also stipulated for all relationship constraint definitions. This meant that an entire child record would be deleted in the event that the corresponding primary key record was deleted in the parent table. Where a record was loaded without an existing parent table primary key, the DBMS would deliver a warning message prompting the necessary remedial action.

4.4.3 Loading Data

Using each established table's data graphical user interface (GUI), bulk data was copied and pasted from the corresponding normalized table in MS Excel.

4.4.4 Access and Administration

As the database was to be stored on the author's local computer drive, for sole personal use only, no additional access was granted for any other users or administrators. Database distribution and security aspects were not addressed and facilitated, as a result.

4.5 Testing and Evaluation

This aspect of the DBLC was primarily facilitated by a results comparison of analyses using the developed database, against results of the same desired analyses, using alternative software. The results of this comparison are presented in Section 5.1 of Chapter 5. Further, successful data loading, without error prompts from the DBMS, after implementing integrity assurance measures in database construction, was seen as further proof of effective implementation.

4.6 Operation

Database operation begun after successful testing and evaluation as well as the facilitation of analysis applications. Analysis application facilitation was effectively achieved using the IDE chosen for database implementation. This was because the IDE also facilitated R Tools for Visual Studio (RTVS) as an extension. This enabled seamless interaction between the DBMS and statistical analysis functionalities of MS SQL Server and R software (R Core Team, 2016*a*). Database views were called using Open Database Connectivity (ODBC). This was facilitated by the RODBC package (Ripley *et al.*, 2010), allowing for direct connectivity to established database objects.

4.7 Maintenance and Evolution

Given future availability and inclusion of additional or improved information, the database needed to be update-able. Thus, if the deal set were to be expanded, given a new search criteria for an altered scope of investigation, or with new deal or determinant information becoming available, new data could be efficiently added. The Database Life Cycle was found to be well suited to such a requirement and would effectively facilitate the redesign and development of an alternative or improved database.

In the case of additional deals being added to the deal set, with the same fields of information, the new deal set would need to be normalized, creating new tables of the same structure. These would need to be compared to existing entities' records. Unique records from new tables would need to added to existing tables, using the INSERT query.

In the case where the database table structure requires alteration, the implications of the change would need to be considered with regard to the established database initial study and design stages (DBLC re-execution). In the event that the database schema is changed through the introduction of new entities or relationships, the DBLC process will require a hard change. This would necessitate a completely new implementation and loading instance according to the new design. The subsequent stages of the DBLC would then be executed. In the event that an additional attributive field is created for existing tables, a soft change would be required. The established database schema would be updated using the necessary data definition and manipulation queries. In a specific example, additional determinant variables could be included by inserting values that correspond to the correct date of an existing Determinants record. This can be achieved using the ALTER TABLE query for the additional field creation within the table, as well as UPDATE, SET and WHERE queries for values that match existing records dates.

Chapter 5

Testing and Analyses

Chapter 5 presents the database testing and analysis results towards thesis Objectives 4 and 5. As stipulated in the overall thesis methodology, this served as an exercise aimed at testing the database's effective design and implementation. It also provided the means for activity analysis and evaluation of the determinants identified through the systematic literature review. Results of the database testing are covered through Section 5.1 while Sections 5.2 and 5.3 cover both activity and determinant analyses.

5.1 Results Comparison

Database testing and evaluation was primarily facilitated by a comparison of analysis results for queries on the developed deals database and the same analysis type on the original de-normalized table of deals, using MS Excel. Here queries were built to match the same functions of statistical analysis performed in MS Excel, using SQL Views and aggregate functions, as well as R Tools for Visual Studio. This was also an opportunity to deliver selected analysis results for the entire deal set sample. The Stargazer package (Hlavac, 2015) from R assisted in developing summary statistics tables.

Table 5.1 shows a sample summary of selected fields from the Deals table. Announced total value, net debt as well as terminal value EBITDA and equity book value ratios are tabled for respective statistics. Here after comparing values for from an MS Excel analysis of the de-normalized table, for mean and standard deviation values of the selected fields, no difference was found for the same values when using the R summary statistics function. Table 5.2 continues to consider company attributes of employee count as well as asset turnover and asset equity ratios for both acquirers and targets in the deal set. Again mean and standard deviation values were compared for an excel analysis and were found to be indifferent from R summary results for the deals sample, shown in the delta columns.

Table 5.1: R Summary Statistics and comparison of MS Excel Descriptive Statistics mean and standard deviation values for a sample of Deals table fields.

	R Summary					**MS Excel**		**Δ**	
Statistics	***Min***	***Mean***	***Med***	***Max***	***St.Dev***	***Mean***	***St.Dev***	***Mean***	***St.Dev***
Tot Value	1	1,739.42	190.04	186,235.80	6,673.66	1739.42	6673.65	0	0
Net Debt	-6,446.58	286.974	0	79,344.00	1,890.32	286.974	1890.32	0	0
TV/EBITDA	-5.29	30.473	12.01	3,392.95	141.659	30.473	141.659	0	0
EqV/BookVal	-6.81	5.726	2.44	1,347.20	31.384	5.726	31.384	0	0

Table 5.2: R Summary Statistics and comparison of MS Excel Descriptive Statistics mean and standard deviation values for a sample of Deals Financials table fields.

		R Summary					**MS Excel**		**Δ**	
	Statistics	***Min***	***Mean***	***Med***	***Max***	***St.Dev***	***Mean***	***St.Dev***	***Mean***	***St.Dev***
ACQS	No.of.Employees	1	27,334.10	3,885	653,300.00	63,798.89	27334.10	63798.89	0	0
	Assets/Turnover	-0.02	0.58	0.409	20.46	0.78	0.58	0.78	0	0
	Assets/Equity	-400.33	12.06	3.198	13,069.00	288.92	12.06	288.91	0	0
TARS	No.of.Employees	1	3,363.23	401.5	653,300	15,459.35	3363.23	15459.35	0	0
	Assets/Turnover	-0.41	0.81	0.638	13.44	0.86	0.81	0.86	0	0
	Assets/Equity	-533.49	4.186	2.473	239.48	13.957	4.19	13.96	0	0

Table 5.3: SQL Aggregate function values and comparison of MS Excel Descriptive Statistics total deal value for the Deals table.

	SQL Aggregate Results						**MS Excel**	**Δ**
Payment Type	***Deal Vol***	***Tot. Value***	***Avg. Value***	***Length***	***TV/EBITDA***	***EqV/B. Val***	***Tot. Value***	***Value***
Cash	1740	1749550.97	1005.49	101	30.06	5.84	1749550.97	0
Cash and Debt	107	59514.92	556.21	102	11.08	4.16	59514.92	0
Cash and Stock	627	2087015.88	3328.57	150	18.64	4.03	2087015.88	0
Cash or Stock	505	643442.62	1274.14	178	33.78	4.05	643442.62	0
Cash,Stock&Debt	30	91051.5	3035.05	185	14.2	3.23	91051.5	0
Debt	8	859.73	107.47	108	43.21	0.43	859.73	0
Stock	1692	3587235.98	2120.12	153	39.21	6.29	3587235.98	0
Stock&Debt	93	170999.29	1838.70	153	23.4	16.66	170999.29	0
Undisclosed	26	8235.48	316.75	81	15.32	5.36	8235.48	0

Table 5.3 delves more deeply into statistics for the Deals table by providing a breakdown of values for each payment type present in the sample. This time statistics results were produced using a database view of the Deals table and the relevant aggregate functions. All values averages per payment type except the sum of deal volume and the sum of total value. The deal length average value was derived using the DATEDIFF aggregate function for deal announcement and completion dates. Additional summary statistics for the numeric variables of the Deals table can be found in Appendix E.

Appendix D delivers summary statistics using R's summary function and continues with a comparison of mean and standard deviation values for the database determinants variables. Both values are found to be indifferent to MS Excel equivalent statistics over the sample deal period for each determinant.

5.2 Activity Analysis

Towards deeper merger activity analysis the following sections set out to explore trends for the sample of deals in terms of deal characteristics. The exploration begins with a broad perspective of deal activity, then continues to consider lower level variables with respect to time. Additionally, characteristics of more recent M&A activity are covered in the final subsection.

5.2.1 Aggregate Merger Activity

Figure 5.1 presents the time series of annual merger activity and value for the gathered data, between 1998 and June 2019. The deal value series, with units in millions of Dollars, shows greater evidence of peaks and troughs, in line with the wave like behaviour of mergers as suggested in literature. A clear trough is evident in the year 2002 in the deal value series, exiting the fifth merger wave. A subsequent depression is not obviously discernible between the years 2008 and 2013, however the peak in activity for 2005 suggests a wave was experienced but with an inconclusive ending year. Although values for the year 2019 were not complete (excluding deals completed after June 2019), a peak in activity is evident for year 2015, both in terms of deal value and volume. With the decline in activity relative to the 2015 peak, through the years 2016 to 2018, both series indicate that a seventh wave may be in progress and potentially be declining. Here, the sum of total deals and deal value for the periods reported are grouped by announcement dates. This is the standard used throughout the remainder of analyses.

Considering data for the 2019 year was incomplete, the average annual deals value is calculated using values from 1998 until 2018. The dotted red line provides a reference for the average annual value of the deal sample at 399,26 billion USD.

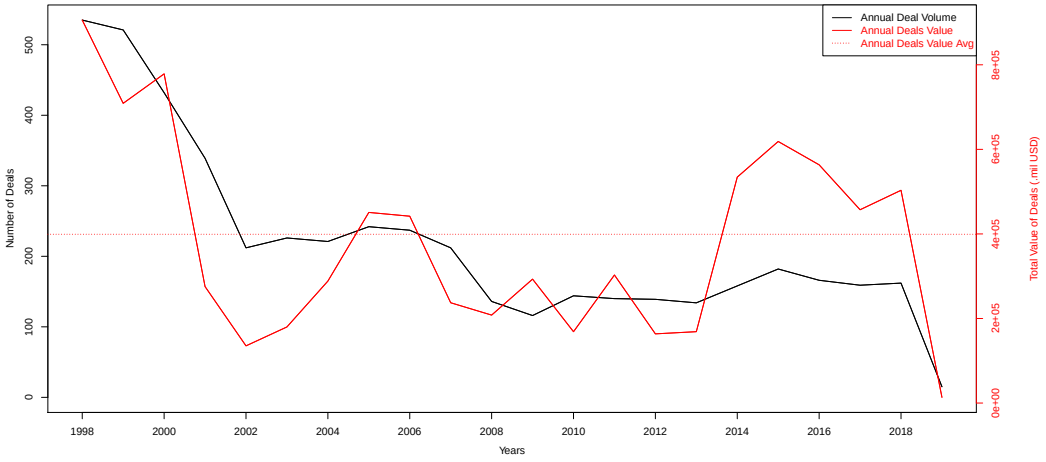


Figure 5.1: Annual total deal value and volume time series.

Overall, the degree of variation within the deal value series seems to exceed that of the volume series, visually. However, by comparing the standard deviation to mean ratio of each series from Table 5.4, they are found to be approximately, equally volatile at 0.61 and 0.59 respectively. By removing the years 1998 to 2001 from the series however, deal volume is greatly exceeded by deal value volatility using this metric. Deal value is close to double the volume volatility at 0.55 and 0.33 respectively. The correlation coefficients between annual deal value and volume was found to be 0.72.

Table 5.4: Annual deal value and volume summary statistics.

Statistic	Min	Mean	Med	Max	St. Dev.
Deals Volume	15	219.455	174	535	129.317
Tot Value	13,480.61	381,723	297,941.9	905,831.6	233,091.1

Figure 5.2 de-aggregates annual merger activity and presents monthly figures for deal value and volume for the gathered M&A data. The red series demonstrates total monthly deal value (in USD millions) behaviour over time, compared to a series for monthly number of deals. A mean value for each series over the sample window is also presented. This highlights the sharp decline in total monthly deals after 2008 and an increasing trend of high deal value for months of the 2014 and 2018 period.

Comparing the standard deviation to mean ratio for monthly deal volume and value, activity volume is found to be less volatile in change than deal value

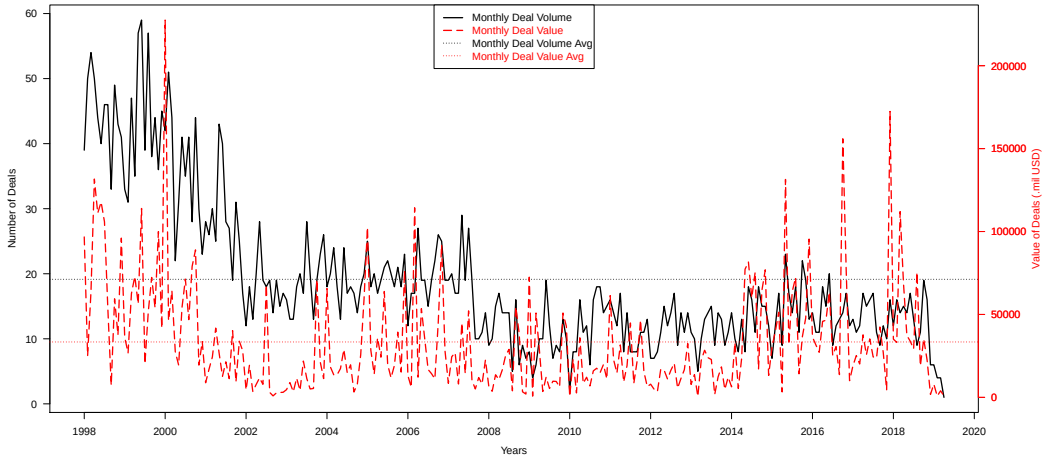


Figure 5.2: Monthly total deal value and volume time series with average annual deal value.

at 0.6 and 1, respectively. These figures are derived from Table 5.5 where deal value is measured in millions of USD and deal volume, by the number of deals.

Table 5.5: Monthly deal value and volume summary statistics.

Statistic	Min	Mean	Med	Max	St. Dev.
Deal Vol	1	18.859	16	59	11.292
Deals Value	480.170	32,804.320	22,535.390	227,533.500	32,683.560

Considering a new activity variable, deal length, Figure 5.3 presents the time series for the average annual deal length, compared to the average value of deals for each year. A valid relationship is evident for the pair, in that the more valuable deals become, the more time will be spent by deal makers and consultants in completing them. A correlation coefficient of 0.44 exists for the two variables. This does not improve for lagged values of both series. An exception to the general correlation occurs in the 2008 and 2009 years where an inverse relationship seems to exist, suggesting deals were completed more quickly through the 2008 financial crisis even as the average value increased.

Data for deals in the 2019 year were not included in the graph owing to the rapid decline in average deal length for the 2019 year (75 days). This seemed out of order but was found to be a fair reflection of the deal set sample. The reason for this was that only shorter deals would have been included in results, owing to the deal completion criteria of the data gathering search protocol. Here, if deals announced in the early months of 2019 were completed after June 2019, they would have been excluded from deal results and therefore,

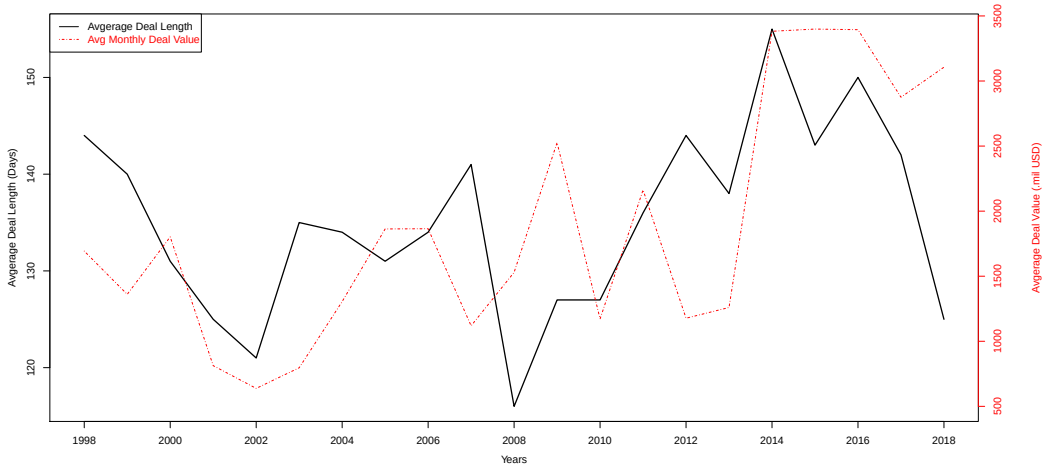


Figure 5.3: Average annual deal length and monthly average value of deals for each year.

biased the resulting sample of 2019 deals. This, as only quickly completed deals would have made the resulting deal set.

Considering companies that made multiple acquisitions throughout the deal set, Appendix F presents a list of the top 30 serial acquirers. The average number of deals and announced deal values, completed by these companies, was found to be 15 deals and 3.884 billion USD, respectively. These companies offered higher premiums in their acquisitions, with an average of 43.22 announced premium percentage, as compared to the deal set sample mean of 38.82 percent. The average deal length for the top 30 serial acquirers was 11 days shorter than that of the greater deal set. Considering the average deal value was more than double the average value for the sample, experience seems to improve deal completion time.

5.2.2 Annual Trends

Results for the annual trends section exclude deals announced in 2019, as figures for the year were incomplete. By considering more detailed activity variables, a break down of deal payment type, annually, is presented in Figures 5.4, 5.5 and 5.6.

The first set of figures show absolute deal volume and value per payment type with respect to time. Here, payment types with fewer than five deals or a total of 3 billion USD (1% of mean annual total value) were excluded for ease in presentation. Figure 5.4 shows a significant decline in stock only payments after 2002 (5th merger wave ending), as acquirers moved towards cash only payments or combinations thereof. This is further confirmed when considering total deal value of each payment type in Figure 5.5, where the shift is more

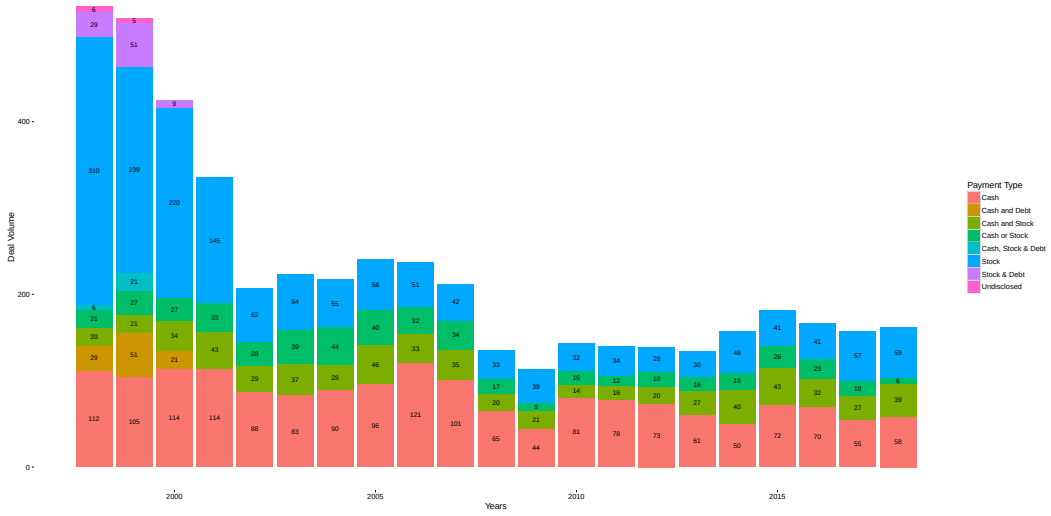


Figure 5.4: Annual absolute deal volume in terms of payment type.

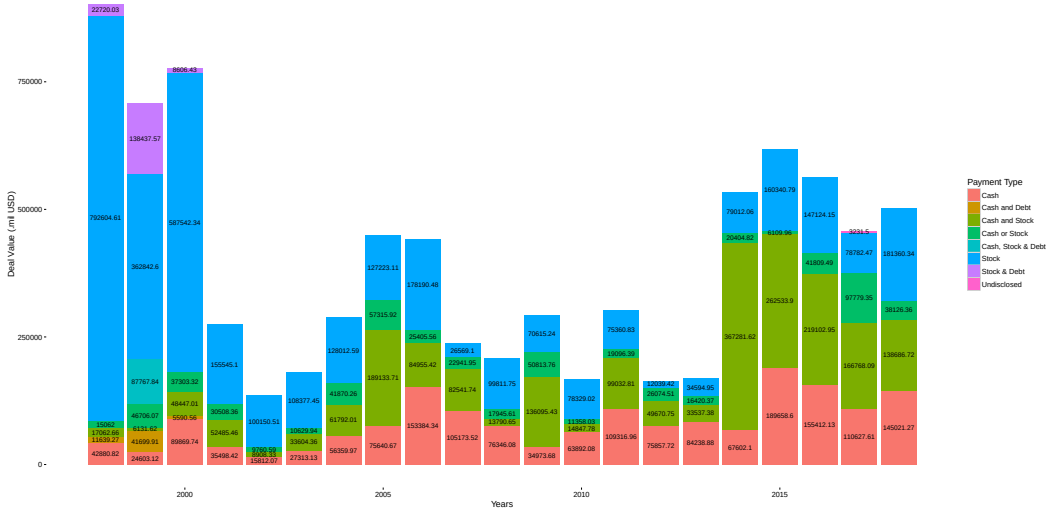


Figure 5.5: Annual absolute total deal value in terms of payment type.

apparent. This may be because larger deals involved more stock payments and thus a sharper value decline in this trend is more evident. Deals involving stock payments still persist through the years 2005 and 2018 with varying levels of cash only and supplementary cash contribution payments offsetting stock only deals.

Upon deeper inspection of relative payment type trends, Figure 5.6 presents proportional total value annually. The shift away from stock only payments into cash only and cash component payments is more evident here. Interestingly, the end of the 5th merger wave is characterized by the immediate shift away from stock only payments while the end of the 6th merger wave sees a

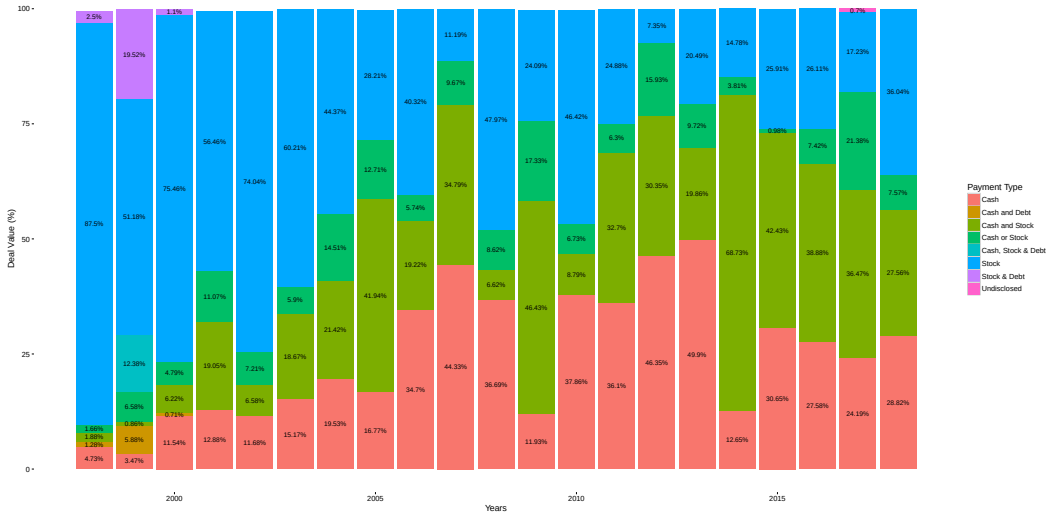


Figure 5.6: Annual proportional total deal value in terms of payment type.

strong reversal in total value for the payment form.

Figures 5.7 and 5.8 explore attributes of deals for the deal set. Company takeovers dominated the sample (4684 deals) as a deal type characteristic and were therefore not considered in the analysis of deal attributes. This meant that the total deal set was not represented in the sample but rather just the deals that had at least one additional attribute categorization. Total deal values for a deal attribute below 3 billion USD were excluded from deal value attribute activity of Figure 5.8.

No obvious trends are evident when considering attributes in terms of volume or value in time, although the value of tender offer deals peaks significantly for the years 2000, 2005 and 2015. When considering the volume of tender offers, the number of deals does not vary as greatly, proving the average deal value for peak tender offer years is higher than normal.

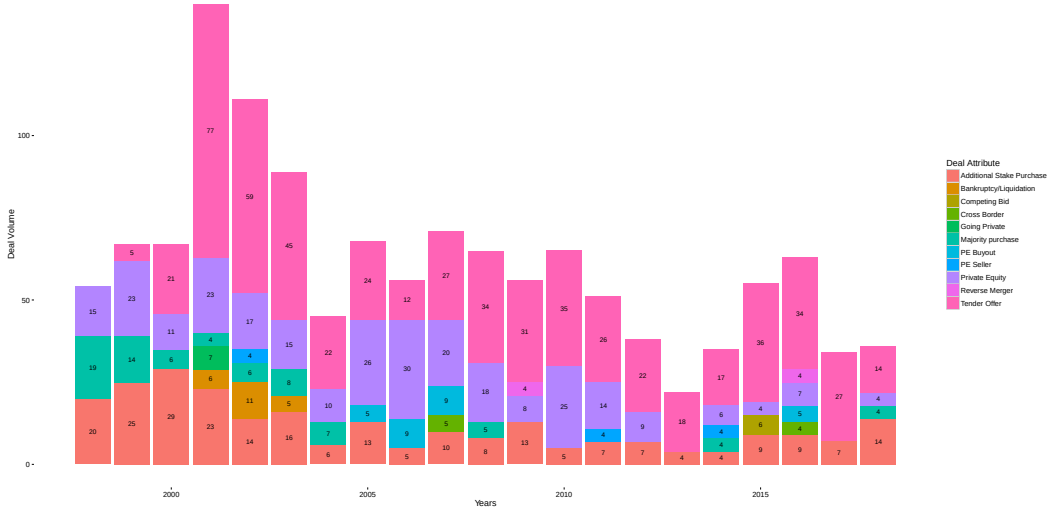


Figure 5.7: Annual absolute total deal volume in terms of deal attributes.

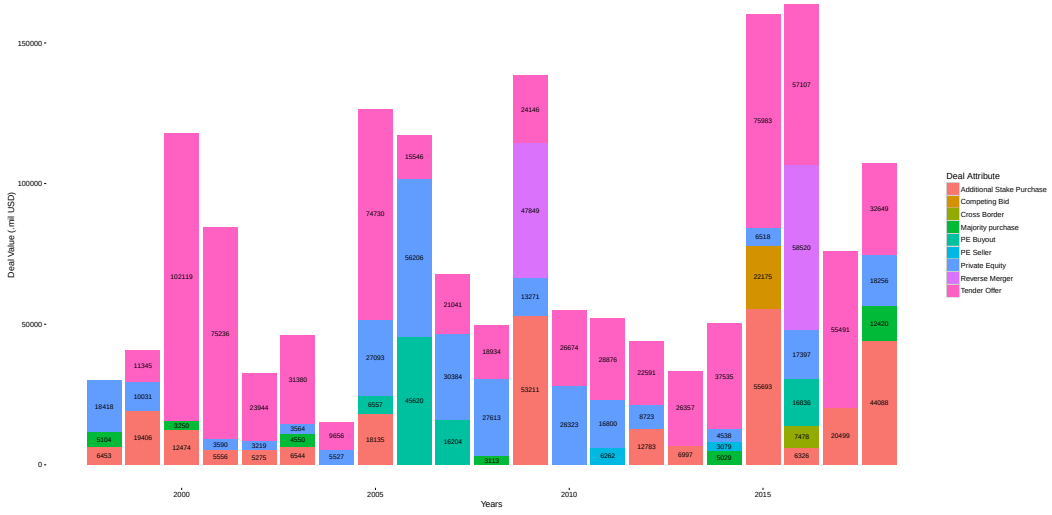


Figure 5.8: Annual absolute total deal value in terms of deal attributes.

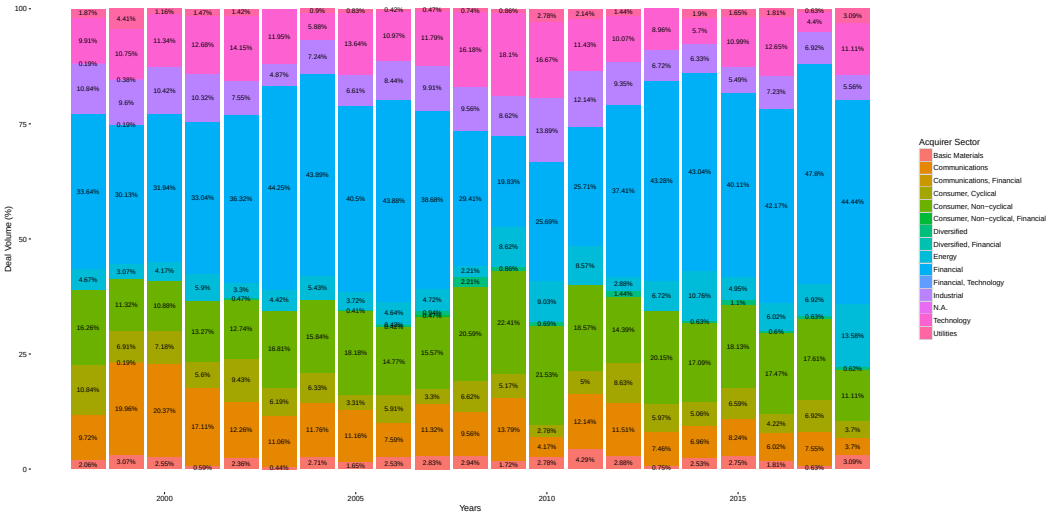


Figure 5.9: Annual proportional total deal volume in terms of acquirer sector.

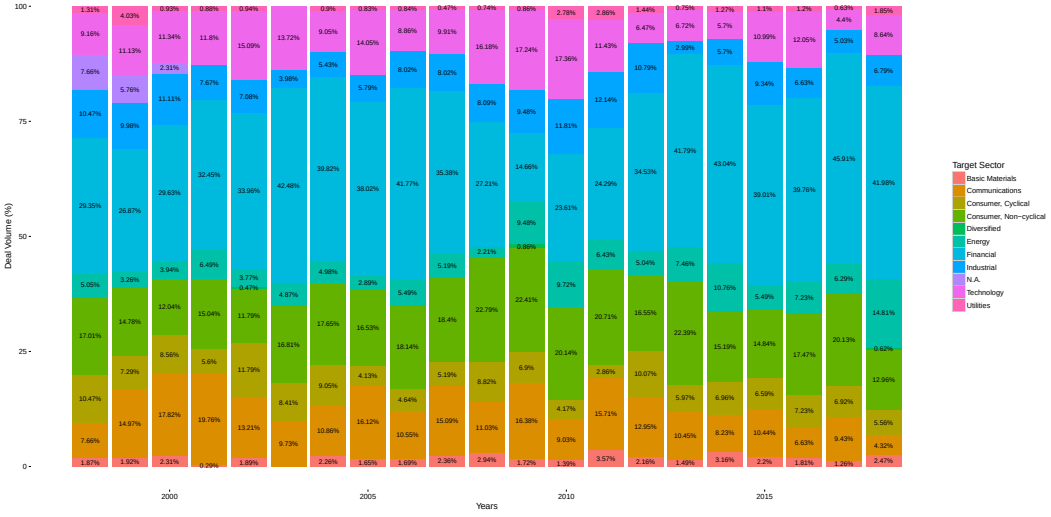


Figure 5.10: Annual proportional total deal volume in terms of target sector.

Figures 5.9 and 5.10 provide the annual proportional breakdown of deal volume for acquirer and target sectors. Steady proportional trends are observed for acquirer and target sectors, with the greatest deviations occurring for financial sector deals, in the past 21 years. The early 2000's, around the "dot-com" bubble, saw an increase in communications sector deals while financial sector deals retreated to their lowest through the great recession of 2008-2009 but recovered soon afterwards. The utilities sector has maintained a steady proportion of deal activity for the period observed.

Delving deeper into level of diversification in deals, Figures 5.11 and 5.12 present total and proportional deal value for deals involving sub-group di-

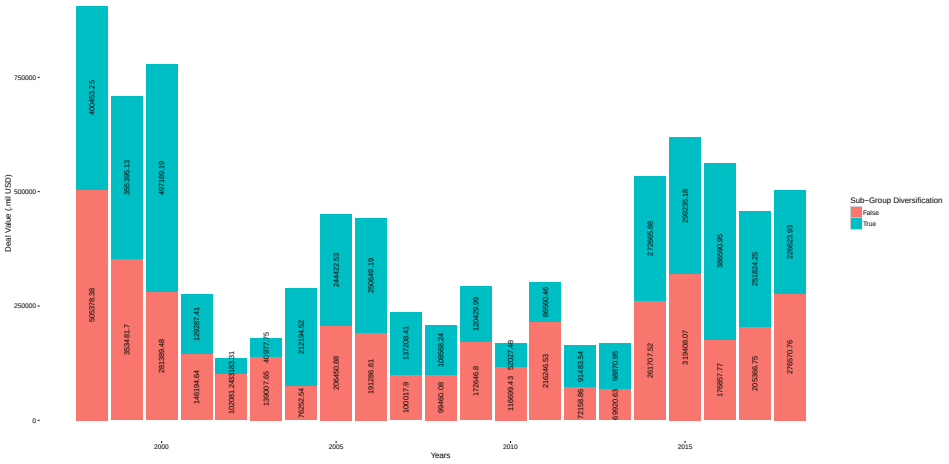


Figure 5.11: Annual absolute total deal value in terms of deal attributes.

Figure 5.12: Annual proportional deal value for sub-group diversification.

versification. By observing the proportional sub-group diversification deals value through the sample period, seasonal trends are discernible, particularly through the years, 2000 to 2011. The highest peaks of industry consolidation occur for the years 2001, 2002 and 2011 after experiencing upward trends, following recessions. Firms seem more likely to acquire companies within their own industry, following periods of economic turmoil, if at all. Figure 5.11 creates context by displaying trends of absolute annual deals value over time.

Isolating sub-group diversification deals and plotting them against the annual absolute value of mergers in Figure 5.13 indicates that correlations exist for both activity volume and value. This suggests that the degree of diversification through merger industry drives deal activity value, also evident in Figure 5.11 where heightened deal value is accompanied by increased diversification.

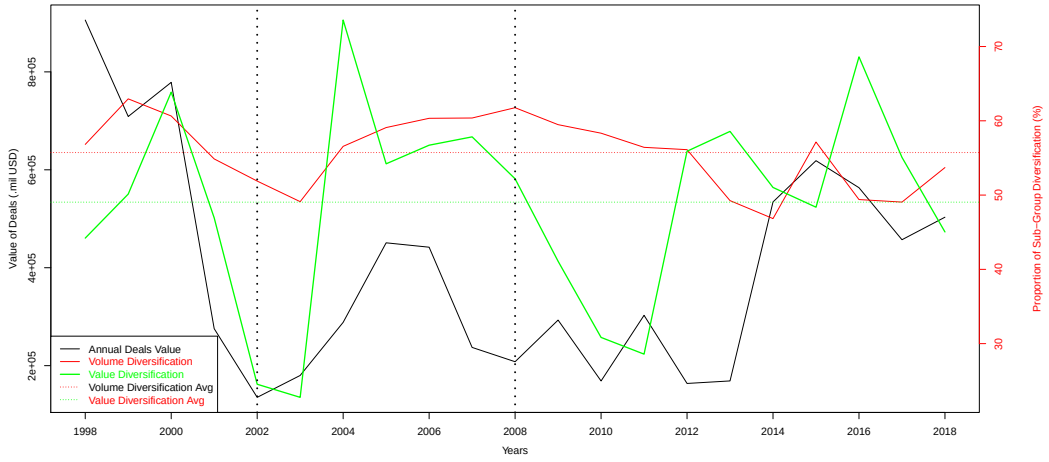


Figure 5.13: Annual proportion of sub-group diversification and total deals value.

5.2.3 The Seventh Merger Wave

Although an absolute trough is not visually discernible following the sixth merger wave, it is still apparent that a definite peak in activity was experienced in 2015, both for merger volume and value. This rise in activity had to stem from somewhere and thus an attempt at justifying a start date for the wave was made. Given two successive years of flat total deal value in 2012 and 2013, followed by a sharp increase in value as well as a slower, but still present, incline for deal volume, 2013 was suggested to be a suitable year for the start of the seventh merger wave. Evidence supporting this can be found in Figure 5.1.

It is also apparent that the characteristics of activity for the 2013 and 2018 period have changed from that of the 6th merger wave. Most noticeably, The average deal value significantly increased for the period, giving way to an era of mega mergers. Average deal length increased significantly in accordance with this.

The merger wave has also seen a significant increase in deals involving combinations of cash and stock payments to compliment higher average deal values, leaving an era dominated by stock only payments behind. Figure 5.5 shows that the value of cash only deals is significantly greater for the seventh merger wave, although used in fewer deals. Tender offers have been the popular deal type experienced through the seventh wave, potentially stimulated by cash contributions in payment. This can be seen in Figure 5.7.

Strong trends in financial sector dominance of activity continue after subsiding, following the sixth merger wave. Other noticeable sector involvement trends evident in Figures 5.9 and 5.10 for more recent years include an increase in energy sector involvement and a decrease in the communications sector. The level of sub-sector diversification has remained stable around a peak in 2016.

More recent years have seen a retreat to mergers within the same sub-sector.

Only once figures for the 2019 year become more complete, could conclusions about the wave's continued support or decline be made. Considering the level of global economic uncertainty, driven by prospects of a trade war, it would be assumed that companies would lean towards organic growth and survival, within the mid-term future. The resulting disruptions to current market equilibrium could provide support for M&A activity through industry consolidation though.

5.3 Determinant Analysis

An analysis of determinant influence was executed in order to better evaluate the effectiveness of introducing additional variables to merger activity prediction. This was achieved using stored determinant proxies and aggregated deal volume and value data from the data warehouse. Before executing the determinant evaluation, the necessary data pre-processing was completed, eradicating irregularities within the data. This is presented in Section 5.3.1, where-after, an analysis rationalization was performed in Section 5.3.2, using a univariate analysis method for merger activity modelling. This provided a reference point within the solution space before considering the effectiveness of multivariate methods involving determinants. The determinant evaluation is presented in Section 5.3.3 where feature selection and multiple regression methods explore the influence on M&A activity. Appropriate packages within R were used to supplement analysis methods and were cited accordingly.

5.3.1 Determinant Data Pre-processing

After finding that the average deal length for the sample was 135 days (approximately 4.5 months) values recorded for 2019 were removed from further analyses. This was because the search criteria in data gathering stipulated that deals were announced between 1998 and June 2019 as well as were completed. This would mean a fair proportion of deals were excluded (given that they had not been completed) for more recent months, but more specifically for January to June of 2019, creating potential biases within the sample. The analysis of determinants would continue to use figures from 1998 until 2018.

Towards creating determinant variables of an annualized frequency, the average annual value for monthly and quarterly figures was used to aggregate data for each determinant. This was done for each variable except the bankruptcy filings variables and annual figures for total housing starts and the fiscal freedom index. The sum function was used to calculate the accumulation of figures for the year in total bankruptcy filings while the annual figures for housing starts and the fiscal freedom index were used in their existing format.

Missing values for the newly created annual determinants variable set were imputed using natural spline method of the spline() function from the Stats package (R Core Team, 2016*b*). This was required for 2018 values of the IMF USA GDP and IMF US GDP Quarterly change indicators as well as 1998 and 1999 values for Total Bankruptcy Filings.

A late discovery of errors in company financial information for the deal set lead to the exclusion of potentially valuable Firm-level determinants for merger activity. This was a flaw found in the extraction of data from Bloomberg. Here, instead of the financial data being returned from financial reports dated closest to the deal, the results returned were for a companies latest financial report, closest to the date of data extraction from Bloomberg. It was decided that the exclusion would be most appropriate as meaningful results would be jeopardized.

5.3.2 Analysis Rationality

Before multivariate models of various determinant configurations were built and tested, it was worthwhile to first establish a benchmark for model usefulness, using a conventional method. By using a univariate time series modelling method, a standard for model accuracy could be determined as a reference point. This helped better define the possible solution space for the problem, before multi-variate models and their respective independent variables could be evaluated on predictive capability as well as relationship nature and significance.

Using the auto.arima() function from the Forecast package (Hyndman, 2016), the best model parameters were estimated for an ARIMA model, capable of explaining respective time series behaviours. The resulting observed and fitted time series for monthly and annual deal volume and value can be found in Figure 5.14. While annual models closely follow observed values, both monthly fits struggle to account for the significant deviation experienced within each activity series.

The resulting parameters for the ARIMA($p,\ d,\ q$) model estimation of number of time lags, degree of differencing and order of moving average, were as follows:

- Monthly value - ARIMA(*0, 1, 1*)
- Monthly volume - ARIMA(*3, 1, 3*)
- Annual volume - ARIMA(*1, 1, 0*)
- Annual value - ARIMA(*1, 0, 0*)

Based on the estimated models, the annual value series was satisfactorily stationary with a differencing factor d of zero. Both annual volume and value

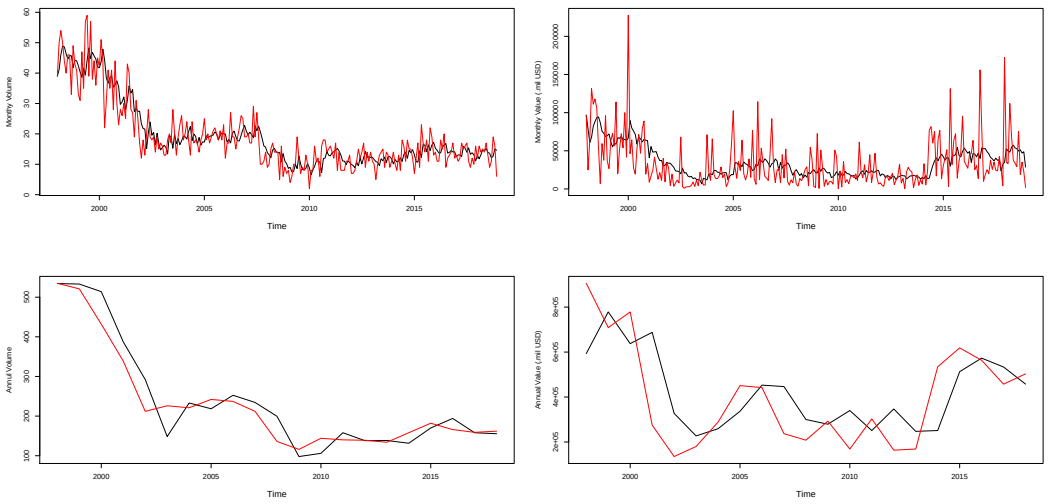


Figure 5.14: ARIMA model fits (black time series) for monthly and annual volume and value (red time series).

Table 5.6: Accuracy statistics for ARIMA models.

Statistic	ME	RMSE	MAE	MPE	MAPE	MASE
Mon Volume	-0.34	5.16	3.86	-9.74	25.37	0.66
Mon Value	-1360.06	29522.71	20702.25	-156.26	183.65	0.75
Ann Volume	-8.67	38.40	28.18	-2.90	13.85	0.94
Ann Value	-21677.81	164628.60	126206.90	-26.28	44.30	0.97

use a single lag factor in predicting values, indicating the presence of auto-correlation within each series. The monthly sets rely more on a moving average window in predicting values. This is evident in the overall trend tracking ability of fitted values from Figure 5.14.

The resulting accuracy statistics for the each model fit can be found in Table 5.6. Each deal value series is expressed in millions of USD while volume series are expressed in number of deals. Further evidence of each monthly model's inability to track higher degrees of deviation is found. Deal value models preform worst when comparing to deal volume predictions. The best resulting model accuracy is produced for annual deal volume with a mean absolute percentage error of 13.85%.

5.3.3 Feature Ranking and Determinant Evaluation

Determinant evaluation began by identifying potential redundancies that may exist for gathered determinant variables in the set. If redundant variables were used in concurrence within an explanatory model, unwanted multi-co-linearity may have been introduced, reducing model effectiveness. Correlation coeffi-

cient statistics for each variable were produced using the ggcor() function of the GGally package (Schloerke *et al.*, 2016). The resulting Pearson correlation coefficient matrix for merger volume, value and gathered determinant proxies can be found in Appendix G. Here, absolute correlation coefficient values above 0.5 are highlighted for emphasis on variables that may introduce multi-co-linearity to a model. The threshold was chosen as a conservative estimate for unwanted correlations.

After considering the nature of determinants and their respective correlation coefficients, Chapter 11, 12 and 13 Bankruptcy Filings were removed from all further determinant evaluation as they were found to be too specific to certain company types. Instead, Total Bankruptcy Filings was used as a suitable proxy, alone. The IMF US GDP indicator was chosen as the sole proxy for evaluation of US GDP influence on merger activity. The decision to exclude quarterly change GDP indicators meant the influence of absolute US GDP behaviour could be evaluated, alone. The money supply indicator, M1 Money, was also removed from evaluation as it was found to be a subset of M2 Money and was determined to be exactly correlated with the latter. Table 5.7 reports the determinants considered through the remainder of analyses along with the respective theoretical classification. Following Table 4.5 of Chapter 4, determinant classifications B, M and N stand for Behavioural, Macroeconomic and Neoclassical theories. Again, where parentheses are present, the allocation is deemed weak as the proxy was not explicitly identified through the literature review, but classified at the authors discretion.

Owing to significant deviation within monthly volume and value merger series, it was decided annualized activity would be used in evaluating determinants and their prospective predictive ability. Additionally, with various drivers potentially playing a role at different times (months), in the lead up to a companies deal announcement, the level of desegregation would force assumptions about causal timing of determinants. Therefore a higher degree of aggregation was deemed more appropriate in evaluating merger determinants. It was also critical to improve upon the best ARIMA model accuracy statistics of annual volume in demonstrating the true usefulness of a multivariate predictive model for merger activity.

The cross validation generalized linear model function, cv.glmnet(), from the Glmnet package (Friedman *et al.*, 2010), was used as a means for initial feature evaluation of determinants. The algorithm was employed as a means for determining the independent variables deemed most useful in predicting the dependent merger activity outcomes by evaluating mean square error results for numerous configurations of predictors. The function's alpha parameter was set to one, establishing the use of the LASSO as the model to drive the algorithm while the default k-fold parameter was left at ten. The set.seed() function of the Base R package (R Core Team, 2016*c*) was used to create reproducible results for the algorithm.

The resulting variable importance ranking for the cross validated LASSO

Table 5.7: Determinant evaluation set.

Abbreviation	Determinant	Classification
Eco_Pol_Unc_Ind	Economic Policy Uncertainty Index	B
Pol_Unc_Ind	Trade Policy Uncertainty Index	B
S.P500	S&P 500 Index	B, M
S.P500_PE	S&P 500 PE Ratio	B, M
S.P_EPS_Q	S&P EPS + Surprise	(B), M
Imp_Int_Rt	Implied Interest Rate	M
US_10YR_Yield	Ten Yr Government Bond Yields	M
US_2YR_Yield	Two Yr Government Bond Yields	M
US_Res_Indic	Reccession Indicator	M
Prime_Int	Prime Interest Rate	M
Ind_Prod_Delta_MOM	Industrial Production MOM %	M
CFNAI	Chicago Fed National Activity Index	M
Mon_Sup_M2	M2 Money Supply	M
Unep_Rt	Unemplyment Rate %	M
CPI_Delta_YOY	NSA CPI YOY Index	(M)
Gold_PX	Gold Spot Price	(M)
Corp_Yield_Ret	Corporate High Yield Total Return Index	(M)
House_Strt	Annual Housing Starts	(M)
IMF_US_GDP_Fit	Real GDP	M, N
Fis_Free	Fiscal Freedom Index	(M),(N)
USD_Ind	US Dollar Spot Index	N
Fed_Fund_Eff_Rt	Federal Funds Effective Rate	N
Tot_Bankrup_Q_Fit	Total Bankruptcy Filings	N

regression, using mean square error (MSE) as the evaluation metric, can be found in Figure 5.15. Here, the varying configurations of independent variables used to predict the respective dependent variables, deal volume and value, are displayed along the top of each plot with the coefficient shrinking parameter, lambda, for each respective model variant on the x-axis. Each graph displays two dashed vertical lines indicating the lambda for lowest resulting MSE (left) and the lambda where error is within one standard deviation. The plots indicate a closely competing number of variables for one standard deviation of MSE for deal volume while a wider variety of variable configurations could be used to determine deal value. By extrapolating for trends within each graph, it is assumed that employing more variables in predicting deal volume improves accuracy while deal value accuracy diminishes through the same tactic. Determinants for the minimum MSE of each LASSO were identified and set aside for further analysis.

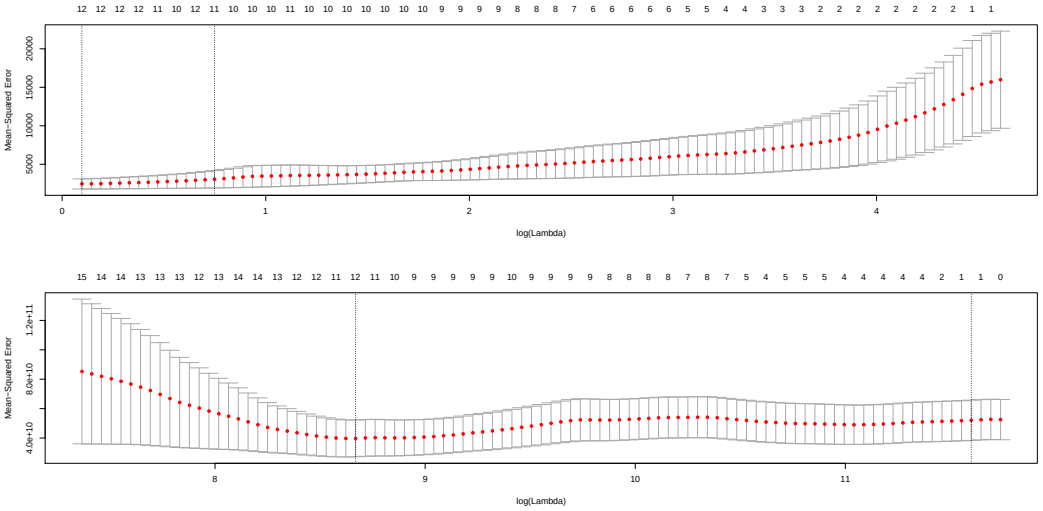


Figure 5.15: Variable importance for LASSO regression in deal volume (top) and deal value (bottom) prediction using MSE evaluation.

Multiple regressions were used to model both annual deal value and volume using resulting variables from respective LASSO models. Figure 5.8 presents the results for each model. Here, independent variable coefficients (first variable value), standard errors (values in parentheses) and significance (starred values) are reported for each regression model. Both multiple regression models were significant with p-values of 0.001796 and 1.005exp(-6) for deal volume and deal value models, respectively. The R squared values indicate a strong goodness-of-fit for each regression with the adjusted R squared value for deal volume suggesting that additional independent variables do not significantly hinder model performance. The lower adjusted R squared value for the deal value regression indicates that certain additional independent variables reduce the models ability to explain variance of the dependent variable, deal value.

Table 5.8: Regression Results

	Dependent variable:	
	Annual Volume	Annual Value
	(1)	(2)
Imp_Int_Rt	0.026	−1,132.600
	(0.566)	(2,006.268)
US_Res_Indic	88.879*	
	(42.288)	
US_2YR_Yield		161,990.200***
		(35,083.740)
CPI_Delta_YOY	−15.090	−78,999.740*
	(9.016)	(38,895.000)
Prime_Int	37.324***	
	(6.471)	
USD_Ind		−5,523.526
		(3,141.624)
CFNAI	160.997***	326,463.300**
	(40.260)	(107,321.600)
Pol_Unc_Ind	−0.133	−1,646.467
	(0.293)	(1,002.309)
Eco_Pol_Unc_Ind		−543.643
		(693.282)
Ind_Prod_Delta_MOM	−116.008*	−393,079.700*
	(54.399)	(188,728.600)
Unep_Rt	−2.704	
	(7.613)	
S.P_EPS_Q	−1.666	−12,815.790
	(2.481)	(10,876.950)
IMF_US_GDP_Fit	−8.605***	−3,522.653
	(1.452)	(6,507.490)
Fis_Free	−12.668***	−59,792.780***
	(2.318)	(12,422.750)
Tot_Bankrup_Q_Fit		0.083
		(0.093)
Constant	1,998.271***	6,444,416.000***
	(263.074)	(1,335,620.000)
Observations	21	21
R^2	0.984	0.934
Adjusted R^2	0.965	0.836
Residual Std. Error	23.301 (df = 9)	90,583.190 (df = 8)
F Statistic	50.651*** (df = 11; 9)	9.478*** (df = 12; 8)
Note:		*p<0.1; **p<0.05; ***p<0.01

Although macroeconomic indicators were more greatly represented after data gathering and subsequent cross validated LASSO model selection, little evidence of significantly influencing determinants for alternative theories is found within each regression model. The variables, Prime Interest Rate, the CFNAI Index, Real GDP and Fiscal Freedom Index were all statistically significant in explaining deal volume behaviour, supporting the macroeconomic theory of mergers. Only Real GDP confirms the relevance of the Neoclassical theory for the model albeit multi-classified as a Macroeconomic constituent as well. The Prime Interest Rate significance supports the assertion by Harford (2005) that necessary capital liquidity be available to stimulate merger waves when they occur. However, the positive coefficient defies the acknowledged effect, that M&A activity is stimulated by lower interest rates. This may have been a bias introduced from within the sample period of their study, being 1980s and 1990s, where companies may have tended to issue corporate bonds as a means for payment more frequently than cash or stock. The significance of the variable still proves its importance as a determinant and the inverse of the assumed relationship with capital liquidity may be a result of trends in increasing cash and stock deal payments.

The negative coefficients for Real GDP and the Fiscal Freedom Index suggest that merger volume rises during tumultuous economic periods and when tax rates increase, where a lower value of the latter proxy implies a higher tax burden imposed by government. While the Fiscal Freedom Index was not found to be explicitly categorized within the Neoclassical theory, there is still merit in assuming it represents a shock to companies in the form of regulatory change. A positive correlation for the CFNAI Index contradicts the established economic hardship effect on mergers (Real GDP coefficient relationship), as it measures growth and decline in activity and consequently, should be negatively related by this assumption. Interestingly, no significant variables were found to support Behavioural theory proxies as drivers of deal volume.

When considering the deal value regression results, fewer determinants were found to be adequately significant. Determinant proxies of Two Year Government Bond Yields, the Chicago Fed National Activity Index and the Fiscal Freedom Index were below the statistical significance threshold. Once again, Macroeconomic proxies were found to play dominant roles, this time in explaining deal value. The positive coefficient for Two Year Government Bond Yields suggest that annual merger value increases when yields increase. As these yields are positively related to prime interest rate, another argument against the role of capital liquidity in lubricating merger activity is raised. With the CFNAI Index measuring growth in economic activity, a positive coefficient indicates that any up turn equates to larger deal value. Once again, a significant Fiscal Freedom Index was found to be negatively related to merger activity. This further confirms the influence of tax obligations and its role merger activity prediction.

The Farrar-Glauber Test was used to test for multi-co-linearity within each

multiple regression model. This was performed using the imcdiag() function of the Mctest package (Imdadullah and Aslam, 2016). No multi-co-linearity was detected for variables in each model.

5.4 Remarks

Through a general exploratory analysis and comparison of results for analysis executed in MS Excel, the implementation of the M&A deals and determinants data warehouse was deemed successful.

Although a suitable statistical method for the seventh merger wave identification was not applied, the activity analysis in Section 5.2.3 still provided evidence of its existence. Considering visual interpretation of the deal value time series primarily as well as a comparison of characteristics for the sixth merger wave, a strong distinction can be made between recent merger activity and that of the previous wave.

Through the analysis and evaluation of determinants in Section 5.3, evidence is found in support of both the Macroeconomic and Neoclassical theories of mergers, when seeking to predict merger volume. Deal value is also found to be most strongly influenced by Macroeconomic factors. No significant predictor variable was found in favour of the Behavioural theory of mergers, using the gathered determinant proxies. However, this evaluation only considers variables significant for the p-value threshold of 0.05 and does not regard other contributing variables to the model. These additional variables, although not all significant, still collectively contribute in explaining aggregate response variables of merger activity better than optimal univariate ARIMA models. For this reason, these determinants should collectively be recognized in their capacity to predict merger activity. However, additional robustness checks are required for the model, beyond the Farrar-Glauber Test for multi-co-linearity, that evaluate the more intricate effects of the CFNAI Index and Real GDP proxies on deal volume prediction.

Table 5.9 presents accuracy statistics for the respective multiple regressions presented in Table 5.8. Additionally, a comparison of accuracy statistics for univariate ARIMA models of merger activity is given through a percentage improvement for the multivariate models. The results prove the overall useful-

Table 5.9: Regression accuracy statistics with percentage improvement on univariate ARIMA models.

Statistic	**RMSE**	**MAE**	**MAPE**
Volume Regression	15.25	11.87	5.73%
Value Regression	55909.16	50001.03	18.07%
Δ Volume ARIMA	60.29%	57.88%	58.62%
Δ Value ARIMA	66.04%	60.38%	59.21%

ness in considering determinant variables when predicting M&A deal volume and value.

Chapter 6

Conclusion

In concluding the thesis, the project summary, a review of initial aims objectives and limitations of the study are presented in Sections 6.1, 6.2 and 6.3. Recommendations for future research follow in Section 6.4.

6.1 Project Summary

A general introduction to aspects of the M&A field, in Section 3.1, provided a foundation for inquisition of activity analysis and subsequent characterizations. Further, the use of the systematic quantitative literature review method, presented in Section 3.2, allowed for a more focused examination of established determinant theory for the field. This proved to be a valuable means for identifying and synthesizing a collection of traditional drivers examined in literature. Ultimately, this assisted in the acquiring necessary knowledge for an analysis of the field.

Once information requirements were established, relevant data was sourced according to a defined search protocol, constrained by the scope of study in Section 1.4. The literature review method also assisted with the demographic analysis of articles in the sample. Various superficial attributes of documents were gathered and analysed to create context for respective contributions to the field.

Supplementing the Data Warehouse Architecture with the Database Life Cycle proved to be appropriate for supporting the development and implementation of a M&A transaction and determinants data warehouse. The process, executed through Chapter 4, detailed the respective design and development stages as informed by Sections 2.1.1 and 2.1.3 of the thesis methodology. Entire data warehouse functional potential was not fully utilized, considering that database Views were the predominant data element for supplied data towards analyses in R. If these tables were replicated in MS Excel and loaded to necessary R data structures, the same analyses could have been facilitated. Views did however, formalize the orientation for structured data elements from

the stored data without unnecessarily storing duplicated information. If the analysis application required greater use of dynamic searches within the data warehouse, these would have been well facilitated by stored routines, where arguments used in R functions could have been passed through the ODBC as arguments of a stored procedure. This would have performed more efficiently than an MS Excel equivalent facilitation.

M&A activity analysis, executed through Sections 5.1 and 5.2 of Chapter 5, provided both cross sectional and longitudinal activity analyses through the use of summary statistics, time series analysis and correlation statistics. The former section also provided an effective means for testing the data warehouse implementation, where no discrepancy was found for results of the same envisioned analysis but using MS Excel on the denormalized Deals data table.

The evaluation of determinants in Section 5.3.3 of Chapter 5 takes an initially naive approach by including a broad array of proxies in evaluating their influence on merger activity as opposed to common subset selection methods. However, by acknowledging association between dominant theories and respective determinants, the interpretation of results for empirical models could be alleviated. While support of the Neoclassical and Macroeconomic theories were found, other non significant variables should not be discredited for their contribution to model performance. Importantly, the determinant analyses executed demonstrate an improvement on univariate M&A activity prediction (ARIMA models), when using multivariate models. This provides evidence of determinant influence on the industry. Broadly, it can be said that determinants recognized in literature, do improve predictive capacities in explaining merger activity.

6.2 Attainment of Initial Aims and Objectives

Broadly, the aim of the thesis, set out in Section 1.3, was to develop a data management facility capable of assisting the analysis of the M&A industry, both in terms of its variables and activity drivers. Through the development and population of a data warehouse capable of connecting effectively to statistical analysis software, the necessary means for M&A activity and driver analysis for the deal set sample was created. Table 6.2 presents the specific method in facilitation, the contributing thesis component and a resulting product summary (output) for each respective thesis objective. The index in the left most column corresponds with each objective stated in Section 1.3.

Table 6.1: Objective facilitation and attainment.

	Facilitation	Contribution	Product
1	Systematic Quantitative Literature Review	Chapter 3, Section 3.2	Synthesized lists of activity and determinant variables
2	Data gathering protocol, informed by the study scope	Chapter 2, Section 2.2	M&A deals sample and determinant proxies
3	Data Warehouse Architecture and DBLC	Chapter 4	M&A deals and determinants data warehouse
4	Descriptive Statistics, Time Series Analysis, Graphical Representation	Chapter 5, Section 5.2	M&A activity trend analysis, seventh wave characterization
5	Time Series Analysis, LASSO, Multiple Regressions	Chapter 5, Section 5.3	Determinant evaluation, predictive model improvement

6.3 General Limitations of Study

Through the exclusion of recognized Firm-level financial attributes, the opportunity to facilitate a more comprehensive evaluation of determinant theory was compromised. As the discovery of errors in company financial information was late in the development and implementation of the data warehouse, it was decided that the structure and relevant data remained in place but would not be utilized in determinant analysis. As such, the existing database schema was maintained, under the assumption that company financial information reflected a financial statement around the time of the deal. Companies financial data was not included in the activity and determinant analysis of Sections 5.1 and 5.3.3 in Chapter 5, as insights from analyses would not fairly represent the deal set sample. If more reliable company financial information were to be obtained, a reiteration of the Database Life Cycle would be required and subsequent design adjustments made.

Another limitation identified was that the study primarily focused on quantitative variables and methods of analysis in the field. The bias in approach disregards the many qualitative influences such as culture, managerial sentiment and regulatory changes. However, these factors are often difficult to measure and subsequently test, limiting an overall inclusive approach. Quantitative indexes, as proxies for policy and economic uncertainty, better bridge this gap though.

6.4 Recommendations for Further Research

An increased number of determinants could be evaluated using the same or a similar approach to methods executed in Section 5.3.3 of Chapter 5. This exercise would be well supported by information from the synthesized list of

determinants, presented in Tables 3.4, 3.5 and 3.6 of Chapter 3. Specifically, if reliable Firm-level determinants could be obtained and included, a more comprehensive analysis of merger determinant theory could be executed. Improved robustness check processes should be implemented towards validating results and conclusions under different assumptions for potentially cumbersome models.

Additionally, a broader deal set sample could be introduced towards testing determinants on a more representative sample of the M&A industry. This could be done by considering a longer time frame for deals and determinants inclusion, as well as by including deals for unlisted companies. By allowing the scope of deals to include unlisted companies, determinants of the Behavioural theory of mergers may be better evaluated, given the significant role of overvaluation within the theory. However, Firm-level analyses prospects may need to be sacrificed, based on financial information availability for unlisted companies.

Finally, in a desegregation of merger activity, sector level trends for annual total value and volume could also be explained. This could be well supported by a similar approach to analyses as that of the one undertaken in this thesis.

List of References

Alexandridis, G., Mavrovitis, C.F. and Travlos, N.G. (2012). How have M&As changed? Evidence from the sixth merger wave. *European Journal of Finance*, vol. 18, no. 8, pp. 663–688. ISSN 1351847X.

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, vol. 23, no. 4, pp. 589–609.

Amel-Zadeh, A. and Zhang, Y. (2015). The economic consequences of financial restatements: Evidence from the market for corporate control. *Accounting Review*, vol. 90, no. 1, pp. 1–29. ISSN 00014826.

Behrens, J.T. (1997). Principles and Procedures of Exploratory Data Analysis. *Psychological Methods*, vol. 2, no. 2, pp. 131–160. ISSN 1082989X.

Benzing, C. (1991). The Determinants of Aggregate Merger Activity Before and After Celler-Kefauver. *Review of Industrial Organization*, vol. 6, no. 1, pp. 61–72. Available at: `https://www.jstor.org/st`

Bloomberg L.P. (2017 November). M& a defintions. Bloomberg Terminal. Bloomberg Terminal Access - Stellenbosch University.

Bloomberg L.P. (2019 July*a*). Mergers and acquisitions: Advanced search. Bloomberg Terminal. Bloomberg Terminal Access - The University of Stellenbosch.

Bloomberg L.P. (2019 July*b*). Price graph data table search function. Bloomberg Terminal. Bloomberg Terminal Access - The University of Stellenbosch.

Carpenter, M.A. and Sanders, W.G. (2006). Strategic management: a dynamic perspective, concepts and cases.

Cartwright, S., Teerikangas, S., Rouzies, A. and Wilson-Evered, E. (2012). Methods in M&A-A look at the past and the future to forge a path forward. *Scandinavian Journal of Management*, vol. 28, no. 2, pp. 95–106. ISSN 09565221.

Chatfield, C. (2004). The analysis of time series : an introduction. Formerly CIP.

Chaudhuri, S. and Dayal, U. (2015). An Overview of Data Warehousing and OLAP Technology. pp. 31–46.

Chidambaran, N.K., Krishnakumar, D. and Sethi, M. (2018). Cross-border vs. domestic acquisitions: Evidence from India. *Journal of Economics and Business*, vol. 95, pp. 3–25. ISSN 01486195.
Available at: `http://dx.doi.org/10.1016/j.jeconbus.2017.10.003`

Cortés, L.M. and Agudelo, D.A. (2017). Waves and Determinants in Mergers and Acquisitions : The Case of Latin America. pp. 1667–1690.

Coyle, B. (2000). *Mergers and Acquisitions*. Financial risk management: Corporate finance. Glenlake Publishing Company. ISBN 9781888998801.
Available at: `https://books.google.co.za/books?id=8-3MmA4KjBIC`

Digeorgio, R.M. (2003). Making mergers and acquisitions work: What we know and do not know - Part Two. vol. 3, pp. 259–274.

Digeorgio, R.M., Digeorgio, R.M. and Digeorgio, R.M. (2002). Practice papers Making mergers and acquisitions work: What we know and don't know - Part One. vol. 3, pp. 134–148.

Ellwanger, N. and Boschma, R. (2015). Who Acquires Whom? The Role of Geographical Proximity and Industrial Relatedness in Dutch Domestic M&As between 2002 and 2008. *Tijdschrift voor Economische en Sociale Geografie*, vol. 106, no. 5, pp. 608–624. ISSN 14679663.

Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, vol. 33, no. 1. ISSN 1548-7660.
Available at: `http://www.jstatsoft.org/v33/i01/`

Fuller, C.S. and Pusateri, N.R. (2018). A holistic approach to merger models with an emphasis on heterogeneity. *Quarterly Review of Economics and Finance*, vol. 69, pp. 260–273. ISSN 10629769.

Gatziu, S. and Vavouras, A. (1999). Data Warehousing: Concepts and Mechanisms. *Wirtschaftsinformatik als Mittler zwischen Technik, Ökonomie und Gesellschaft*, pp. 61–69.

Golubov, A. (2012). Empirical M & A research : A review of methods , evidence and managerial implications. , no. January, pp. 1–46.

Gomes, E. and Angwin, D.N. (2012). Critical success factors through the mergers and acquisitions process : Revealing pre- and post- M & A connections for improved performance. , no. December, pp. 1–48.

Gugler, K., Mueller, D.C. and Weichselbaumer, M. (2012). The determinants of merger waves: An international perspective. *International Journal of Industrial Organization*, vol. 30, no. 1, pp. 1–15. ISSN 01677187.
Available at: `http://dx.doi.org/10.1016/j.ijindorg.2011.04.006`

Harford, J. (2005). What drives merger waves? *Journal of Financial Economics*, vol. 77, no. 3, pp. 529–560. ISSN 0304405X.

Haspeslagh, P.C. and Jemison, D.B. (1991). *Managing acquisitions: Creating value through corporate renewal*, vol. 416. Free Press New York.

Hlavac, M. (2015). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Harvard University, Cambridge, USA. R package version 5.2.
Available at: `http://CRAN.R-project.org/package=stargazer`

Hoaglin, D.C. and Velleman, P.F.A. (1949). *Applications, basics and computing of exploratory data analysis*.

Hyndman, R. (2016). *Forecast: Forecasting functions for time series and linear models*. R package version 7.3.
Available at: `http://github.com/robjhyndman/forecast`

IMAA (2019 May). Mergers & acquisitions m & a - worldwide. Global Report for Deal Value and Deal Volume.
Available at: `https://imaa-institute.org`

Imdadullah, M. and Aslam, M. (2016). *mctest Multicollinearity Diagnostic Measures*. R package version 1.0.
Available at: `https://CRAN.R-project.org/package=mctest`

Inmon, B. and Kelly, C. (1994). The twelve rules of data warehouse for a client. *Server World, Data Management Review*, vol. 4, no. 5, pp. 6–16.

Kendall, K.E. and Kendall, J.E. (2011). *Systems analysis and design*. 8th edn. Pearson Prentice Hall, Upper Saddle River, N.J. ISBN 9780135094907.

Kitagawa, G. (2010). *Introduction to time series modeling*. Chapman & Hall, CRC press.

Koerner, J., Business, F. and Republic, C. (2014). The M & A Process Revisited: Identifying a Suitable Phase Model.

Komlenovic, S., Mamun, A. and Mishra, D. (2011). Business cycle and aggregate industry mergers. *Journal of Economics and Finance*, vol. 35, no. 3, pp. 239–259. ISSN 10550925.

Malik, M.F., Khan, A.W., Melati, A.A., Khan, S. and Khan, F. (2014). Mergers and acquisitions : A conceptual review. *International Journal of Accounting and Financial Reporting*, vol. 4, no. 2, pp. 520–533.

Masulis, R.W. and Simsir, S.A. (2013). Deal Initiation in Mergers and Acquisitions. *Ssrn*, vol. 53, no. 6, pp. 2389–2430.

Montgomery, D.C. and Runger, G.C. (2003). *Applied Statistics and Probability for Engineers*. John Wiley and Sons.

Moschieri, C. and Campa, J.M. (2018). The European M&A Industry: A Market in the Process of Construction. *Academy of Management Perspectives*, vol. 23, no. 4, pp. 71–87. ISSN 1558-9080.

Motis, J. (2007). Mergers and acquisitions Motives. *Hospitals & health networks / AHA*, vol. 75, no. 3, pp. 49, 51–54. ISSN 1068-8838.

Multiple.Com (2019 July). S&p 500 pe ratio by month. Electronic Data Set. Available at: `https://www.multpl.com/s-p-500-pe-ratio/table/by-month`

Pickering, C. and Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. vol. 4360.

Polasek, W. (2013). Time series analysis and its applications: With r examples, third edition by robert h. shumway, david s. stoffer. *International Statistical Review*, vol. 81, no. 2, pp. 323–325. ISSN 1751-5823.

Polemis, M.L. and Paleologos, J.M. (2014). Too big to merge? Evidence from the US banking sector. *Applied Economics Letters*, vol. 21, no. 11, pp. 782–785. ISSN 14664291.
Available at: `http://dx.doi.org/10.1080/13504851.2014.889797`

R Core Team (2016*a*). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Available at: `https://www.R-project.org/`

R Core Team (2016*b*). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Available at: `https://www.R-project.org/`

R Core Team (2016*c*). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Available at: `https://www.R-project.org/`

Ripley, B., and from 1999 to Oct 2002 Michael Lapsley (2010). *RODBC: ODBC Database Access*. R package version 1.3-2.
Available at: `http://CRAN.R-project.org/package=RODBC`

Rob, P., Coronel, C. and Crockett, K. (2008). *Database systems : design, implementation & management*. International ed. edn. Cengage Learning, London. ISBN 9781844807321.

Salus, N. (1989). Public relations before and after the merger. *Journal of*.

Schloerke, B., Crowley, J., Cook, D., Briatte, F., Marbach, M., Thoen, E., Elberg, A. and Larmarange, J. (2016). *GGally: Extension to 'ggplot2'*. R package version 1.2.0.
Available at: `https://CRAN.R-project.org/package=GGally`

Schmidt, S. (2013). Post-Merger Integration: Achieving Success in M&A. , no. May.

Takechi, K. (2011). R&D intensity and domestic and cross-border M&A of Japanese firms before domestic M&A deregulation. *Japan and the World Economy*, vol. 23, no. 2, pp. 112–118. ISSN 09221425.
Available at: `http://dx.doi.org/10.1016/j.japwor.2011.01.001`

Tang, J., Alelyani, S. and Liu, H. (2014). Feature selection for classification: A review. *Data classification: Algorithms and applications*, p. 37.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288. ISSN 0035-9246.

Wong, T.T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, vol. 48, no. 9, pp. 2839–2846. ISSN 00313203.
Available at: `http://dx.doi.org/10.1016/j.patcog.2015.03.009`

Yaghoubi, R., Yaghoubi, M., Locke, S. and Gibb, J. (2016). Mergers and acquisitions: a review part 1. *Studies in Economics and Finance*, vol. 33, no. 1, pp. 147–188.

Appendices

Appendix A

Twelve Rules That Define a Data Warehouse

The twelve rules that define a data warehouse as developed and presented by Inmon and Kelly (1994), referred to in literature coverage, Section 3.3.1.

1. "The data warehouse and operational environments are separated.
2. The data warehouse data are integrated.
3. The data warehouse contains historical data over a long time.
4. The data warehouse data are snapshot data captured at a given point in time.
5. The data warehouse data are subject oriented.
6. The data warehouse data are mainly read-only with periodic batch updates from operational data. No online updates are allowed.
7. The data warehouse development life cycle differs from classical systems development. The data warehouse development is data-driven; the classical approach is process-driven.
8. The data warehouse contains data with several levels of detail: current detail data, old detail data, lightly summarized data, and highly summarized data.
9. The data warehouse environment is characterized by read-only transactions to very large data sets. The operational environment is characterized by numerous update transactions to a few data entities at a time.
10. The data warehouse environment has a system that traces data sources, transformations, and storage.

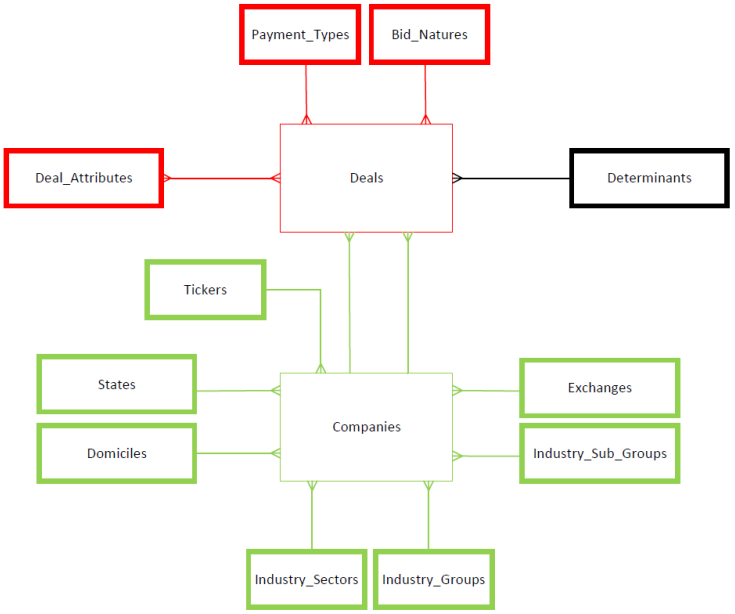
11. The data warehouse's metadata are a critical component of this environment. The metadata identify and define all data elements. The metadata provide the source, transformation, integration, storage, usage, relationships, and history of each data element.

12. The data warehouse contains a charge-back mechanism for resource usage that enforces optimal use of the data by end users."

Appendix B

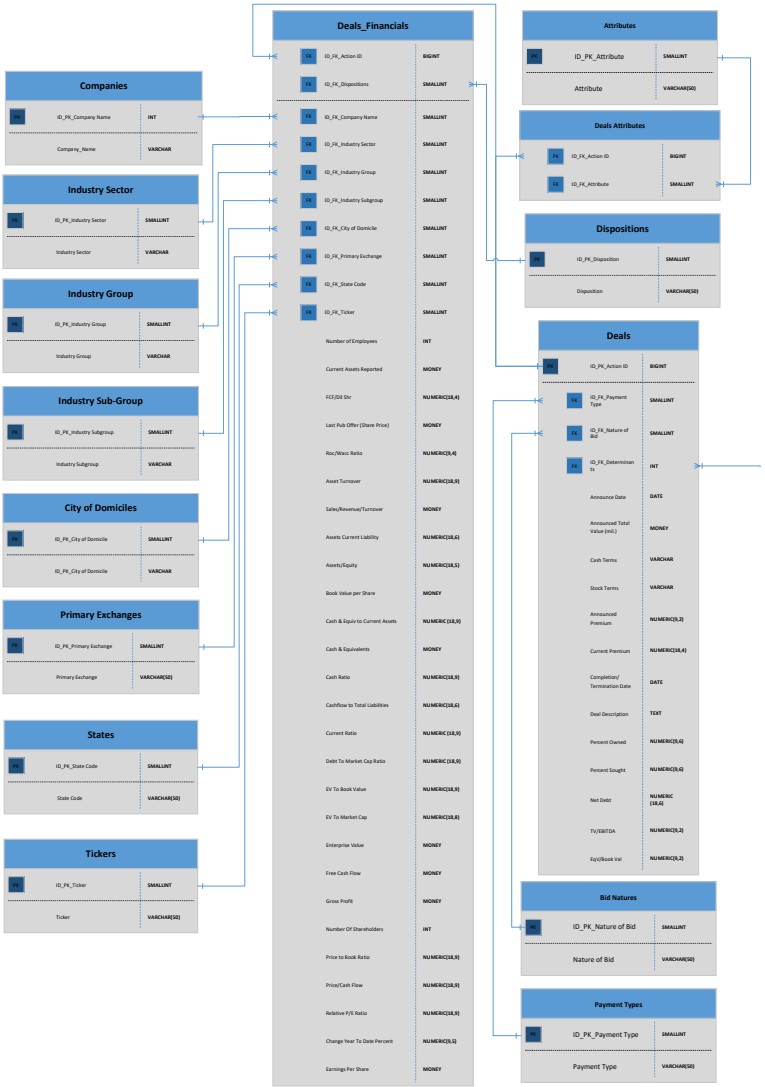
Initial ERD Design

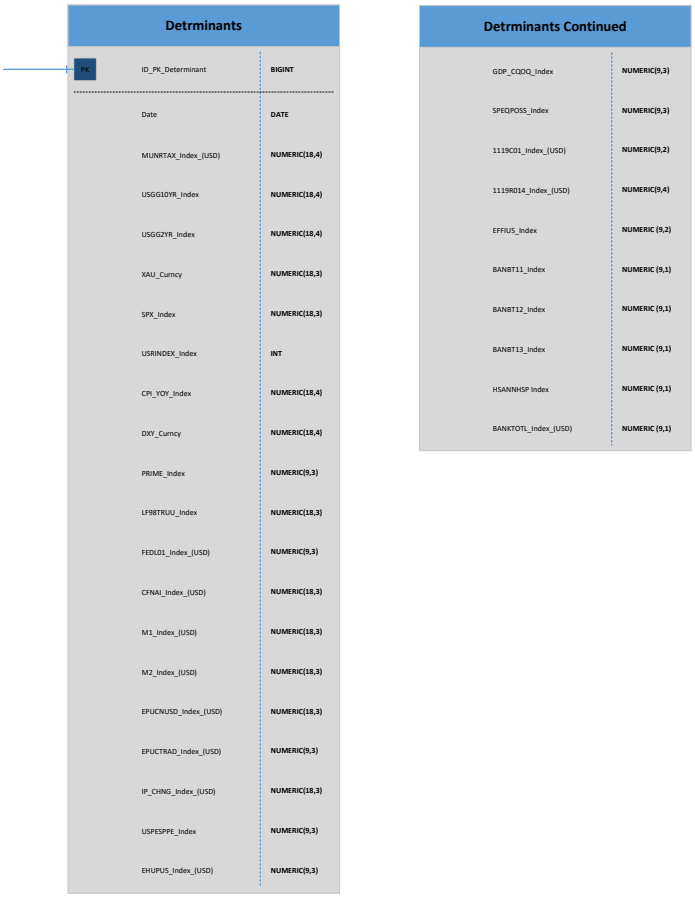
Primary entities included, Deal Attributes, Payment Types, Bid Natures and Determinants for the Deals attributive entity. The Companies attributive entity was characterized by the primary entities of States, Domiciles, Industry Sectors, Groups and Sub-groups as well as Exchanges and Tickers. Each primary entity was to be of a one-to-many relation, where each entity could relate to many records of the attributive entities and each attributive entity was to have at most one corresponding primary entity record, excepting Deal Attributes, which was to be governed by a many-to-many relationship. The repeated one-to-many relationship between Deals and Companies was used to distinguish for acquirer and target companies involved in each deal. The design was found to assume incorrect transitive dependencies for all green primary entities on the Companies attributive entity which were actually dependent on Deals primary key.



Appendix C

Database Schema





Appendix D

Determinants Summary Statistics Comparison

	R Summary						Excel Descriptive		Delta	
Statistic	***N***	***Min***	***Median***	***Max***	***Mean***	***St.Dev.***	***Mean***	***St.Dev.***	***Mean***	***St.Dev***
MUNRTAX	259	-14.78	35.454	97.541	39.563	22.224	39.563	22.224	0	0
USGG10YR	259	1.453	3.593	6.665	3.611	1.315	3.611	1.315	0	0
USGG2YR	259	0.2	1.872	6.676	2.391	1.884	2.391	1.884	0	0
XAU	259	255.68	916.9	1,825.55	866.428	477.473	866.428	477.473	0	0
SPX	259	735.09	1,325.83	2,980.38	1,515.07	539.55	1515.069	539.55	0	0
USRINDEX	259	0	0	1	0.108	0.311	0.108	0.311	0	0
CPI	258	-2.097	2.082	5.6	2.151	1.212	2.151	1.212	0	0
DXY	259	71.802	90.07	120.21	91.321	11.601	91.321	11.601	0	0
PRIME	259	3.25	4.25	9.5	5.155	2.044	5.155	2.044	0	0
LF98TRUU	259	495.43	906.34	2,110.98	1,093.39	504.851	1093.388	504.851	0	0
FEDL01	259	0.07	1.26	6.54	2.096	2.092	2.096	2.092	0	0
CFNAI	258	-4.75	-0.08	1.59	-0.206	0.828	-0.206	0.828	0	0
M1	258	1,074.00	1,465.25	3,824.00	1,963.30	887.883	1963.302	887.883	0	0
M2	258	4,041.20	7,875.45	14,771.20	8,506.11	3,145.25	8506.111	3145.248	0	0
EPUCNUSD	259	18.15	85.65	626.03	106.766	73.177	106.766	73.177	0	0
EPUCTRAD	258	5.42	50.07	635.38	74.703	82.588	74.703	82.588	0	0
IP	258	-4.34	0.14	2.05	0.095	0.656	0.095	0.656	0	0
USPESPPE	259	13.5	22.18	123.73	26.026	16.222	26.026	16.222	0	0
EHUPUS	86	3.63	5.2	9.93	5.792	1.774	5.792	1.774	0	0
GDP	86	-8.4	2.4	7.5	2.308	2.394	2.308	2.394	0	0
SPEQPOSS	85	48.9	67.2	81.4	66.195	6.69	66.195	6.69	0	0
X1119C01	77	76.6	98.7	114.1	96.905	9.891	96.905	9.891	0	0
X1119R014	77	-4.062	2.3	5.266	2.208	1.818	2.208	1.818	0	0
BANBT11	86	564	1,569.50	4,455	1,675.14	797.387	1675.14	797.387	0	0
BANBT12	82	1	86.5	279	87.476	56.775	87.476	56.775	0	0
BANBT13	86	379	695	1,358	740.86	278.496	740.86	278.496	0	0
BANKTOTL	78	116,771	299,771	667,431	301,813.00	101,188	301813	101188.2000	0	0
EFFIUS	21	65.1	69.8	79.4	71.9	5.347	71.9	5.347	0	0
HSANNHSP	21	554	1,249.90	2,068.30	1,298.30	467.91	1298.3	467.91	0	0

Appendix E

R Summary Statistics for Numeric Deals Table Variables

Statistic	N	Min	Mean	Median	Max	St. Dev.
Ann Tot Value mil	4,828	1.000	1,739.417	190.040	186,235.800	6,673.655
Announced Premium	4,093	−100.000	38.819	30.590	903.570	45.176
Current Premium	3,251	−100.000	60.013	0.280	103,650.000	2,313.082
Deal Length	4,828	−3	135.871	120	2,354	97.182
Percent Owned	4,828	0.000	2.923	0.000	98.900	13.737
Percent Sought	4,828	0.000	95.713	100.000	100.000	16.940
Net Debt	4,828	−6,446.576	286.974	0.000	79,344.000	1,890.322
TV/EBITDA	2,207	−5.290	30.473	12.010	3,392.950	141.659
EqV/Book.Val	3,705	−6.810	5.726	2.440	1,347.200	31.384

Appendix F

Top 30 Serial Acquirers for Deal Set Sample

Acquirer	**Number of Deals**	**Average Deal Value**	**Average Announced Premium(%)**	**Average Deal Length**
General Electric Co	28	1155.30	40.23	93
BB&T Corp	27	598.54	34.31	179
Oracle Corp	24	2163.39	35.55	106
Berkshire Hathaway Inc	20	6484.80	26.88	116
Danaher Corp	19	1727.12	48.03	59
Johnson & Johnson	18	1416.04	40.39	77
International Business Machines Corp	18	693.22	41.74	62
PacWest Bancorp	18	272.73	34.47	163
FNB Corp/PA	16	221.93	54.60	181
Thermo Fisher Scientific Inc	15	2783.13	23.24	220
Pfizer Inc	15	17191.71	70.92	122
Wells Fargo & Co	15	3566.29	24.84	141
Apollo Global Management LLC	15	2085.48	27.39	94
AT&T Inc	15	24877.02	69.82	242
Cisco Systems Inc	14	2251.21	31.25	81
Blackstone Group Inc/The	14	4481.46	20.11	100
HP Inc	14	3958.80	46.71	90
Medtronic PLC	13	4787.36	54.71	111
Sprint Communications Inc	13	4974.82	35.24	121
Liberty Interactive LLC	12	966.95	43.73	173
Allergan PLC	12	8522.16	85.37	97
Banknorth Group Inc	12	272.51	37.93	172
Fifth Third Bancorp	12	1189.01	59.41	221
Avis Budget Group Inc	11	797.36	55.57	83
IAC/InterActiveCorp	11	805.90	22.90	133
Apartment Investment & Management Co	11	107.12	NULL	64
UnitedHealth Group Inc	10	3707.98	17.31	130
EMC Corp	10	1064.62	38.05	65
Wachovia Corp	10	6617.24	19.53	124
Merck & Co Inc	9	6796.17	113.28	78
Average	**15.03**	**3884.58**	**43.22**	**124**
Deal Set Average	**1.92**	**1739.42**	**38.82**	**135**

Appendix G

Pearson Correlation Matrix for Determinant Proxies