University of Surrey

School of Management

The Measurement of Financial Performance and its Effect on the Relationship between Gearing and Over-investment

by

Hassan Hamadi

Thesis submitted in part fulfillment of the requirements for the award of Doctor of Philosophy

2005

©
Abstract

This thesis investigates the relationship between financial performance, over-investment and gearing for UK and US firms. In this thesis, we test the proposition that debt disciplines managers and mitigates the overinvestment problem and therefore, this should improve firm performance. A sample of panel data from FTSE 500 and S&P 500 companies covering a 15 year period (1988 - 2002) was downloaded from Datastream. Regression tests were conducted using different estimation models such as fixed effects and logit model estimation to test the hypotheses of this thesis.

In the first stage, we find that the Residual Income model (RI) explains more than 75% of the cross sectional variation in stock prices for both UK and US firms. Secondly we used this model to determine performance and we find that the RI component derived from the RI model outperforms both the first difference of the RI model and the simple RI model as a measure of financial performance.

In the second stage, we find that the higher the gearing the lower the probability of over-investment. Moreover, we find that the relationship between gearing and over-investment is non-linear. At low levels of gearing, as the gearing increases so does the over-investment problem, while at high levels of gearing, the higher the gearing the lower the probability of over-investment. In line with Jensen’s free cash flow theory, we find evidence of over-investment and our results suggest that the over-investment problem can be mitigated using mid-levels of debt. Furthermore, we find that the relationship between over-investment and firm performance is significantly negative. Finally, where gearing mitigates the overinvestment, the interaction of gearing and overinvestment has a positive influence on firm performance.

This study introduces a new proxy to measure the over-investment problem, it uses a performance measure based on residual income model and tests the non-linearity between gearing, over-investment and performance using quadratic and cubic equations.
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Trade-off in balance sheet format</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Description of FTSE 500 and S&amp;P 500</td>
<td>102</td>
</tr>
<tr>
<td>5.1</td>
<td>Definition of the variables as in data-stream</td>
<td>113</td>
</tr>
<tr>
<td>5.2</td>
<td>Criteria used in analysing panel data</td>
<td>115</td>
</tr>
<tr>
<td>5.3</td>
<td>UK and US descriptive statistics for ‘P’ &amp; ‘V’</td>
<td>116</td>
</tr>
<tr>
<td>5.4</td>
<td>Correlation between ‘P’ and ‘V’</td>
<td>116</td>
</tr>
<tr>
<td>5.5</td>
<td>UK and US results from equation (1.1)</td>
<td>116</td>
</tr>
<tr>
<td>5.6</td>
<td>Cross-sectional results from equation (1.1)</td>
<td>118</td>
</tr>
<tr>
<td>5.7</td>
<td>Cross-sectional results (R^2) for UK portfolios</td>
<td>119</td>
</tr>
<tr>
<td>5.8</td>
<td>Cross-sectional results (R^2) for US portfolios</td>
<td>120</td>
</tr>
<tr>
<td>5.9</td>
<td>Descriptive statistics for performance models</td>
<td>124</td>
</tr>
<tr>
<td>5.10</td>
<td>Results from equation (2.1) first difference</td>
<td>125</td>
</tr>
<tr>
<td>5.11</td>
<td>Results from equation (3.1) the RIC model</td>
<td>125</td>
</tr>
<tr>
<td>5.12</td>
<td>Comparison between the two performance models</td>
<td>126</td>
</tr>
<tr>
<td>6.1</td>
<td>Definition of the variables as in data-stream</td>
<td>138</td>
</tr>
<tr>
<td>6.2</td>
<td>UK and US descriptive statistics</td>
<td>141</td>
</tr>
<tr>
<td>6.3</td>
<td>Correlation between over-investment &amp; gearing</td>
<td>141</td>
</tr>
<tr>
<td>6.4</td>
<td>Correlation between performance &amp; interaction variables</td>
<td>142</td>
</tr>
<tr>
<td>6.5</td>
<td>Modelling D by logit: results from UK &amp; US data</td>
<td>144</td>
</tr>
<tr>
<td>6.6</td>
<td>Modelling D by logit: industry-adjusted</td>
<td>146</td>
</tr>
<tr>
<td>6.7</td>
<td>Modelling D by logit: results from UK data</td>
<td>147</td>
</tr>
<tr>
<td>6.8</td>
<td>Modelling D by logit: results from US data</td>
<td>148</td>
</tr>
</tbody>
</table>
6.9 Modelling D by logit: UK top and mid-percentile 149
6.10 Modelling D by logit: The cubed results 151
6.11 Modelling D by logit: The squared results 152
6.12 Modelling performance and over-investment 156
6.13 Modelling performance & over-investment (UK+US) 157
6.14 Modelling performance & the interaction variable:
The pooled UK and US results 159
7.1 Summery of the tested hypotheses 171
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>MM proposition II</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Graphical illustration of MM's 1958 theory</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>MM proposition II (with corporate taxes)</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Cost of capital &amp; the optimal capital structure</td>
<td>19</td>
</tr>
<tr>
<td>2.5</td>
<td>Effect of leverage on the value of firm's stock</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>Total agency cost &amp; the optimal capital structure</td>
<td>28</td>
</tr>
<tr>
<td>2.7</td>
<td>Agency cost &amp; asymmetric information</td>
<td>43</td>
</tr>
</tbody>
</table>
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>BG</td>
<td>Book Value of Gearing</td>
</tr>
<tr>
<td>BV</td>
<td>Book Value of equity</td>
</tr>
<tr>
<td>CAPM</td>
<td>Capital Assets Pricing Model</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Density Function</td>
</tr>
<tr>
<td>CSR</td>
<td>Clean Surplus Relation</td>
</tr>
<tr>
<td>CV</td>
<td>Continuing Value</td>
</tr>
<tr>
<td>DCF</td>
<td>Discounted Cash Flow</td>
</tr>
<tr>
<td>DDM</td>
<td>Dividend Discount Model</td>
</tr>
<tr>
<td>EBO</td>
<td>Edwards-Bell-Ohlson</td>
</tr>
<tr>
<td>EP</td>
<td>Economic Profit</td>
</tr>
<tr>
<td>EPS</td>
<td>Earnings per Share</td>
</tr>
<tr>
<td>EVA</td>
<td>Economic Value Added</td>
</tr>
<tr>
<td>FASB</td>
<td>Financial Accounting Standards Board</td>
</tr>
<tr>
<td>F&amp;F</td>
<td>Fama and French</td>
</tr>
<tr>
<td>F&amp;L</td>
<td>Frankel and Lee</td>
</tr>
<tr>
<td>FCF</td>
<td>Free Cash Flow</td>
</tr>
<tr>
<td>FROE</td>
<td>Future Return on Equity Forecasts</td>
</tr>
<tr>
<td>GAAP</td>
<td>Generally Acceptable Accounting Principles</td>
</tr>
<tr>
<td>IBCLs</td>
<td>Interest-Bearing Current Liabilities</td>
</tr>
<tr>
<td>LIFO</td>
<td>Last in First Out</td>
</tr>
<tr>
<td>MM</td>
<td>Modigliani and Miller</td>
</tr>
<tr>
<td>MNCs</td>
<td>Multinational Corporations</td>
</tr>
<tr>
<td>MG</td>
<td>Market Value of Gearing</td>
</tr>
<tr>
<td>MV</td>
<td>Market Value of Equity</td>
</tr>
<tr>
<td>MVA</td>
<td>Market Value of the firm’s equity plus the book value of the firm’s total debt</td>
</tr>
<tr>
<td>NIBCLs</td>
<td>Non-Interest-Bearing Current Liabilities</td>
</tr>
<tr>
<td>NOPAT</td>
<td>Net Operating Profit after Tax</td>
</tr>
<tr>
<td>NOPLAT</td>
<td>Net Operating Profit less Adjusted Taxes</td>
</tr>
<tr>
<td>NPV</td>
<td>Net Present Value</td>
</tr>
</tbody>
</table>
OLS  Ordinary Least Square
PRF  Population Regression Function
R&D  Research and Development
RI   Residual Income
RIC  Residual Income Components
REVA Refinement of Economic Value Added
ROA  Return on Assets
ROCE Return on Capital Employed
ROE  Return on Equity
ROIC Return on Invested Capital
SIC  Standard Industrial Classification
SRF  Sample Regression Function
TV   Terminal Value
WACC Weighted Average Cost of Capital
Table of Contents

Abstract........................................................................................................................... I

List of Tables ................................................................................................................ II

List of Figures .............................................................................................................. IV

List of Abbreviations .................................................................................................... V

Table of Contents ........................................................................................................ VII

Acknowledgments ........................................................................................................ X

Chapter 1 ........................................................................................................................ 1
  Introduction ................................................................................................................... I
  1.1 Introduction ....................................................................................................... 1
  1.2 The Aim of this Study ....................................................................................... 3
  1.3 Outline of the thesis ........................................................................................... 5

Chapter 2 ........................................................................................................................ 7
  Capital Structure ......................................................................................................... 7
  2.1 Introduction ......................................................................................................... 7
  2.2 Modigliani and Miller Theory ........................................................................... 9
    2.2.1 Introduction .................................................................................................... 9
    2.2.2 MM Proposition I ....................................................................................... 9
    2.2.3 MM Proposition II ..................................................................................... 12
    2.2.4 MM with Corporate Taxes ........................................................................ 14
    2.2.5 Implications of the MM Theory .................................................................... 16
  2.3 The Miller Model ............................................................................................. 17
  2.4 Cost of Financial Distress ................................................................................ 18
  2.5 The Trade-Off Theory ..................................................................................... 20
  2.6 Agency Costs ................................................................................................... 23
    2.6.1 Introduction ............................................................................................. 23
    2.6.2 Agency Cost of External Equity .............................................................. 24
    2.6.3 Agency Cost of Debt ............................................................................... 25
  2.7 Signalling and Asymmetric Information Theory ............................................. 29
  2.8 The Pecking Order Theory .............................................................................. 31
  2.9 Free Cash Flow Theory: Jensen (1986) ........................................................... 34
  2.10 Empirical Evidence on Capital Structure ...................................................... 35
  2.11 Conclusion ....................................................................................................... 41

Chapter 3 ...................................................................................................................... 46
  Financial Performance Measures ................................................................................. 46
  3.1 Introduction ......................................................................................................... 46
  3.2 Financial Ratios ................................................................................................. 47
  3.3 Discounted Cash Flow Model (DCF): The Free Cash Flow Model .......... 48
  3.4 Discounted Cash Flow: The Economic Profit Model ........................................ 53
### Chapter Three

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Residual Income: The Economic Value Added Model</td>
<td>55</td>
</tr>
<tr>
<td>3.8</td>
<td>Comparison between the Models</td>
<td>69</td>
</tr>
<tr>
<td>3.8.1</td>
<td>Theoretical Equivalence between EVA and RI in Valuation</td>
<td>69</td>
</tr>
<tr>
<td>3.8.2</td>
<td>Theoretical Equivalence between the RI and DCF Models</td>
<td>71</td>
</tr>
<tr>
<td>3.8.3</td>
<td>Differences between the RI and DCF Models</td>
<td>73</td>
</tr>
<tr>
<td>3.9</td>
<td>Empirical Research</td>
<td>77</td>
</tr>
<tr>
<td>3.10</td>
<td>Conclusion</td>
<td>83</td>
</tr>
</tbody>
</table>

### Chapter Four

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>86</td>
</tr>
<tr>
<td>4.2</td>
<td>Research Objectives</td>
<td>86</td>
</tr>
<tr>
<td>4.3</td>
<td>The Hypotheses</td>
<td>87</td>
</tr>
<tr>
<td>4.4</td>
<td>Research Methods</td>
<td>90</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Regression Analysis</td>
<td>90</td>
</tr>
<tr>
<td>4.4.1.1</td>
<td>Sample Regression Function</td>
<td>92</td>
</tr>
<tr>
<td>4.4.1.2</td>
<td>Quality of the Sample Regression Function</td>
<td>92</td>
</tr>
<tr>
<td>4.4.1.3</td>
<td>Fixed and Random Effects Models</td>
<td>93</td>
</tr>
<tr>
<td>4.4.1.4</td>
<td>Other Issues to be Considered</td>
<td>95</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Regression on Dummy Dependent Variable</td>
<td>95</td>
</tr>
<tr>
<td>4.4.2.1</td>
<td>Binary Logit/Probit Model</td>
<td>97</td>
</tr>
<tr>
<td>4.4.2.2</td>
<td>Should the Logit or the Probit Model be Chosen?</td>
<td>98</td>
</tr>
<tr>
<td>4.5</td>
<td>Panel Data</td>
<td>99</td>
</tr>
<tr>
<td>4.6</td>
<td>Data and Sample Description</td>
<td>101</td>
</tr>
<tr>
<td>4.7</td>
<td>Statistical Packages</td>
<td>104</td>
</tr>
<tr>
<td>4.8</td>
<td>Conclusions</td>
<td>105</td>
</tr>
</tbody>
</table>

### Chapter Five

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Testing the Residual Income Model</td>
<td>106</td>
</tr>
<tr>
<td>5.2</td>
<td>Definition of the Variables</td>
<td>106</td>
</tr>
<tr>
<td>5.3</td>
<td>The Hypothesis and the Regression</td>
<td>110</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Panel Data Analysis</td>
<td>114</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Pooling the Data Cross-Sectionally</td>
<td>117</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Conclusions from Both Panel and Cross-Sectional Data</td>
<td>121</td>
</tr>
<tr>
<td>5.5</td>
<td>Determination of Performance</td>
<td>121</td>
</tr>
<tr>
<td>5.5.1</td>
<td>The Hypotheses and the Regressions</td>
<td>123</td>
</tr>
<tr>
<td>5.5.2</td>
<td>Results</td>
<td>124</td>
</tr>
<tr>
<td>5.5.3</td>
<td>Comparison between the two Performance Models</td>
<td>126</td>
</tr>
<tr>
<td>5.6</td>
<td>Conclusions</td>
<td>127</td>
</tr>
</tbody>
</table>

### Chapter Six

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>128</td>
</tr>
<tr>
<td>6.2</td>
<td>The Relationship between Firm Performance, Debt and Over-investment</td>
<td>129</td>
</tr>
<tr>
<td>6.3</td>
<td>Developing the Model</td>
<td>132</td>
</tr>
<tr>
<td>6.4</td>
<td>Definition of the Variables</td>
<td>134</td>
</tr>
<tr>
<td>6.5</td>
<td>The Hypotheses</td>
<td>138</td>
</tr>
</tbody>
</table>
6.6 Analysis and Results from the Regression Analysis ..................................... 140
6.6.1 Over-investment and Gearing Assuming Linearity .................................. 144
6.6.2 Over-investment and Gearing Assuming Linearity (Industry Effect) ... 145
6.6.3 Over-investment and Gearing Assuming Linearity (Gearing Classification) ........................................................................................................ 146
6.6.4 Over-investment and Gearing Assuming Non-Linear Variables .......... 150
6.7 Regression Analysis: Over-investment, Gearing and Firm Performance Assuming Non-Linearity ........................................................................................... 154
6.7.1 Over-investment and Firm Performance ............................................... 155
6.7.2 Firm Performance, Over-investment and Gearing Assuming Non-Linearity: 157
6.8 Conclusions ................................................................................................... 160

Chapter 7 .................................................................................................................... 163
Summary and Conclusions .................................................................................... 163
7.1 Introduction .................................................................................................... 163
7.2 Summary of Chapters .................................................................................. 163
7.3 Contributions and Managerial Implications of the Study ......................... 171
7.4 Limitations of this Study ............................................................................. 173
7.5 Suggestions for Further Research ................................................................ 173
Bibliography .............................................................................................................. 175
Appendices ................................................................................................................ 191
Appendix 3.A: Derivation of the Continuing Value Formula .............................. 191
Appendix 3.B: Balance Sheet before and after Adjustments ................................ 193
Appendix 3.C: Development of REVA ................................................................. 194
Appendix 5.A: Non-Parametric Test- Spearman’s Correlation......................... 198
Appendix 5.B: The Transformed Variables ............................................................ 199
Appendix 5.C: Price / Book Value in Comparison with Price / Residual Income Model .............................................................................................................. 200
Appendix 5.D: The Simple Residual Income Model in Comparison with First Difference and Residual Income Component .......................................................... 201
Appendix 6.A: Over-investment and Debt-Equity Ratio Assuming Linearity. ...... 202
Appendix 6.C: Over-investment and Debt-Equity Ratio Classification Assuming Linearity ............................................................................................................... 204
Appendix 6.D: Over-investment and Debt-Equity Ratio Assuming Non-Linearity. 206
Appendix 6.E Over-investment and Gearing Assuming Non-Linear Variables ...... 207
Acknowledgements

I wish to thank my supervisors Professor Simon Archer and Dr. Julinda Nuri very much for their invaluable continuous support, guidance, inspiration and patience. I am grateful for all the meetings I had with them during the last years. These meetings never failed to be constructive, encouraging and stimulating. I could not have produced this thesis without their supervision. (“If I have seen further, it is by standing on the shoulders of a giant” – Sir. Isaac Newton; 1642-1727).

Dr Yimaz Guney provided invaluable help on models of panel data. I thank him for his generous help and successful discussions. I also wish to thank Dr. Paul Masters for his academic support, friendship, understanding and patience with my continuous ‘nagging’. I thank my colleagues in the department for all the time, feedback and discussions. I also would like to thank Dr. Samer Kharroubi for his kind support through out my thesis.

I also wish to thank my best friends Mr. Ausama Salman, Mrs Murooj Salman, their daughter Nabaa, and Mr fadhel Houssein for the precious moments and moral support that they have lovingly gave me when I needed them most. I will not forget you. I also wish to thank my good friend Ali Jammoul for his generous moral and emotional support.

As always, though words do not suffice endless love and gratitude to my family and especially to my parents and my sister Fatima without whom this work would not have been possible; their love and support will always be appreciated and never forgotten. Hugs and kisses to my lovely son Ali for his charming smile. And last but not least, I would like to thank my beloved wife ‘Mirvat Hamadi’ for always being there when I needed her and for her relentless efforts and support over the frustrating moments of this PhD. Thank you Mirvat for your comforting words, lute playing, and all other means you have used to make my arduous work bearable. I hope I can repay you one day for all what you have given me and more.
Chapter 1

Introduction

1.1 Introduction

A central theme in corporate finance relates to the choice that firms make between issuing equity or debt. The theoretical literature on this topic includes both static models that date back to Modigliani and Miller (1958), and dynamic models. Under perfect capital market conditions, Modigliani and Miller (1963) showed that, as the result of gains from leverage, the value of a levered firm is higher in the presence of corporate taxes than that of a non-levered firm in the same risk class. However, in practice, firms are not all-debt financed despite the tax shields benefits, and significant differences in the cross-sectional and time-series variations in their capital structure are observed. It appears that when making their financial decisions, firms consider the costs and benefits associated with each financing method (Bradley et al, 1984; Titman and Wessels, 1988).

As a result, the trade-off theory has emerged. This theory states that the optimal level of leverage would emerge from a trade-off between the tax deductibility of interest and the expected bankruptcy costs (Brennan and Schwartz, 1978; Chen, 1978), firms’ taxable capacity (DeAngelo and Masulis, 1980), and the agency cost of debt (Jensen and Meckling, 1976; Jensen, 1986). More recently, agency theory variables (as explained below) are regarded as the main determinants of capital structure (Barclay et al, 1995; Harris and Raviv, 1990; Stulz, 1990). Since the effects of taxes and direct bankruptcy costs, though well developed theoretically, appear to play a minor role empirically, the major determinants of capital structure appear to relate to agency theory (Titman and Wessels, 1988; Barclay and Smith, 1988).
Therefore, the empirical evidence has shown that the choices that firms make in issuing equity or debt are interdependent. Accordingly, some theories that explain this have been developed. These theories are based on capital-market imperfections; with respect to investment decisions, the existence of asymmetric information between the main stakeholders constitutes the foremost imperfection. Agency problems play an important role in this thesis. Information differences and divergences of interest lead to agency problems, and to costs associated with these problems.

The role of asymmetric information in investment decisions has as its primary basis the theoretical works of Jensen and Meckling (1976), Myers (1977) and Myers and Majluf (1984). The first two papers emphasise the consequences of the existence of post-contract asymmetric information between shareholders and bondholders, while the paper of Myers and Majluf (1984) emphasise the role of the pre-contract asymmetric information between current and prospective shareholders. All the above-mentioned papers show that information asymmetries may lead to some investment projects with a positive net present value (NPV) not being undertaken.

A second foundation in the study of inefficient investment decisions is the work of Jensen (1986). His work, starting from the hypothesis of the existence of asymmetric information between managers and shareholders, introduces the so-called problem of over-investment, as a basic argument of his free cash flow theory. According to this theory there can be negative NPV investment projects that end up being completed because managers derive benefits from investment (the larger the organisation, the greater the economic and political power of top management team). Managers are empire builders and continue to choose investment projects even after all positive NPV projects have been taken. Accordingly, the purpose of this thesis is to test whether a certain level of debt can, to a certain degree, mitigate the over-investment problem and therefore, lead to better firm performance.

While agency theory has generated insights into capital structure, measuring its effect on companies' financial performance represents a real challenge. Researchers usually use one of the two ways, presented below, to measure financial performance:

1- Market returns i.e. stock price return.
2- Commonly used financial ratio measures of returns, such as Return on Assets (ROA) or Return on Capital Employed (ROCE) as proxies of financial performance.

Financial ratio measures of returns are inadequate instruments because they ignore the cost of capital, motivating dysfunctional behaviour causing managers to pay attention to the "wrong" things (Aggarwal 2001). On the other hand, market returns models (stock price or returns based on stock price) impound market expectations about a firm's future and thus may not be able to capture managers' current performance. It also may not be an efficient contracting parameter because it is driven by many factors beyond the control of the firms' executives (Bacidore et al 1997). In this thesis, a residual income-based model, namely the Frankel and Lee (1998) model, with a solid theoretical underpinning will be used to determine firm performance (for reasons explained in Chapter 5) instead of returns models. However, the Frankel and Lee (1998) model is a valuation model, but the first difference can be used as a performance measure with correction for dividend. A comparison between various Residual Income models (RI), i.e. Economic Value Added (EVA), Economic Profit (EP), the Ohlson (1995) model and the Frankel and Lee (1998) model on the one hand, and the Discounted Cash Flow (DCF) model on the other hand will be carried out in the literature review section. The RI and the DCF models are based on theoretical arguments and take into consideration the cost of capital when calculating the value of a company. Although these models are preferable to ROA, they measure value rather than performance; but by correcting for dividend, the first difference can measure performance.

1.2 The Aim of this Study

The aim of this study is to find a relationship between capital structure and companies' performance for a sample of listed non-financial firms from FTSE 500 and S&P 500. Agency theory variables appear to play a major role in inducing

---

managers to perform better and it is hypothesised that capital structure affects company performance. Furthermore, according to Jensen (1986), debt can discipline managers and mitigate the problem of over-investment because debt provides a means of bonding managers’ promises to pay out free cash flows rather than investing in wealth-destroying ventures. Debt also provides means for controlling opportunistic behaviour by reducing cash flows available for discretionary spending. Therefore, managers’ attention is then clearly focused on those activities necessary to ensure that debt payment is met. If this is true, then in this situation debt should induce better firm performance. In this thesis, we will adapt Jensen’s (1986) argument and empirically test the proposition that debt disciplines managers and mitigates the over-investment problem and ultimately leads to better firm performance. The measurement of financial performance represents a real challenge for researchers and academics. As mentioned above, a performance measure based on the Frankel and Lee (1998) model will be used to measure companies’ performance. Previous studies regarding the effect of capital structure on companies’ performance have focused on the earnings figure. However, less attention has been paid to the informative value of the balance sheet items. Alternatively, researchers have used the market returns to examine the effects of capital structure on performance.

There are several important features of our analysis, which we believe extend the literature on the empirical aspects of agency problems; more precisely, Jensen’s free cash flow argument. Firstly, whereas most studies investigate US firms our sample comprises both US and UK firms. Secondly, we will use a performance measure with a solid theoretical underpinning and empirically test and account for the cost of capital. Thirdly, we will show the possibilities of using the Residual Income Components (RIC) as a performance measure. The RIC, derived from the Frankel and Lee (1998) model, represents the residual income components of that model and accounts for the cost of capital. Fourthly, our analysis, distinct from previous empirical studies, introduces a simple way of measuring the over-investment problem through a combination of capital expenditure and growth opportunities. Fifthly, we will test for the non-linearity of the relationship between gearing, performance and over-investment. Finally, our analysis incorporates the dynamic nature of response of firm’s performance to investment decisions and controls for the endogeneity.
1.3 Outline of the thesis

This thesis consists of a review of capital structure literature, review of valuation and performance models literature, two main empirical studies, and the conclusions.

Chapter 1 states the aims and objectives of this study as well as describes the structure of this study.

Chapter 2 discusses the theoretical and empirical literature on capital structure. Starting from the seminal paper by Modigliani and Miller (1958), we describe their theoretical and empirical contributions. The description of their work provides the background for the subsequent chapters, with respect to the theory that is tested and the results of the existing empirical studies. The chapter concludes with a critical assessment of empirical capital structure research and discusses alternative directions in empirical research.

Chapter 3 discusses performance models. It begins with highlighting the main weaknesses of profitability ratio analysis. We then empirically and theoretically compare most of the valuation and performance models in the literature. The chapter concludes with a critical assessment of empirical performance models and discusses the most appropriate one, which will be used in this thesis.

Chapter 4 discusses data collection, methodology and research design in addition to the statistical software packages that will be used in this thesis.

The empirical study in Chapter 5 investigates the relationship between the residual income model and market stock prices. We use a set of publicly available data of FTSE 500 and S&P 500 listed non-financial firms over the 1990-2002 period. We aim to test the hypothesis that the defined Residual Income model can significantly capture cross-sectional differences in stock market prices over the stated period of time. We apply regression analysis on panel data as well as cross-sectional analysis to test the behaviour of the model year by year. We further divide our sample into industry portfolios to assess industry differences. Chapter 5 also provides empirical tests for determining firm performance and assesses the possibilities of using the
Residual Income Component as well as the first difference of the Frankel and Lee model (1998) as a performance measure.

The study in Chapter 6 aims to test the relationship between gearing, over-investment and firm performance. We use non-financial firms from FTSE 500 and S&P 500. Firstly, we investigate the relationship between gearing and over-investment, which is based on the proposition that gearing mitigates over-investment. Secondly, we test the relationship between firm performance and over-investment. The aim is to ascertain whether over-investment destroys firm value or not. Thirdly, we test the effects of the interaction variable between gearing and over-investment on firm performance. This will enable us to see the effect of gearing and over-investment as one variable on performance. The aim is to find out whether gearing, given that it mitigates the over-investment problem, leads to better firm performance.

A summary and the conclusions are provided in Chapter 7. In this chapter we summarise the results from the preceding chapters, and discuss the findings. In addition, we describe the contributions to the existing literature, limitations of this study and extensions for further research.
Chapter 2

Capital Structure

2.1 Introduction

The capital structure issue is concerned with the optimum balance (if any) between equity and debt used to finance companies. It is one of the key areas in the economics of corporate finance, as it has implications for new security issues, the financing of takeover and buyout activities, and dividend policy, as retained earnings (the profits retained by a firm after payment of dividends) are a major source of equity funding. Modigliani and Miller’s (1958) analytical approach challenged the conventional view that a moderate amount of debt finance (in the form of corporate bonds, debentures or loan stock) was beneficial, but that higher levels were not prudent - indeed, high corporate debt levels had been cited as one of the factors causing the great stock market crash of 1929. Their analysis pointed initially to the irrelevancy of the debt-equity split, and in later work to the advantages of debt finance through corporate tax effects subject to financial distress costs. Miller (1977) subsequently integrated personal taxes into the framework and argued that the tax advantages of borrowing were small. In summary, in a world of perfect capital markets, Modigliani and Miller (1958) demonstrated that investment, financing and dividend decisions are independent.

During the last decades, however, the empirical evidence has shown that those decisions are interdependent. Accordingly, some theories that explain the previous evidence have been developed. These theories are based on capital-market imperfection; with respect to the investment decision, agency costs and the existence of asymmetric information between the main stakeholders constitute the foremost imperfection. The role of asymmetric information in investment decisions has as its primary basis the theoretical works of Jensen and Meckling (1976), Myers (1977) and
Myers and Majluf (1984). All the above mentioned papers show that information asymmetries may lead to some investment projects with a positive net present value (NPV) not being undertaken.

A second foundation in this study of inefficient investment decisions is the work of Jensen (1986). This paper, starting with the hypothesis of the existence of asymmetric information between managers and shareholders, introduces the so-called problem of over-investment, as a basic argument of his free cash flow theory. According to this theory, there can be negative NPV investment projects that end up being undertaken.

The objective of this chapter is to present the above-mentioned theories. It starts with Modigliani and Miller (1958, 1963) and Miller (1977) in the first and second sections, then moves to the cost of financial distress and trade-off theory in the third and fourth sections. In the fifth section, agency costs and asymmetry of information will be presented followed by free cash flow theory. This chapter ends with the conclusion, which aims to find a direct link between gearing and firms' performance.
2.2 Modigliani and Miller Theory

2.2.1 Introduction

There are many ways for the firm to raise its required funds. However, the most basic and important instruments are stocks (equity) and bonds (debt). The firm's mix of different securities is its capital structure. A natural question arises: What is the optimal debt-equity ratio? For example, if a firm needs £100 million for a project, should all this money be raised by issuing stocks, or 50% of stocks and 50% of bonds (debt-equity ratio 1:1) or any other combination?

Modigliani and Miller (MM) (1958) showed that financing decisions do not matter in perfect capital markets. Their famous proposition I states that:

'The total value of a firm is the same whatever its debt-equity ratio'

(assuming no corporate taxes).

If this is true, the basic exercise in capital budgeting can be directly applied to project evaluation for firms with different debt-equity ratios. However, in practice, capital structure does seem to matter. Then why do we bother to learn the MM theory? This theory is valid under certain conditions. An understanding of MM theory helps in understanding these conditions, which in turn helps us to understand why one particular capital structure is better than another. In addition, the theory tells us what kinds of market imperfections to attend to. The imperfections that are most likely to make a difference are taxes, the costs of bankruptcy and the costs of writing and enforcing complicated debt contracts.

2.2.2 MM Proposition I

Modigliani and Miller's (1958) propositions earned them the honour of two Nobel prizes. They provide a good start in understanding capital structure decisions. Modigliani and Miller (1958) argued that under certain conditions, the total value of all financial securities is the same regardless of the mix of different securities. They
proved their argument by showing that there would be arbitrage opportunities if the value of the firm depended on its financial structure. Because there should be no arbitrage in perfect capital market conditions, it follows that firms should be able to choose any mix of securities without changing their value (Rajan and Zingales, 1995). Their proposition on the capital structure is as follows:

"The market value of any firm is independent of its capital structure and is given by capitalizing its expected returns at the rate appropriate to its risk class" (Modigliani and Miller, 1958)

According to Brealey and Myers (2000), Modigliani and Miller (1958) made the following assumptions about the markets in which they were working:

1) Capital markets are perfect
2) There is no cost to bankruptcy.
3) Firms issue risky debt and equity, and have the same β risk.
4) All cash flow streams are perpetuities, and no growth is allowed.
5) Managers always maximize the shareholders’ wealth (implies no agency costs).
6) Homemade leverage is a perfect substitute for corporate leverage. That is, there is no difference between corporate and personal borrowing (necessary for arbitrage arguments).

In modern terms, capital structure is irrelevant, and firm value is equal to the present value of its free cash flow discounted at the relevant cost of capital.

**MM Proposition I**

\[
\text{Value of the levered firm} = \text{Value of un-levered Firm} \\
VL = VU
\]


The above figure is Modigliani and Miller’s proposition 1 and is one of the most important results of theory in corporate finance (Wald, 1999). It is a simple arbitrage
argument that, in a world without taxes, any investor of the unlevered firm can borrow money at the same interest as the levered firm and make money by investing in the unlevered company. Therefore, the value of the two companies will be the same (Copeland and Weston: 1992).

The key to their model, according to Megginson (1997), is the assumption which states that shares of firms within a given risk class have both the same expected return and the same probability distribution of expected returns, and can therefore be considered perfect substitutes for each other. Companies within a risk class thus differ from each other only in scale, they have the same expected profit per dollar of invested capital, and investors can expect their per share returns to be identical. MM suggested that these classes might be comparable to industrial classifications, and this is a useful and intuitive analogy. In addition, Higson (1998) pointed out that MM employ two concepts that are used in financial decisions to prove their argument, namely arbitrage and homemade alternative. Arbitrage is the process that ensures that two firms differing only in capital structure must have the same value. Homemade gearing describes that an investor holding an equity stake in a leveraged firm can sell his stake, raise a personal loan equal to the share that he held in the leveraged company, spend the proceeds on a firm that is not geared and increase his income at no cost (Higson, 1998).

Despite the fact that many of these assumptions are far from the reality of capital markets, relaxing them does not always lead to violation of the MM theory. It is difficult to find an example where a firm’s value might plausibly depend on financing. The most serious violations of MM’s proposition I create a moneymaking opportunity for firms and financial intermediaries. Any distortion of the normal function of capital markets creates unsatisfied investors that can become the clientele for new more attractive securities. Once the clientele is satisfied, proposition I is restored, until the next distortion (Brealey and Myers, 1991).

Although perfect capital market assumptions are unrealistic, Garvey (1992) argues that most of them do not pose serious problems for the theory. However, there are two assumptions that need highlighting as they have a significant effect on the results:
1. It is assumed there is no taxation: this is a serious problem because arguably one of the key advantages of debt is the tax relief on interest payments.

2. Risk in Modigliani and Miller’s theory is measured entirely by variability of cash flows. They ignore the possibility that cash flows might cease because of bankruptcy. This is another significant problem for the theory if borrowing is high (Garvey, 1992).

In addition, the assumption about homemade leverage also seems to contradict the fact that many individuals are constrained in the amount of credit they can borrow, and in any case cannot borrow at the same terms that firms can (Stiglitz, 1988).

### 2.2.3 MM Proposition II

In proposition II, the required return on equity is a linear function of the debt-equity ratio.

> “The expected yield of share of stock is equal to the appropriate capitalisation rate for a pure equity stream in the same risk class, plus a premium related to financial risk equal to the debt-to-equity ratio times the spread between the capitalisation rate and the cost of debt.” (Modigliani and Miller, 1958).

MM proposition II is represented graphically below:
Figure 2.1 MM Proposition II

Re = Expected return on equity

Rates of Return

Ra = Expected return on assets

Rd = Expected return on debt

Risk-free debt

Risky Debt D/E = debt/equity

Adopted from: Brealey and Myers (2000)

It can be seen from the figure above, as the debt-equity ratio increases, the expected return on equity increases so long as the debt is risk-free. However, "if leverage increases the risk of the debt, the demand for a higher return will cause the increase in return on equity to slow down" (Brealey and Myers, 2000). Indeed, according to Dickerson et al (1995), MM shows that the use of cheap debt gives shareholders a higher rate of return, but this higher return is precisely what they need to compensate for the increased risk from financial leverage. The graph of cost of capital against gearing (as measured by debt/equity ratio) is shown in Figure 2.3.
Solomon (1963) points out that as far as the leverage effect is concerned, there exists a clearly definable optimum position, namely the point at which the marginal cost of more debt is equal to a company’s average cost of capital. This point will be shown later in the Trade-off theory.

2.2.4 MM with Corporate Taxes

MM published a follow-up paper in 1963 in which they relaxed the assumption that there are no corporate taxes. The US tax code allows corporations to deduct interest
payments as an expense, but dividend payments to stockholders are not tax deductible. This differential treatment encourages corporations to use debt in their capital structures i.e. the value of the levered firm will equal that of an unlevered firm, plus the present value of the tax shield provided by debt (Modigliani and Miller: 1963).

A significant positive relationship between market values of firms and their debt tax shield, found in research made by Modigliani and Miller (1963), emphasises the importance of the leverage decision to firm value. MM theory with corporate taxes is represented below:

**MM Theory with Corporate Taxes**

\[
\text{Value of levered firm} = \text{Value of unlevered firm} + \text{Value of tax savings}
\]

Adopted from: Modigliani and Miller: (1963)

MM (1963) demonstrate that if the other assumptions hold, this treatment leads to a situation that calls for 100% debt financing. This means that a firm financed with 100% debt would have optimal capital structure.

**Figure 2.3: MM Proposition II (with corporate taxes)**

Rates of Return

\[ R_e = \text{Expected return on equity} \]

\[ \text{WACC} \]

\[ (1-Tc) \times r_{\text{debt}} = \text{after-tax expected return on debt} \]

Debt-Equity Ratio D/E

Adopted from: Brealey *et al* (1999)
Brigham and Houston (1999) point out that this suggestion is not observed in the real world due to the fact that the bondholders will become owners of the company and it would be unrealistic for them to ask a rate of return smaller that the previous owners. Baxter (1967) and Solomon (1963) point out that the tax correction model is unreasonable since it implies that the firm should utilise the maximum amount of debt in its capital structure. In addition, a question that should be raised is, what is happening to risk and residual claim?

2.2.5 Implications of the MM Theory

The implication of including corporate taxes in the model is that the firm's value is maximised when it is financed entirely by debt. This is not very plausible; no firm is financed wholly by debt. A number of real world constraints need consideration. Firstly, there are institutional and legal restrictions – some institutions will not purchase stock of a firm that has a debt-equity ratio exceeding a set cut-off. Secondly, there are costs imposed for going bankrupt that might persuade the firm's management not to increase the debt-equity ratio too high. Thirdly, the interest tax shield, which might increase as the company uses more debt, may exhaust taxable income (this suggests an upper boundary to the amount of debt). Finally, there may be conflicts of interest between stockholders and bondholders.

Each of these points suggests that the 100% debt policy may not be optimal for a firm. "If we look to the U.S market, the average debt-to-value ratio is less than 40%" (Alderson and Betker, 2000). Furthermore, "a survey of 768 of the largest industrial firms shows that 126 (16%) have no debt in their capital structures" (Alderson and Betker, 2000). This empirical evidence suggests that the 100% debt policy is clearly not what is observed. The wide range of debt-equity ratios in the market could indicate that the original proposition about the irrelevance of the capital structure may have more merit than we initially gave it.
2.3 The Miller Model

Miller (1977) highlights the limitations of his and Modigliani’s 1963 arguments by additionally considering the effect of personal taxation. “Miller argues that the debt Irrelevancy Theory could be resuscitated even in the presence of corporate taxes, if taxes on the dividends and interest income that individuals receive from firms were factored in the analysis” (Damodaran, 1997). Miller (1977) modified the MM theory by introducing both corporate and personal taxes. Accordingly, the gearing from debt is as follows:

\[ G = \frac{1 - (1 - T_c)(1 - T_{ps})}{1 - T_pB} ] BL. \]

Where:
- \( G \) is the tax advantage of gearing,
- \( T_c \) is the corporation tax rate,
- \( T_{ps} \) is the personal tax rate on dividends,
- \( T_pB \) is the personal tax rate on interest and
- \( BL \) is the market value of the levered firm’s debt

Adopted from Miller (1977)

In the Miller model, there is a personal tax advantage to equity because capital gains are taxed only on realization, and a corporate tax advantage to debt because interest is tax deductible. In equilibrium, people with personal tax rates above the corporate tax rate hold equity while those with rates below hold debt. This prediction is not consistent with what occurred in the U.S. in the late 1980’s and early 1990’s when there were no personal tax rates above the corporate rate. The Miller model suggests that there should have been a very large increase in debt used by corporations but there was only a small change.

As companies begin to borrow, managers must persuade investors to hold bonds instead of stocks. The bigger the tax bracket of the investor, the bigger the rate of interest payments that the firm must give in order to attract the particular investor. Companies can afford to bribe investors as long as the personal tax rate is smaller than the corporate tax rate. Migration of investors from being equity-holders to being bond holders stops when the corporate tax savings are equal to the personal tax loss. Consequently, the debt to equity ratio of a company depends on the corporate tax rate.
and the funds available to investors in the various tax brackets (Brealey and Myers, 1991).

Optimal capital structure can be explained by a Trade-Off between a gain from leverage and relative costs such as bankruptcy costs. Furthermore, in the observed market equilibrium, the effects of interest rates are seen before those of tax rates that are "grossed up", so that most or all of the interest rate tax shield is lost. Finally, it implies equilibrium of aggregate debt outstanding in the economy, determined by relative corporate and personal tax rates (Copeland and Weston, 1992).

De Angelo and Masulis (1980) analysed Miller’s personal tax theory and introduced the accounting depreciation and investment tax credits, where they stated that these non-debt corporate tax shields were sufficient to overturn the leverage irrelevancy theory. They remarked that these would lead to a market equilibrium, in which each firm has a unique interior optimum leverage decision solely due to the interaction of personal and corporate tax treatments of debt and equity.

2.4 Cost of Financial Distress

The Modigliani and Miller (1958) theory of capital structure is such that the product-market decisions of firms are separate from financial-market decisions. Essentially this is achieved by assuming there is perfect competition in product markets. In an oligopolistic industry where there are strategic interactions between firms in the product market, financial decisions are also likely to play an important role. Financial distress is defined as a firm’s inability to meet part or all of its financial obligations, a situation that may or may not lead to bankruptcy. Moreover, the firm is exposed to certain costs, direct or indirect, when it faces financial distress. Direct costs include expenses related to courts, lawyers, experts and accountants in addition to administrative expenses in case of bankruptcy proceedings.

Indirect costs, on the other hand, are expenses or economic losses that result from bankruptcy but are not cash expenses spent on the process itself. These include the diversion of management’s time while bankruptcy is underway, lost sales during and
after bankruptcy, constraints on capital investment and R&D spending, and loss of key employees after a firm becomes bankrupt (Megginson, 1997). The probability of financial distress increases as the debt financing increases, while the benefits of the debt tax shield increase as the use of debt financing increases; the former decreases the firm’s market value whereas the later increases it.

Baxter (1967), Kraus and Litzenberger (1973), Scott (1976) and Kim (1978) showed that the firm should trade off the tax advantage from debt financing against the risk of bankruptcy, and that an optimal capital structure would maximize the firm’s total market value. The optimal capital structure is represented graphically below:

**Figure 2.4: Cost of Capital and the Optimal Capital Structure**

![Graph showing cost of capital and optimal capital structure](image)

Optimal capital structure is determined by taking on increasing amounts of debt until the gain from leverage is equal to the present value of the expected loss from bankruptcy costs (Copeland and Weston, 1992)

Allen (1986), Brander and Lewis (1986), Maksimovic (1986) and others have considered various different aspects of these interactions between financing and product markets. Allen (1986) considers a duopoly model where a bankrupt firm is at
a strategic disadvantage in choosing its investment because the bankruptcy process forces it to delay its decision.

Brander and Lewis (1986) and Maksimovic (1986) analyse the role of debt as a pre-commitment device in oligopoly models. By taking on a large amount of debt a firm effectively pre-commits to a higher level of output. Titman (1984) and Maksimovic and Titman (1993) have considered the interaction between financial decisions and customers' decisions. Titman (1984) looks at the effect of an increased probability of bankruptcy on product price because of, for example, the difficulties in obtaining spare parts and servicing should the firm cease to exist. Maksimovic and Titman (1993) consider the relationship between capital structure and a firm's reputation incentives to maintain high product quality.

2.5 The Trade-Off Theory

The preceding arguments led to the development of "the Trade-Off theory of capital structure". The Trade-Off theory of capital structure has been the textbooks' staple for many years (see Kraus and Litzenberger (1973), Scott (1977), Taggart (1977), Haugen and Senbet (1978), Marsh (1982), Kane et al (1984) and Bradley et al (1984)). According to Brigham et al (1999), Trade-Off theory states that the optimal capital structure is the trade-off between the benefit of debt (i.e. the interest tax shelter) and the costs of debt (i.e. financial distress and agency costs). A summary of the Trade-Off theory is given in the figure below:
As can be seen from the figure 2.7, the diagonal straight line represents the value of the firm in a world without bankruptcy costs. The curved line represents the value of the firm with these costs. The curved line rises as the firm moves from all equity to a small amount of debt. Here the expected present value of the distress costs is minimal because the probability of distress is so small. However, as more and more debt is incurred, the present value of these costs rises at an increasing rate. At some point, the rise in the present value of these costs from an additional dollar of debt equals the rise in the present value of tax shield. This is the debt level maximising the value of the firm and is represented by (B) in the figure above. In other words, (B) is the optimal amount of debt. Bankruptcy costs increase faster than the tax shield advantages beyond this point, implying a reduction in firm value from further leverage. Accordingly, value of levered firms will be as follows:

\[
\text{Value of levered firm} = \text{value of un-levered firm} + \text{value of tax savings} - \text{Present value of expected costs of financial distress}
\]

Adopted from: Brealey and Myers (2000)
Brealey and Myers (2000) point out that the Trade-Off theory recognises that target debt ratios may vary from firm to firm. They stated that companies with safe tangible assets and plenty of taxable income to shield ought to have high target ratios. Unprofitable companies with risky, intangible assets ought to rely primarily on equity financing. The table below represents trade off in a balance sheet format:

**Table 2.1: Trade-Off in a Balance Sheet Format**

<table>
<thead>
<tr>
<th>Advantages to Borrowing</th>
<th>Disadvantages to Borrowing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Tax benefit:</strong></td>
<td><strong>A. Bankruptcy costs:</strong></td>
</tr>
<tr>
<td>Higher tax rates = Higher tax benefit</td>
<td>Higher business risk = Higher cost</td>
</tr>
<tr>
<td><strong>B. Added discipline:</strong></td>
<td><strong>B. Agency cost:</strong></td>
</tr>
<tr>
<td>Greater the separation between the managers and stockholders = greater the benefit</td>
<td>Greater the separation between stockholders and lenders = Higher cost</td>
</tr>
<tr>
<td></td>
<td><strong>C. Loss of future financing flexibility:</strong></td>
</tr>
<tr>
<td></td>
<td>Greater the uncertainly about the future financing needs = Higher cost</td>
</tr>
</tbody>
</table>

Adopted from: Damodaran (1997)

Myers (1990) argues that Trade-Off theory avoids a corner solution (i.e. it does not give a clear explanation to what is observed in reality) and rationalises moderate borrowing in a way that can easily be understood. Nevertheless, closer analysis of financial distress gives a testable prediction from the Trade-Off theory, since these costs should be most serious for firms with valuable intangible assets and growth opportunities (Myers, 1990).

However, the Trade-Off theory of capital structure does not provide a satisfactory explanation of actual practice, because the empirical magnitudes of bankruptcy costs and interest tax shields do not match observed capital structures (Bernanek et al, 1995). Ross et al (1993) pointed out that no formula exists in Trade-Off theory at this time to exactly determine the optimal debt level for any particular firm. This is because bankruptcy costs cannot be expressed in a precise way.
2.6 Agency Costs

2.6.1 Introduction

Agency theory defines the firm as "simply one form of legal fiction which serves as a nexus for contracting relationships and which is also characterized by the existence of divisible residual claims on the assets and the cash flow of the organization which can generally be sold without permission of the other contracting individuals" (Jensen and Meckling, 1976: p311). In their paper Jensen and Meckling discuss agency costs as the key tool in evaluating alternative designs of principal-agent relations. They define agency costs as the sum of: 1) monitoring expenditures by the principal; 2) the bonding expenditures by the agent and; 3) the residual loss i.e. the monetary equivalent of the reduction in the welfare of the principal as the result of the differences between the agent’s decisions and those decisions which would maximize the welfare of the principal. Accordingly, major insights into the problems of capital structure can be gained if they are understood in terms of principal-agent theory. The outcomes of the firm are no longer independent of financing decisions, not only because of tax advantages but also for reasons of conflicts of interests, which may result in different optimal investment decisions.

The main issues that agency theory tries to address are:

1. To study the influence of the sharing rule in itself through a detailed analysis of the agent’s actions if a certain capital structure is assumed and if the firm has already had the financing.

2. To devise a certain kind of framework that will enable us to identify the ultimate bearers of the losses resulting from the agency problems.

3. To advise different instruments that can be used to tackle these agency problems in different scenarios.
2.6.2 Agency Cost of External Equity

The agency cost theory of financial structure was put forward by Jensen and Meckling in their 1976 paper. They observe that when an entrepreneur owns 100% of the stock of a company there is no separation between corporate ownership and control. This means that the entrepreneur bears all of the costs, and reaps all of the benefits of his or her actions. Once a fraction (α) of the firm’s stock is sold to outside investors, however, the entrepreneur bears only 1-α of the consequences of his or her actions. This gives the entrepreneur a clear incentive, in Jensen and Meckling’s tactful phrasing to “consume perquisites”. However, the entrepreneur is charged in advance for the “perk” he or she is expected to consume after the equity sale. Since selling stock to outside investors creates agency costs of equity, which are born solely by the entrepreneur, but which also harm society by reducing the value of corporate assets and discouraging additional entrepreneurship.

The agency conflict that derives from the manager’s tendency to appropriate perquisites out of the firm’s resources for his own consumption is not the only or the most important conflict. It is likely that the most important conflict arises from the fact that as the manager’s ownership falls, his effort to devote significant effort to creative activities such as searching out new profitable projects falls, i.e. he may shirk. He may in fact avoid such projects simply because it requires too much trouble or effort on his part to learn about them. Avoidance of these personal costs and the anxieties that go with them represent a source of on-job utility to him, and this shirking can result in the value of the firm being substantially lower than it otherwise could be. In practice, it is possible by expending resources to alter the opportunity the manager has for receiving non-money-related benefit. These methods include auditing, formal control systems, budget restrictions and the establishment of incentive compensation systems, which serve to bring the manager’s interests closer to those of outside shareholders.

Selling external equity is vital for entrepreneurs both because of individual portfolio diversification demands and because of the need to finance corporate growth once it outstrips personal wealth constraints. Jensen and Meckling point out that using debt finance can help overcome the agency problem costs of external equity in two ways.
Firstly, using debt by definition means that less external equity will have to be sold to raise a given amount of external financing. If agency costs of outside equity rise more than proportionally as \( \alpha \) increases, then economising on the amount of outside equity sold will reduce the deadweight agency costs of the manager/stockholder relationship. The second and more important effect of employing outside debt rather than equity financing is that this reduces the scope for excessive managerial perquisite consumption. The burden of having to make regular, contractually enforceable, debt service payments serve as a very effective tool for disciplining entrepreneurs. With debt outstanding, the cost of excessive perk consumption might well include the entrepreneur losing control of his or her company following default and bondholder seizure of the company’s assets.

Jensen and Meckling (1976) raise the question “Why don’t we observe large corporations individually owned with a tiny fraction of the capital supplied by the entrepreneur in return for 100% of the equity and the rest simply borrowed?” The reasons they give for this are: 1) the incentive effects associated with highly geared firms, 2) the monitoring costs these incentive effects lead to and 3) bankruptcy costs. All these costs are simply particular aspects of the agency costs associated with the existence of debt claims on the firm. This argument has led to another agency problem, i.e. the agency cost of debt which will be discussed in the next paragraph.

2.6.3 Agency Cost of Debt

Jensen and Meckling (1976) pointed out another agency problem in corporations, the agency problem between equity-holders, managers and bondholders. As the fraction of debt in a firm’s capital structure increases, bondholders begin taking on an increasing fraction of the firm’s business and operating risk, but shareholders and managers still control the firm’s investment and operating decisions. This gives managers a variety of incentives to expropriate bondholder wealth for the benefit of themselves and the shareholders they represent. The easiest way to do this would be to float a bond issue, and then pay out the money raised to shareholders as a dividend. After default, the bondholders would be left with an empty corporate shell, and
limited liability would prevent the bondholders from trying to collect directly from shareholders. As reported by Ross *et al* (1999):

"When a firm has debt, conflicts of interest may arise between stockholders and bondholders. Because of this, stockholders are tempted to pursue selfish strategies. These conflicts of interest, which are magnified when financial distress is incurred, impose agency costs on the Firm."

Another way the shareholders can separate bondholders from their wealth is to borrow money on the promise that it will be used to finance a safe investment and then actually invest in a risky project. If these investments are successful, shareholders can fully repay bondholders and pocket any excess project returns. If the project is unsuccessful, shareholders simply default and bondholders take over an empty corporate shell. The most effective preventive steps bond investors can take, according to Megginson (1997), involve writing very detailed covenants into bond contracts, which sharply constrain the ability of the borrowing firm’s managers to engage in inappropriate behaviour. However, these covenants make bond agreements immensely costly to negotiate and enforce, and in constraining management’s ability to make valuing value-increasing investment (Smith *et al*, 1979).

The second problem generated from the conflict between shareholders and bondholders is what is called, according to Myers (1977), the “Moral Hazard”. If the firm were totally equity funded, shareholders would accept all positive NPV projects. However, when partially funded with debt the shareholders may have an incentive to reject some positive NPV investments (Myers 1977). Myers argues that when a firm’s assets are largely made up of growth opportunities it would be difficult to fund the firm with debt because of the shareholders’ incentive to under-invest. Myers pointed that if there is a large amount of debt outstanding which is not backed by cash flows from the firm’s assets, i.e. a ‘debt overhang’, equity-holders may be reluctant to invest in safe, profitable projects because the bondholders will claim the lions’ share of the cash flow i.e. equity-holders may not receive anything.

Under the NPV rule a firm would accept a project if the present value of the generated cash flows exceeded the present value of the cost of investment. Some positive NPV
projects could generate enough proceeds to cover the basic cost of purchasing the
required inputs and capital goods but still not generate enough cash flow to also pay
the outstanding debt. The shareholders would not accept the project unless its
expected proceeds would both cover the cost of investment and pay-off the debt-
holders.

The third problem that is generated from the conflict between shareholders and
bondholders is the “Adverse Selection” problem developed by Stigliz and Weiss
(1981). Stigliz and Weiss illustrate that the inability of lenders to distinguish between
good and bad risks ex ante prevents them from charging variable interest rates
dependent on the actual risk. In this event, lenders are forced to increase the general
cost of borrowing, which will tend to induce a problem of adverse selection, as good
risks are driven from the market by the high costs of borrowing. Due to this
information asymmetry, companies will tend to prefer internal to external financing,
where available for financing good risks.

The study of Jensen and Meckling (1976) predicts that managers of an individual firm,
starting from an all-equity position, will substitute bonds for stock in the firm’s capital
structure in order to reduce the agency cost of equity. As this process continues,
however, the agency costs of debt begin to rise at an increasing rate. The firm’s
optimal debt to equity ratio is reached at the point where the agency cost of an
additional dollar of debt exactly equals the agency cost of the dollar of equity retired.
This situation is represented below:
The agency perspective has also led to a series of important papers by Hart and Moore (1989, 1994), Aghion and Bolton (1992), Berglof and von Thadden (1994), Von Thadden (1995) and Hart (1995) on financial contracts. Hart and Moore (1989) consider an entrepreneur who wishes to raise funds to undertake a project. Both the entrepreneur and the outside investor can observe the project payoffs at each date, but they cannot write explicit contracts based on these payoffs because third parties such as courts cannot observe them. The focus of their analysis is the problem of providing an incentive for the entrepreneur to repay the borrowed funds. Among other things, it is shown that the optimal contract is a debt contract and incentives to repay are provided by the ability of the creditor to seize the entrepreneur’s assets.

Titman (1984) points out that agency costs are important for contracts between the firm and the customer or between the firm and its employees. If, for example, a firm produces durable goods that require future servicing, the customer does not only buy...
ownership, but also future servicing. Therefore, customers must assess the probability of bankruptcy and weigh it in their decision to purchase durable goods. Firms that produce durable goods will have lower demand for debt. Consequently, if labour markets are competitive, then labourers will charge lower wages to work for a firm that has a lower probability of bankruptcy. Thus, one should expect to find that firms which use a larger percentage of job-specific human capital carry less debt.

In Harris and Raviv (1990) and Stulz (1990), managers disagree over an operational decision. In particular, in the Harris and Raviv model managers are assumed to want always to continue the firm’s current operations even if investors prefer the firm’s liquidation. In the Stulz model, managers are assumed to want always to invest all available funds even if paying out cash is better for investors, therefore creating an over-investment conflict. In both cases, it is assumed that the conflict cannot be resolved through contracts based on cash flow and investment expenditure. Debt mitigates the problem in the Harris and Raviv model by giving investors (debt-holders) the option to force liquidation if cash flows are inadequate, whereas in the Stulz model, debt payment reduces free cash flow. The cost of debt in Stulz’s model is that debt payments may more than exhaust free cash flow, reducing the funds available for profitable investment (under-investment). The optimal capital structure in Harris and Raviv trades off improved liquidation decisions versus higher investigation costs. A larger debt level improves the liquidation decision because it makes default more likely.

2.7 Signalling and Asymmetric Information Theory

MM assumed that investors have the same information about a firm’s prospects as its managers i.e. symmetric information, because those who are inside the firm (managers and employees) and those who are outside the firm (investors) have identical information. However, in fact managers generally have better information about their firms than do outside investors; i.e. asymmetric information (Besley and Brigham, 2000). A signal, as defined by Megginson (1997), is an action that imposes deadweight costs on the signaller in order to convey value to relatively poorly
informed outsiders (usually investors). The signal is credible if it is prohibitively costly for a weaker firm to attempt to mimic.

An early contribution is the Ross (1977) paper on the “incentive signaling approach”, where he developed a signaling model of corporate capital structure based on asymmetric information problems between well-informed managers and poorly informed outsider shareholders. This model is based on the idea that corporate executives with favourable inside information about their firms have a clear incentive to convey this positive information somehow to outside investors, in order to increase the market price of shares. However, managers cannot simply announce that they have good news because every other manager has the incentive to do the same, and the market will be appropriately cautious towards any self-serving statement which can only be proved to be true as the time passes. One solution for this problem is for managers having “good news”, i.e. those of high-value firms, to signal it to the investors by taking some action that is prohibitively costly for the managers that have “bad news” i.e., of low-value firms, to mimic.

Ross shows that it is possible to design an incentive based-compensation contract for managers of high value firms that will induce them to use a heavily leveraged capital structure for their companies. Less valuable companies are unwilling to assume so much debt because they are much more likely to fall into bankruptcy. Based on these assumptions, a separating equilibrium occurs where high value firms use more debt financing and less valuable companies rely more on equity. Investors are able to differentiate between high and low value firms by looking at their capital structure and are willing to assign higher valuations to highly levered firms, since weaker firms are unwilling to mimic the stronger ones by borrowing extra debt, thus, the equilibrium is enforced.

Leland and Pyle (1977) consider a situation where entrepreneurs use their retained share of ownership in a firm to signal its value. Owners of high value firms retain a high share of the firm to signal their type. Their high retention means they do not get to diversify as much as they would if there was symmetric information and this makes it unattractive for low value firms to mimic them.
Megginson (1997) points out that the signalling models explain market responses to the different types of security issues. Debt issues signal good news (managers are confident about the future, and are greeted with a positive stock price response, while equity issues signal bad news (earnings per share will decline in the future), and are met with significant stock price declines. However, leverage ratios are inversely related to profitability in almost every industry, not directly related as the signalling models predict. Furthermore, the signalling model predicts that industries rich in growth options and other intangible assets should employ more debt than mature, tangible-asset-rich industries because growth companies have more severe information asymmetry problems, and thus greater need to signal. In reality, exactly the opposite pattern is observed; asset-rich companies use far more debt than do growth companies. Two subsequent papers based on asymmetric information have been very influential (Myers, 1984; Myers and Majluf, 1984). If managers are better informed about the prospects of the firm than the capital markets they will be unwilling to issue equity to finance investment projects if the equity is undervalued. Instead, they will prefer to use equity when it is overvalued. Thus, equity is regarded as a bad signal. Myers (1984) uses this kind of reasoning to develop the “pecking order” theory of financing, which will be discussed in the next section

2.8 The Pecking Order Theory

Myers (1984) proposed the Pecking Order Theory of Leverage, which is based on four observations and assumptions about corporate behaviour:

1) Dividend policy is “sticky”. Managers try at all costs to maintain a constant dollar-share dividend payment, and will neither increase nor decrease dividends in response to temporary fluctuation in current profit.

2) Firms prefer internal financing (retained earnings) to external financing of any sort, debt or equity.

3) If a firm must obtain more external financing, it will choose the safest (least risks) security first.
4) As a firm is required to obtain more external financing, it will work down the pecking order of securities, beginning with very safe debt, and then progressing through risky debt, convertible securities, preferred stock, and finally common stock as a last resort.

This model focuses on the motivations of the corporate manager, rather than on capital market valuation principles. Megginson (1997) points out that this model had been largely ignored by modern economists because it seemed to be based on irrational, value-decreasing corporate behaviour that financial natural selection should have existed long before. Indeed, the simple pecking order model presumes severe market imperfections (very high transactions costs, uninformed investors, and managers who are completely insensitive to the firm’s stock market valuation) that are hard to accept as accurate portraits of modern capital markets.

Myers (1984) provides a viable theoretical justification for the pecking order theory, based on asymmetric information; Myers and Majluf (1984) make two key assumptions about corporate managers. Firstly, they assumed a firm’s managers know more about the company’s current earnings and investment opportunities than do outside investors. Secondly, they assumed managers act in the best interest of the firm’s existing shareholders. These two assumptions, according to Megginson (1997), are crucial because the asymmetric information assumption implies that managers who develop or discover a marvelous new positive-NPV investment opportunity are unable to convey that information to outside shareholders because the managers’ statements will not be believed.

After all, every management team has an incentive to announce wondrous new projects in order to bid up the firm’s stock price, so they can sell shares at an unjustifiably high price. In addition, since investors are unable to verify these claims until long after the fact, they will assign a low average value to the stocks of all firms and will buy new equity issues only at a large discount from their equilibrium values without informational asymmetries. Corporate managers understand these problems, and in certain cases will refuse to accept positive-NPV investment opportunities if this would entail issuing new equity, since this would give away too much of the project’s value to the new shareholders at the expense of the old.
Furthermore, information problems in financial markets are problems caused by human nature, and thus are not soluble through reductions in transaction costs or other capital market innovations. Therefore, the solution to this pervasive problem of modern corporate finance, according to Myers and Majluf, is to retain sufficient financial slack\(^2\) to be able to fund positive-NPV projects internally. Firms with sufficient financial slack will never have to issue risky debt or equity securities in order to fund their investment projects, and they are thus able to finesse asymmetric information problems between managers and investors.

Bruner (1988) points out that this model provides an explanation for the observed pattern of profitable firms retaining their earnings as equity and building up their cash reserves; they are building both financial slack and financial flexibility. Furthermore, Besley and Brigham (2000) state that the Myers and Majluf model explains stock market reactions to leverage-increasing and leverage-decreasing events. Since firms with valuable investment opportunities find a way to finance their projects internally, or use the least risky securities possible if they have to obtain financing externally, the only firms that will issue equity are those with managers who consider the firm’s shares to be over-valued. Investors understand these incentives, and also realize that managers are better informed about a firm’s prospects than they are, and therefore investors always greet the announcement of a new equity issue as “bad news” (as a sign that management considers the firm’s shares to be over-valued or that it has exhausted its debt capacity).

Unfortunately, the Pecking Order Theory cannot explain all the capital structure regularities observed in practice. For example, it suffers in comparison with the trade-off theory in its ability to explain how taxes, bankruptcy cost, security issuance costs, and the individual firm’s investment opportunity set influence that company’s actual debt ratio. Furthermore, the theory ignores significant agency problems that can easily arise when a firm’s managers accumulate so much financial slack that they become immune to market discipline (Jensen, 1993). On the other hand, Shyam-Sunder and Myers (1993) empirically compare the Pecking Order and trade-off models of corporate leverage, and find the Trade-Off model wanting. In fact their tests show the power of some usual tests of the Trade-off model to be virtually nil.

---

\(^2\) Financial slack is defined to include a firm’s cash and marketable securities holdings, as well as unused (risk free) debt capacity.
Nonetheless, the Pecking Order theory of capital structure seems to explain certain aspects of observed corporate behaviour better than any other model does, and this is particularly true of corporate financing choice and market responses to security issues.

2.9 Free Cash Flow Theory: Jensen (1986)

Jensen (1986) expanded on the issues of management's deviations from optimal investment policies and perquisite consumption raised in Jensen and Meckling (1976). The work of Jensen starting from the hypothesis of the existence of asymmetric information between managers and shareholders, introduces the so-called problem of over-investment, as a basic argument of his free cash flow theory. According to this theory, negative NPV investment projects can be undertaken.

The over-investment process arises from the conflict between managers and shareholders. When information asymmetries exist, and taking into account that mechanisms used to align the interests between shareholders and managers may not be fully efficient, managers may use the free cash flow to undertake negative NPV projects in their own best interests (Jensen, 1986). Note that free cash flow is cash flow in excess of that required to fund all positive NPV projects, hence managers waste these funds instead of paying them to shareholders. Managers will have incentives to over-invest because of the pecuniary and non-pecuniary benefits associated with the larger dimension of the firm (Jensen, 1986).

In mature industries where firms generate large cash flows, but have only few growth opportunities, managers tend to over-invest and diversify into industries in which they have little knowledge. By issuing debt in exchange for stock, managers are bonding their promise to pay future cash flows in a way that cannot be accomplished by simple dividend increases. Jensen calls this the 'Control Hypothesis' for debt creation. Debt creation, without retention of the proceeds of the issue, enables managers to effectively bond their promise to pay out future cash flows. Debt provides a means of bonding managers' promises to pay out future cash flows rather than investing in wealth-destroying ventures. It also provides the means for controlling opportunistic
behaviour by reducing the cash flows available for discretionary spending. Top managers’ attention is then clearly focused on those activities necessary to ensure that debt payments are met. This use of debt as a disciplinary tool makes survival the central issue for all concerned.

2.10 Empirical Evidence on Capital Structure

Modigliani and Miller (1963) argue that, due to the tax deductibility of interest payments, companies may prefer debt to equity. This would suggest that highly profitable firms would choose to have high levels of debt in order to obtain attractive tax shields. However, others such as Miller (1977) highlight the limitations of his and Modigliani’s 1963 arguments by additionally considering the effect of personal taxation. Moreover, De Angelo and Masulis (1980) argue that interest tax shields may be unimportant to companies with other tax shields, such as depreciation. They stated that introduction of such non-debt tax shields leads to the conclusion that each firm has a unique interior capital structure that maximises its value. This capital structure is determined only by the interactions of personal and corporate taxes as well as positive default (financial distress) costs.

Ashton (1989) argues that if there is a UK Tax advantage of debt, it is likely to be much smaller than the traditional MM value. The tax advantage of debt is considerably less under the UK imputation system than it is under the US classical system. Giner and Reverte (2001) used the Ohlson (1995) framework to examine the usefulness of the information about debt level on stock prices. They take into account the relationship between the cost of debt and the return on investment. Thereafter, they consider the degree of company leverage relative to the optimal level and the future prospects facing the firm are included in their analysis to further examine the value of debt. Their results seem to support the optimal capital structure theory in that deviations of the debt-equity ratio from its optimal level are negatively perceived by investor. In addition, their results also seem to be consistent with the signalling theory, as debt seems to be a positive signal for firms with higher future expected earnings relative to those where expected earning are not so high.
An alternative hypothesis regarding the relationship between profitability and gearing relates to Myers and Majluf (1984) and Myers (1984) pecking-order theory. Based on asymmetric information, they predict that companies will prefer internal to external capital sources. Consequently, companies with high levels of profits will prefer to finance investments with retained earnings rather than by the raising of debt finance. The finding of Rajan and Zingales (1995) of a negative relationship between gearing and profitability is consistent with Myers' (1984) pecking-order theory.

In their cross-sectional study of the determinants of capital structure, Rajan and Zingales (1995) examine the extent to which, at the level of the individual firm, gearing may be explained by four key factors, namely market-to-book, size, profitability and asset tangibility. Their analysis is based upon a firm-level sample from each of the G-7 countries, and although the results of their regression analysis differ slightly across countries, they appear to uncover some strong conclusions. The market-to-book ratio is used by Rajan and Zingales (1995) as a proxy for the level of growth opportunities available to the enterprise. This is in common with most studies, which tend to apply proxies, rather than valuation models, to estimate growth opportunities (Danbolt et al. 1999). Rajan and Zingales suggest that, a priori, one would expect a negative relationship between growth opportunities and the level of gearing.

This is consistent with the theoretical predictions of Jensen and Meckling (1976) based on agency theory, and the work of Myers (1977), who argues that, due to information asymmetries, companies with high gearing would have a tendency to pass up positive NPV investment opportunities\(^3\). Myers (1977), therefore, argues that companies with large amounts of investment opportunities (also known as growth options) would tend to have low gearing ratios. Moreover, as growth opportunities do not yet provide revenue, companies may be reluctant to take on large amounts of contractual liabilities at this stage. Similarly, as growth opportunities are largely intangible, they may provide limited collateral value or liquidation value (in a similar spirit to the discussion of tangibility below). Companies with growth options may thus not wish to incur — nor necessarily be offered — additional debt financing.

---

\(^3\) The reasons for passing up positive NPV investments are explained before in this chapter in section 2.6.3 Agency cost of Debt.
However, the empirical evidence regarding the relationship between gearing and growth opportunities is rather mixed. While Titman and Wessels (1988), Chung (1993) and Barclay et al. (1995) find a negative correlation, Kester (1986) does not find any support for the predicted negative relationship between growth opportunities and gearing. Despite this controversy, however, Rajan and Zingales (1995) uncover evidence of negative correlations between market-to-book and gearing for all G-7 countries. This is thus consistent with the hypotheses of Jensen and Mekling (1976) and Myers (1977), and lends weight to the notion that companies with high levels of growth opportunities can be expected to have low levels of gearing.

Secondly, Rajan and Zingales include size (which is proxied by the natural logarithm of sales) in their cross-sectional analysis. There is no clear theory to provide *ex ante* expectations as to the effect which size should have on gearing. Rajan and Zingales (1995) state that:

"The effect of size on equilibrium leverage is more ambiguous. Larger firms tend to be more diversified and fail less often, so size (computed as the logarithm of net sales) may be an inverse proxy for the probability of bankruptcy".

Rajan and Zingales (1995)

In addition, larger companies are more likely to have a credit rating and thus have access to non-bank debt financing, which is usually unavailable to smaller companies. While the prior empirical evidence with regard to the relationship between size and gearing is rather mixed, Rajan and Zingales (1995) find gearing for UK companies to be positively related to sales, as hypothesized.

Thirdly, consistent with Toy et al. (1974), Kester (1986) and Titman and Wessles (1988), Rajan and Zingales (1995) find profitability to be negatively related to gearing. Given, however, that the analysis is effectively performed as an estimation of a reduced form, such a result masks the underlying demand and supply interaction which is likely to be taking place. Although on the supply-side one would expect that firms that are more profitable would have better access to debt, the demand for debt may be negatively related to profits. Stiglitz and Weiss (1981) illustrate, as explained before, that the inability of lenders to distinguish between good and bad risks *ex ante*
prevents them from charging variable interest rates dependent on the actual risk. In this event, lenders are forced to increase the general cost of borrowing, which will tend to induce a problem of adverse selection, as good risks are driven from the market by the high costs of borrowing. Due to this information asymmetry, companies will tend to prefer internal to external financing, where available.

Consistent with the findings of Bradley et al. (1984) and Titman and Wessels (1988), Rajan and Zingales' (1995) study of capital structure in the G-7 economies produces evidence to suggest a positive relationship between asset tangibility, which they define as the ratio of fixed to total assets, and gearing. Following the theories of Scott (1977), Williamson (1996) and Harris and Raviv (1990), Rajan and Zingales (1995) suggest this may reflect the fact that debt may be more readily available to a firm which has high amounts of collateral upon which to secure debt, thus reducing agency problems.

Almost all the above studies find a strong dependence of investment on the availability of internal funds, this dependence being interpreted as evidence of the under-investment problem due to adverse selection. However, the positive relationship found between investment and cash flow may not arise only from the under-investment problem due to adverse selection. It may also indicate that high levels of cash flow allows managers to undertake negative NPV projects, which would not happen if they had to raise external funds and explain the rationality of their investments. Thus, the over-investment hypothesis is confirmed whenever the positive relationship between investment and cash flows is maintained for firms whose investment opportunities are of low quality. On the contrary, for firms with valuable investment opportunities, a positive relationship indicates an under-investment problem.

The over-investment idea has gained much support in the literature. For instance, De Jong and Veld (2001) find that Dutch managers avoid the disciplining role of debt i.e. they avoid using debt and over-invest, however, the market reaction shows that this over-investment behaviour is recognised, and the authors interpret this as evidence of over-investment. Moreover, the relation between firm performance and managerial ownership has been used in previous work to support the over-investment model. Morck et al (1988) estimate a piecewise linear relation between managerial ownership
and firm performance. They find that firm performance is increasing in managerial ownership for ownership levels below 5 percent or over 25 percent but decreasing in ownership for ownership levels between 5 and 25 percent of the firm. They interpret their results as evidence that managers make investment decisions that entrench them in their positions for ownership in this range. As a result of entrenchment, firm performance is lower. Many subsequent papers (McConnell and Servaes (1990), Himmelberg, Hubbard, and Palia (1999) and Palia (2001) have conducted similar analysis with mixed results. Other support for over-investment comes from Jensen (1993), who provides illustrative calculations of the destruction of shareholders’ value at a number of the world’s largest corporations.

Roy and Mingfang (2000) investigated the argument that decisions concerning the choice of capital structure need to be linked with a firm’s competitive environment, more specifically to environmental dynamism, the degree and the instability of changes in a firm’s competitive environment. Roy and Mingfang argued that while the tax advantages of increased debt are recognised, by increasing debt a firm introduces a stakeholder group of lenders who, by definition, have a short-term orientation. This group is potentially able to limit the freedom of choice available to managers in the selection of strategies to contend with competitive threats or opportunities especially when firms need and depend on creative and innovative strategic choices to thrive and succeed. The authors integrated complementary elements from agency theory and transaction cost economics with elements of strategic management in an effort to provide a more holistic view of the capital structure decision as it relates to the nature of the firm’s competitive environment, which provides opportunities to examine the linkage among environment, capital structure and organisational performance.

From the literature review, the authors argued that agency theory does not take into consideration competitive environments, or the necessity for managers to make choices beyond a stakeholder wealth-maximizing perspective. They added that debt and equity are not just financial instruments but are more a means of corporate governance;

"From a strategic management perspective, there is a clear indication that external factors can influence the efficacy of the capital structure"
decision with respect to the ability of the firm to make critical choices in response to competitive pressure. However, with agency theory, transaction cost economics does not take into consideration the competitive environment such that it can provide adequate prescriptive advice as to how to create a capital structure that will ensure the long-term survival of the firm”

(Roy and Mingfang, 2000)

Drastic changes are taking place in the competitive landscape; advanced technologies are beginning to alter the effectiveness of traditional competitors’ approaches, and to introduce a new array of competitive weapons. Accordingly, the authors argued that whether firms adopt some of the more traditional strategic actions or emerging actions, these actions seem to be associated with greater risk, novelty, the need for investment in specialized assets, and more importantly, they may be more difficult to value by outside groups.

Miguel and Pindado (2001) analysed the firm characteristics such as: tangible and intangible fixed assets, tax aspects of capital structure, financial distress and aspect related to the interdependence between investment and debt, which are determinants of capital structure according to different explanatory theories. They also analyze how institutional characteristics (such as tax code, bankruptcy laws, state of development of bond markets and patterns of ownership) affect capital structure. Their purpose is to provide additional international empirical evidence on capital structure in two aspects. First, they provide evidence of firm characteristics that determine capital structure for a non-G-7 country (Spain). Second, they examine how the institutional characteristics affect capital structure. They have developed a target adjustment model, in order to study the debt of Spanish firms, which allows them to explain the firm’s debt in terms of its debt in the previous period and the firm’s target debt level, the letter being a function depending on the firm characteristics, which according to the different theories explain the capital structure.

Chkir and Cosset (2001) examined the relationship between the debt level of multinational corporations (MNCs) and their diversification strategy. By integrating both the international market and the product dimension of diversification into the analysis and by utilizing a switching regression model that allows the effect of the
determinants of the capital structure of MNCs to vary with the strategy of diversification, their paper sheds new light on the debt policy of MNCs. The switching regression methodology that they used, allows one to examine the trade-off between the debt-reducing effect of agency costs due to international diversification and the debt-increasing effect of risk reduction due to industrial diversification.

2.11 Conclusion

The main conclusion to be drawn from the diverse literature on capital structure and financing decisions is the absence of any universal consensus on optimal financing structure. The seminal work in this area was Modigliani and Miller (1958). They showed that with perfect markets (i.e., no frictions and symmetric information) and no taxes, the total value of a firm is independent of its debt/equity ratio. Similarly, they demonstrated that the value of the firm is independent of the level of dividends. In their framework, it is the investment decisions of the firm that are important in determining its total value.

The purpose of the Modigliani and Miller theorems was not as a description of reality. Instead, it was to stress the importance of taxes and capital market imperfections in determining corporate financial policies. Incorporating the tax deductibility of interest (but not of dividends) and bankruptcy costs led to the Trade-Off theory of capital structure. Some debt is desirable because of the tax shield arising from interest deductibility, but the costs of bankruptcy and financial distress limit the amount that should be used. In addition, debt, from the agency's point of view, is preferable, since it reduces the agency cost between the managers and the shareholders. However, as the level of debt increases the agency problem between the bondholders and the shareholders increase. Hence, the result is a trade-off between debt and equity.

The Trade-Off theory of capital structure does not provide a satisfactory explanation of actual practice. The tax advantage of debt, relative to the magnitude of expected bankruptcy costs, would seem to imply that firms should use more debt than is actually observed. Attempts to explain this, such as Miller (1977), which incorporates
personal as well as corporate taxes into the theory of capital structure, have not been successful. In the Miller (1977) model, there is a personal tax advantage to equity because capital gains are only taxed on realization and a corporate tax advantage to debt because interest is tax deductible. In equilibrium, people with personal tax rates above the corporate tax rate hold equity while those with rates below hold debt. This prediction is not consistent with what occurred in the U.S. in the late 1980’s and early 1990’s when there were no personal tax rates above the corporate rate. The Miller (1977) model suggests that there should have been a very large increase for debt used by corporations, but the change was small.

In imperfect capital markets, financing and investment decisions are not independent. In fact, capital market imperfections, such as information asymmetries and agency costs, could lead to either under-investment or over-investment processes, i.e., not all positive NPV projects will be undertaken and some negative NPV projects not rejected. Informational asymmetries contribute to several conflicts between the main stakeholders, which give rise to either over-investment or under-investment processes and hence either increases or decreases firm’s performance. These topics and how they are connected are described in the figure 2.7 below:
The under-investment process is caused by agency conflicts between main stakeholders, which are facilitated by informational asymmetries. In particular, the financial literature describes three problems that give rise to under-investment processes: 1) Asset Substitution arises from the conflict between shareholders and bondholders, since shareholders are encouraged to undertake riskier investment projects than those initially proposed when asking bondholders for funds (Jensen and Meckling, 1976). 2) Moral Hazard also stems from the shareholder-bondholder conflict, given the incentives of shareholders not to undertake or abandon profitable projects whenever their net present value is lower than the amount of debt issued (Myers, 1977).

Finally, the adverse selection problem has a double source: the conflict between bondholders and shareholders (Stiglitz and Weiss, 1981), and the conflict between current and prospective shareholders (Myers and Majluf, 1984), since neither bondholders nor prospective shareholders have enough information on the quality of the investment project proposed by the firm when asking for funds. In summary, all these problems create a cost disadvantage of external finance that may give rise to financing constraints in capital markets and, consequently, that may lead firms to forgo positive NPV when there are no internal funds available.

Alternatively, over-investment processes arising from the misalignment of the interest of owners and managers may also explain the dependence of investment on internal funds. When ownership and control are separated, managers have great discretion in the decision-making process and, rather paying out dividends to shareholders, they prefer to use cash flow to maximize their personal wealth. Consequently, as pointed out by Jensen (1986), managers have incentives to use the firm's free cash flows to undertake negative NPV projects, while this would not happen if they had to raise external capital at a higher cost.

It follows from Jensen's (1986) free cash flows hypothesis that debt is valuable for firms with large cash flows and few growth opportunities, because it commits managers to pay out cash in the future, thereby reducing the “free” cash flow at their disposal for empire-building investment (and other expenditures which increase managerial private benefits at the expense of firm profits). Thus, debt will tend to

4 This reasoning is explained in section 2.6.3 Agency Cost of Debt.
control for opportunistic behaviour, as debt provides a means of bonding managers' promises to pay out future cash flows rather than investing in wealth-destroying ventures. Therefore, debt in this case, mitigates the over-investment problem, and thus, should be positively related to high firm performance. In summary, using Jensen’s (1986) argument, a direct link between firms’ performance and debt can be seen. Therefore, this project will empirically test the Jensen model to ascertain whether a certain level of gearing can lead to a better firm performance. However, before we proceed to the empirical tests employed, firms’ performance will be discussed in the next chapter.
Chapter 3

Financial Performance Measures

3.1 Introduction

Corporate managers now face a period in which a new economic framework that better reflects value and profitability must be implemented in their companies. Many managers, some times, are confounded by the conflicting messages the markets appear to send them, improvements in the reported financial performance of a company can be followed by a sharp fall in the price of its shares. By contrast, results that moderately exceed consensus forecasts can propel its share price to new heights, leaving managers to wonder how they can possibly achieve the superhuman feats the market expects from them. Looking at the academics and practitioners working in the field of business performance measurement, there is little agreement about which methods or models are the important in this field.

This chapter looks at the valuation and financial performance models that are frequently used. Market returns models (stock price or returns based on stock price) are excluded because they impound market expectations about a firm’s future and may not be an efficient contracting parameter because they are driven by many factors beyond the control of the firms executives as explained in Chapter 1. Section one highlights the limitation of ratios analysis in assessing firms’ performance. Section two discusses the Discount Cash Flow models. Section three presents the Residual Income models. Section four compares all the above-mentioned models and highlights their advantages and disadvantages. Section five summarises the main important points that have been mentioned in this chapter and indicates the appropriate models to be used in assessing firm’s performance.
3.2 Financial Ratios

Financial ratios are one of the tools commonly used to evaluate a company’s performance. Generally speaking, financial information relating to the status of a company’s business operations will be reported in the yearly financial statements, and it is the ratio of any two accounting items in the financial statements that compose a financial ratio. The observation and analysis of appropriate financial ratios can serve as a preliminary reference for the diagnosis of the results of business operation.

Moreover, these ratios directly or indirectly demonstrate certain aspects of a company’s operating situation. For example, are funds being used properly? Are profit earnings at an average level? These financial ratios are still extensively used from both academics and practitioners as indicators of firms’ performance. However, as Lev and Sunder (1979) comment in their comprehensive review:

*It appears that the extensive use of financial ratios by both practitioners and researchers is often motivated by tradition and convenience rather than resulting from theoretical considerations or from careful statistical analysis.*

Indeed, looking at ratios, such as Return on Equity (ROE) or Return on Assets (ROA), the reader can notice that ROE mixes operating performance with financial structure, making peer group analysis or trend analysis less focused. ROA is inadequate because it includes a number of inconsistencies between the numerator and denominator. Non-interest-bearing liabilities are typically not deducted from the denominator, total assets. Yet, the implicit financing cost of these liabilities is included in the expenses of the company and, therefore, deducted from the numerator.

The profit stream in Return on Capital Employed (ROCE) is not properly related to the investment that produced it. In addition, as Aggrawal (2001) points out, if the incentives facing managers and employees are based on accounting measures of profit and do not reflect the cost of capital the firm uses in generating those profits, such capital is unlikely to be used most efficiently. In addition, Copeland et al (2000, 1994) stated that the main weakness of ratios, such as ROA, is that they do not consider the
investment required to generate earnings or its timing. Keen (1999) points out that earnings per share (EPS) tells nothing about the cost of generating those profits. If the cost of capital (loans, bonds, equity) is, say 15\% then a 14\% earning is actually a reduction, not a gain, in economic value. ROA, as Keen (1999) adds, is a more realistic measure of economic performance, but it ignores the cost of capital. For instance, IBM's return on assets was over 11\% but its cost of capital was almost 13\%.

The above arguments have led managers to think about profit calculated by the excess of the rate of return over the cost of capital i.e. economic profit. The concept of economic profit dates back at least to the economist Alfred Marshall who wrote in 1890: “what remains of his (the owner or manager) profits after deducting interest on his capital at the current rate may be called his earnings of undertaking or management”. Marshall is saying that the value created by a company in any time period (its economic profit) must take into account not only the expenses recorded in its accounting records but also the opportunity cost of capital employed in the business. Accordingly, the above-mentioned ratios ignore the cost of capital and fail to measure value creation. In such cases, these ratios may appear to overstate returns; their internal use for measuring and motivating managerial and business unit performance may even lead to the destruction of shareholders' value. As the purpose of this study is to look at a model that can capture company performance, such ratios will not be sufficient and will not be used in this project.

3.3 Discounted Cash Flow Model (DCF): The Free Cash Flow Model

The dividend discount model (DDM) of Williams (1938) provides the basis for most equity valuation models. When investors buy stocks, they expect to receive two types of cash flow: dividend in the period during which the stock is owned, and the expected sale price at the end of the period. In the extreme example, the investor keeps the stock until the company is liquidated; in which case, the liquidating dividend becomes the sale price. Under the assumption of an infinite time horizon, the DDM can be expressed as:
Where $V$ is equal to the present value of all expected future dividends; $\text{DIV}$ discounted at the firm’s cost of equity capital $r_e$, which is generally assumed constant through time. According to Herz et al (2001), the key ingredients necessary to apply the DDM are dividend forecasts and estimated $r_e$. Lee (1996) points out that more than 25% of firms listed on the New York Stock Exchange do not pay any dividends at all. When firms do pay dividends, the amount is discretionary and often does not reflect current firm prospects. Indeed, under the DDM, dividends are treated as the distribution rather than the creation of wealth. Penman (1992) describes it as the dividend conundrum; ‘price is based on future dividends but observed dividends do not tell us anything about price’. These practical constraints greatly limit the usefulness of DDM. As a result, alternative forms of the DDM emerged with the goal of improved practical implementation. The most commonly used model is the DCF model because, as noted by Herz et al (2001), of its direct link to finance theories of Modigliani and Miller (1958).

The DCF model can be found in most financial management textbooks (Copeland et al (2000) for example). Most specifications of the DCF model require estimates of free cash flow (FCF). Free cash flow is the cash flow available for distribution to a defined set of capital providers after all operating and investing needs of the firm are met. Hence, FCF according to Copeland et al (2000) is equal to the Net Operating Profit Less Adjusted Taxes (NOPLAT) minus Net Investment. Although the DCF model has many variants, FCF in the most commonly applied version, is defined as the cash flow available for distribution to both debt and equity holders, and the discount rate is the weighted average cost of capital (WACC). This model estimates the value of the combined debt and equity of the firm; the market value of the firm’s debt net of the firm’s excess cash must be subtracted from the total value of the firm to obtain the value of the equity.

Copeland et al (2000, 1994) state that a company’s expected cash flow could be separated into two time periods and the company’s value defined as follows:

\[
V = \sum_{t=1}^{\infty} \frac{E_t(\text{DIV}_t)}{(1+r_e)}
\]
"Value = Present value of cash flow during explicit forecast period +
Present value of cash flow after explicit forecast period".

The second term in this equation is the continuing value. It is the value of the company’s expected cash flow beyond the explicit forecast period. They mention that by using simplifying assumptions about the company’s performance during this period, for example assuming constant growth rate, permits estimation of continuing value with one of several formulas. Using a continuing value formula eliminates the need to forecast in detail the company’s cash flow over an extended period. A high-quality estimate of continuing value is essential to any valuation, because continuing value often accounts for a large percentage of the total value of the company. Although these continuing values are large, this does not mean that most of a company’s value will be realised in the continuing value period. It often just means that the cash inflow in the early years is offset by outflow for capital spending and working capital investment, investments that should generate higher cash flow in later years. Copeland et al (2000, 1994) mention three techniques to calculate the continuing value which are as follows:

1- **Explicit forecast for long period of time:** One approach to continuing value is to avoid it altogether by carrying out the explicit forecast for a very long period of time (75 or more years) so that any value beyond the explicit forecast would be insignificantly small. Since such a forecast is unlikely to be very detailed, the two formulas below often work just as well with less effort.

2- **Growing free cash-flow perpetuity formula:** According to Copeland et al (2000, 1994), the growing free cash flow perpetuity formula assumes that the company’s free cash flow will grow at a constant rate during the continuing value period using the following formula:

\[
\text{Continuing value (C V)} = \frac{\text{FCF}_{T+1}}{\text{WACC} - g}
\]

Where: \(\text{FCF}_{T+1}\) = the normalised level of free cash flow in the first year after the explicit forecast period.

\(\text{WACC}\) = the weighted average cost of capital.
g = the expected growth rate in free cash flow in perpetuity.

This technique provides the same result as a long explicit forecast when the company’s free cash flow is forecasted to grow at the same rate. This formula is only valid if “g” is less than WACC and is easily misused. Copeland et al (2000, 1994) stress the importance of correctly estimating the normalised level of free cash flow that is consistent with the growth rate that is being forecasted. For example, if growth in the continuing value period is forecasted to be less than the growth in the explicit forecast period (as is normally the case) then the proportion of NOPLAT$^5$ that needs to be invested to achieve growth is likely to be less as well. Hence, in the continuing value period more of each dollar of NOPLAT becomes free cash flow available for the investors. If this transition is not taken into consideration, the continuing value could be significantly understated.

3- **Value-driver Formula**: Copeland et al (2000, 1994) argue that the third technique expresses the growing free cash flow perpetuity formula in terms of the value drivers, ROIC and Growth, as follows:

$$CV = \frac{NOPLAT_{t+1} (1 - \frac{g}{ROIC})}{WACC - g}$$

(see Appendix 3.A for the derivation of this formula)

Where: $NOPLAT_{t+1} =$ the normalised level of NOPLAT in the first year after the explicit forecast period,

g = the expected growth rate in NOPLAT in perpetuity,

ROIC = the expected rate of return on net new investment.

A variation of the value-driver formula is the two-stage value-driver formula. This formula allows the company to break up the continuing value period into two periods with different growth and ROIC assumptions. For example, it might be assumed that during the first eight years after the explicit forecast period the company would grow

$^5$ NOPLAT is the net operating profit less adjusted taxes, and represents the after tax operating profits of the company after adjusting the taxes to a cash basis.
at 8 percent per year and earn an incremental ROIC of 15 percent. After those eight years, the company’s growth would slow to 5 percent and incremental ROIC would drop to 11 percent.

\[
CV = \left[ \frac{NOPLAT_t\left(1 - \frac{g_A}{ROIC_A}\right)}{WACC - g_A} \right] \left[1 - \left(\frac{1 + g_A}{1 + WACC}\right)^{v-1}\right] + \left[ \frac{NOPLAT_{t+1}(1 + g_A)^{v-1}\left(1 - \frac{g_B}{ROIC_B}\right)}{(WACC - g_B)(1 + WACC)^{v-1}} \right]
\]

Where: CV = continuing value.

N = the number of years in the first stage of the continuing value period.

g_A = the expected growth rate in the first stage of the CV period.

g_B = the expected growth in the second stage of the CV period.

ROIC_A = the expected ROIC during the first stage of the CV period.

ROIC_B = the expected ROIC during the second stage of the CV period.

Note that: g_B must be less than WACC for this formula to be valid.

Copeland et al (2000, 1994) state that they generally use the value-driver formula, because it is easier than developing a 75 year projection and is not as easy to misuse as the growing FCF perpetuity. In addition, it forces the company to think about the value drivers explicitly in estimating the continuing value.

Copeland et al (2000, 1994) argue that the explicit forecast period should be long enough so that the business will have reached a steady state of operations by the end of the period. This is because any continuing value approach relies on the following key assumptions:

- The company earns constant margins, maintains a constant capital turnover, and henceforth earns a constant rate of return on existing invested capital.
• The company grows at a constant rate and invests the same proportion of its gross cash flow in its business each year.

• The company earns a constant return on all new investments

The business must be operating at an equilibrium level for the continuing value approaches to be useful. The forecast period should be extended until the free cash flow growth becomes constant. In other words, select longer rather than shorter explicit forecast periods. It is rare for the explicit-forecast period to be shorter than seven years. Furthermore, the explicit-forecast period should not be determined by the company’s internal planning period. Just because the company forecasts only three years ahead does not justify a three-year forecast period for purposes of valuation. A rough forecast beyond three years is better than no forecast. Finally, the continuing value should be discounted to the present value at the WACC before it can be added to the present value of the explicit free cash flow Copeland et al (2000, 1994).

3.4 Discounted Cash Flow: The Economic Profit Model

Economic Profit\(^6\) (EP), according to Copeland et al (2000), measures the value created in a company in a single period of time and is defined as follows:

\[
EP = Invested \text{ Capital} \times (\text{ROIC} - \text{WACC})
\]

EP translates the value drivers, ROIC and growth, into a single dollar figure (growth is ultimately related to the amount of invested capital or size of the company). With the EP approach, continuing value does not represent value of the company after the explicit forecast period; instead, it represents incremental value over the company’s invested capital at the end of the explicit forecast period. According to Copeland et al (1994), the total value of the company is as follows:

---
\(^6\) EP model is another version of residual income model and conceptually is similar to EVA, Ohlson (1995) and Frankel and Lee (1998) model but under different name.
Value = Invested capital at beginning of forecast + present value of forecasted economic profit during explicit forecast period + present value of forecasted economic profit after explicit forecast period.

The recommended continuing value (CV) for economic profit formula according to Copeland et al (2000, 1994) is as follows:

$$CV = \frac{\text{Economic Profit}_{t+1}}{WACC} + \frac{(\text{NOPLAT}_{t+1}) \left( \frac{g}{\text{ROIC}} \right) (\text{ROIC} - \text{WACC})}{WACC(WACC - g)}$$

Where: Economic Profit\textit{t+1} = economic profit in the first year after the explicit forecast period.

NOPLAT\textit{t+1} = NOPLAT in the first year after the explicit forecast period.

\(g\) = the expected growth rate in NOPLAT in perpetuity.

\(\text{ROIC}\) = the expected rate of return on net new investment.

\(\text{WACC}\) = the weighted average cost of capital.

This formula says that the value of economic profit after the explicit forecast equals the present value of economic profit in the first year after the explicit forecast in perpetuity, plus any incremental economic profit after that year created by additional growth at returns exceeding the cost of capital. If expected \(\text{ROIC} = \text{WACC}\), the second half of the equation equals zero, and the continuing economic profit value is the value of the first year’s economic profit in perpetuity. The forecasting horizon should be set the same as in the DCF model; i.e. not less than seven years.

Copeland et al (2000) claim that a manager who is interested in maximising share value should use DCF analysis i.e. FCF and EP models, not earnings per share, to make decisions. They also claim that the DCF models are conceptually superior to the accounting models (discussed in remaining section). They also reflect how the stock market actually behaves and they claim that substantial evidence supports the view that the market takes a sophisticated approach such as DCF. They grouped their evidence into three classes:
1- Accounting earnings are not well correlated with share price

2- Earnings ‘window-dressing’ does not improve share price

3- The market evaluates management decisions based on their expected long-term cash flow impact, not their short-term earnings impact.

The claimed conceptual advantages of the DCF method are based on its corporate finance roots that emphasise cash flows (Brealey and Myers, 2000). Practical valuation “handbooks” such as Copeland et al, (2000, 1994) maintain that cash flows dominate accounting earnings for valuation purposes and thus advocate the DCF model over accounting-based models i.e. the RI model. Lee (1996) argues that valuation models based on discounted future earnings and cash flows have shortcomings. They typically ignore much of the information contained in the balance sheet by ascribing all of a firm’s value to its future earnings (cash flow) stream.

In effect, the DCF method pushes the portion of firm value in the balance sheet into future projections of cash flows (or earnings). This causes a much greater proportion of the firm value to appear in later periods of the forecast. As a result, DCF valuations tend to be plagued by significant practical problems associated with terminal value estimations. These terminal values are higher and more volatile than they need to be because a large portion of the projected cash flow pertains to the current capital base.

3.5 Residual Income: The Economic Value Added Model

The concept of Economic Value Added (EVA) is well established in financial theory, but only recently has the term moved into the mainstream of corporate finance, it is becoming increasingly popular for measuring and maximizing shareholder wealth, as more and more firms adopt it as the base for business planning and performance monitoring (Keen 1999). According to Stewart & Co. (2000), EVA is net operating
profit minus an appropriate charge for the opportunity cost of all capital invested in an enterprise:

\[
EVA = \text{Net Operating Profit After Taxes (NOPAT)} - \left[ \text{Capital} \times \text{the Cost of Capital} \right]
\]

Stewart & Co. (2002)

As such, EVA is an estimate of true "economic" profit, or the amount by which earnings exceed or fall short of the required minimum rate of return, which shareholders and lenders could get by investing in other securities of comparable risk. Accordingly, EVA is the after-tax cash flow generated by a business minus the cost of the capital it has deployed to generate that cash flow. Ferguson and Leistikow (1998), point that EVA is earnings net of a capital charge based on the firm's cost of capital and net asset value. Net asset value is the value of the firm's assets (book value), as opposed to the market value of the firm's business.

The calculation of EVA usually involves a number of adjustments to accounting data. Bacidore et al (1997) point out that because of variety of accounting distortions, the total asset value on the typical balance sheet does not accurately represent either the liquidation value or the replacement-cost value of the assets in place. Therefore, it is of limited use for firm valuation purpose and must be transformed. The proponents of EVA, most notably Stern Stewart, are careful to adjust this accounting balance sheet before arriving at an estimate of the value of the firm's assets in place. The adjustments include netting the non-interest-bearing current liabilities against the current assets, adding back to equity the gross goodwill (i.e. adding cumulative amortised goodwill back to total assets), other write-offs, capitalised value of Research and Development (R&D) (and possibly advertising), LIFO (Last In First Out) reserve, and so forth. The debt balance is increased by the capitalised value of operating lease payments. The goal is to produce an adjustment balance sheet that reflects the economic values of assets in place more accurately than the inherently conservative, historical-cost-based balance sheet, guided by generally accepted accounting principles. Please refer to appendix 3.B to see the two tables that represent the balance sheet after and before the adjustments have been made.

Stewart & Co. (2002) argue that EVA is the financial measure that comes closer than any other to capturing the true economic profit of an enterprise. Stewart & Co add
that EVA is also the performance measure most directly linked to the creation of shareholder wealth over time.

"The best practical periodic performance measure is EVA; forget EPS, ROE and ROI. EVA is what drives stock prices. It stands well out from the crowd as the single best measure of wealth creation on a contemporaneous basis and EVA is almost 50% better than its closest accounting-based competitor in explaining changes in shareholder wealth".

(Stern et al 1995)

Companies that have adopted EVA for performance measurement and/or incentive compensation include AT & T, Coca-Cola, Eli Lilly, George Pacific, Polaroid, Quaker Oats, Sprint, Teledyne and Tenneco (Biddle et al, 1997).

According to McDaniel et al (1988), EVA differs from the conventional financial accounting tools, such as return on investment (ROI), in three ways:

1. EVA is not bound by the generally acceptable accounting principles (GAAP) as laid out by the Financial Accounting Standards Board (FASB). Users can customise the framework to their specific business conditions. For example, many companies consider training to be an investment. The EVA framework can be modified to show training as an investment rather than an expense item.

2. EVA can support every decision throughout a company, ranging from capital investment, to employee compensation, to business unit performance.

3. Finally, the simple yet robust structure of EVA allows its use by engineering, environmental and other personnel as a common tool for communicating different aspects of financial performance.

EVA, like other residual income models, has an advantage over earnings in taking into account the cost of capital consumed, which assists this study in elucidating any relationship between the operating performance of the company and its capital structure. However, its major disadvantage when compared to free cash flow is that it is still based on historical asset values and only measures a single period in time.
Despite adjusting the asset value, there is a serious danger that the adjustment fails to represent its fair value. This short-termism, according to O’Hanlon and Peasnell (1998), can manifest itself in earnings management games.

There is also the more serious danger that the failure of the accounting system to reflect economic reality (i.e., its fair value) might cause the managers to run the business without proper regard to the long term. While these deficiencies are encountered with other accounting-based performance measures, the case for EVA is diminished if it is similarly plagued. Much of Stern Stewart’s effort has been devoted to addressing this problem. O’Hanlon and Peasnell (1998) demonstrate that, in general, the owners’ excess gain during period $t$ is not equal to that period’s residual income. They add that, a sufficient condition for a single-period residual income to equal single-period excess return is that the book value is equal to economic value, or equivalently that unrecorded goodwill is zero. The answer to this problem is to adjust the accounting book value of the entity such that it is equal to the entity’s economic value.

Aggarwal (2001) argues that any new performance measurement system must be able to balance the need for managerial compensation to reflect factors that managers can influence, with the need to respond to capital market signals. Although changes in market value theoretically equal changes in the present value of future EVA, in practice they can be quite noisy, as changes in stock prices often reflect market-wide changes that may have little to do with any given firm. Thus, it is often difficult to relate changes in EVA measured within the company to changes in the market value of equity. Moreover, EVA focuses on the efficient use of capital. This is certainly very important for manufacturing and many other traditional industries. However, most firms now operate in a new era dominated by service, high technology and knowledge-based work. Business success is generally determined less by physical capital and more by efficient development and deployment of intangible human and intellectual capital. It may be difficult to modify an EVA system based on traditional accounting data so that it is optimal for such firms.

Bacidore et al (1997) state that financiers must earn at least their opportunity cost of capital on the invested capital. This condition implies that this cost of capital needs subtraction from operating profits to gauge the firm’s financial performance. For that
reason EVA defines net operating profit after tax (NOPAT) and subtracts a capital charge for the economic book value of assets in place; this is the measure of the capital provided to the firm by its financiers. However, does this amount truly represent the capital used to generate this operating profit? They believe not; the creation of a true “operating” surplus for the financiers in a given period, the operating profit at the end of the period must exceed a capital charge that based on the total market value of the capital used at the beginning of the period, not simply the economic book value of its assets in place. The total market value of the firm, not simply the economic book value of the assets in place, represents the capital commitment of the firm’s financiers.

For that reason, Bacidore et al (1997) modified the EVA measurement and defined a new performance measure, a Refinement of EVA (REVA). The motivation for the refinement of EVA stems partly from EVA’s use of the economic book value of assets while the capital charge for the firm derives from a market-based WACC. The REVA for a given time $t$ is:

$$REVA_t = NOPAT_t - k_w (MVA_{t-1})$$

Bacidore et al (1997)

Where $NOPAT_t$ is the firm’s NOPAT at the end of period $t$ and $MVA_{t-1}$ is the total market value of the firm’s assets at the end of period $t-1$ (beginning of period $t$). $k_w$ is the weighted-average cost of capital and $MVA_{t-1}$ is given by the market value of the firm’s equity plus the book value of the firm’s total debt, less non-interest-bearing current liabilities, all at the end of period $t-1$. However, REVA is criticised for failing to be consistent with financial theory. For further discussion refer to appendix 3.C.

### 3.6 Residual Income: The Ohlson (1995) Model

Ohlson’s (1995) valuation model defines stock prices as a direct function of both earnings and book values. The model includes the bottom-line items in the balance sheet and income statement “book value and earnings”, and its format requires the
change in book value to equal earnings minus dividends (net of capital contributions). He refers to this relation as the "clean surplus relation" because all changes in net assets unrelated to dividends or new equity paid in must pass through the income statement. The analysis starts from the assumption that value equals the present value of expected dividends. Then the clean surplus relation can replace dividends with earnings and book values in the dividend discount formula. Assumptions about the stochastic behaviour of the accounting data then lead to a multiple-data uncertainty model such that earnings and book value act as complementary value indicators. Specifically, the core of the valuation function expresses value as a weighted average of (i) capitalised current earnings (adjusted for dividends) and (ii) current book value. Extreme parameterisations of the model yield either (i) or (ii) as the sole value indicators. The combination is of conceptual interest because it brings both the bottom-line items into valuation through the clean surplus relation.

Ohlson (1995) points out discounted abnormal earnings represent the difference between market and book values; they signify goodwill. In fact, a straightforward two-step procedure derives a particularly parsimonious expression for goodwill as it relates to abnormal earnings. Firstly, following Peasnell (1982) and others, the clean surplus relation implies that goodwill equals the present value of future expected abnormal earnings. Secondly, according to Ohlson, if one assumes that abnormal earnings obey an auto-regressive process, then it follows that goodwill equals the current abnormal earnings scaled by a (positive) constant. The result highlights that one can derive value by assuming abnormal earnings processes that do not refer to past or future expected dividends.

Two closely related Modigliani and Miller (MM) (1958, 1961) properties are satisfied: dividends displace market value on a dollar-for-dollar basis, so that dividend payment irrelevancy applies. The model accordingly separates the creation of wealth from the distribution of wealth. Given the importance generally attached to MM properties in valuation analysis, the requirement that dividends reduce book value but not current earnings enhances the economic significance of owners’ equity accounting (Ohlson 1995). The theory rests directly on the clean surplus relation and the feature that dividends reduce book value but leave current earnings unaffected. Three straightforward analytical assumptions formulate the valuation model:
1- Price is equal to the present value of expected dividends:

\[ P_t = \sum_{r=t}^{\infty} R_f^{-r} E_t[\tilde{d}_{t+r}] \]  

(PVED)

Where:

- \( P_t \) = the market value, or price, of the firm’s equity at date \( t \).
- \( d_t \) = Net dividends paid at date \( t \).
- \( R_f \) = the risk-free rate plus one.
- \( E_t[\cdot] \) = the expected value operator conditioned on the date \( t \) information.

2- Accounting data and dividends satisfy the clean surplus relation, and dividends reduce book value without affecting current earnings. Thus, the clean surplus relation is:

\[ y_{t-1} = y_t + d_t - x_t \]  

(A2)

- \( x_t \) = Earnings for the period \((t-1, t)\)
- \( y_t \) = (Net) book value at date \( t \).

The clean surplus relation (A2) can be applied to express \( P_t \) in terms of future (expected) earnings and book value in lieu of the sequence of (expected) dividends in the PVED formula, and “residual income” or “abnormal earnings” as follows:

\[ x_t^a = x_t - (R_f - 1)y_{t-1} \]

Where \( x_t^a \) are abnormal earnings for period \( t \).

Combined with the clean surplus restriction (A2), the definition implies

\[ d_t = x_t^a - y_t + R_f y_{t-1} \]

(See Appendix 3.1.D for the derivation of this formula)

Using this expression to replace \( d_{t+1}, d_{t+2}, \ldots \) In the PVED formula yields the equation:
\[ P_t = y_t + \sum_{r=1}^{\infty} R_f^{-r} E_t \tilde{x}_{t+r} \]  (1)

(See appendix 3.2. D for the derivation of this formula)

So, the price can be expressed as the sum of book value and the present value of future abnormal earnings:

\[ P_t = y_t + \sum_{r=1}^{\infty} R_f^{-r} E_t \tilde{x}_{t+r} \] or \[ P_t = y_t + \sum_{r=1}^{\infty} E_t \tilde{x}_{t+r} (1 + r)^{-r} \].

Ohlson refers to \( x_t^a \) as abnormal earnings.

Ohlson’s third assumption is that abnormal earnings satisfy the following modified autoregressive process:

\[ \tilde{x}_{t+1} = \omega x_t^a + v_t + \tilde{\epsilon}_{t+1} \]  (2a)

\[ \tilde{v}_{t+1} = +p v_t + \tilde{\epsilon}_{2t+1} \]  (2b)

\( v_t \) = Information other than abnormal earnings, the disturbance terms, \( \epsilon_{2t}, \epsilon_{2t+1}, \tau \geq 1 \), are unpredictable, zero-mean, variables; that is, \( E_t [\tilde{\epsilon}_{t+1}] = 0 \), \( k = 1,2 \) and \( \tau \geq 1 \).

\( \omega \) and \( \gamma \) are fixed and known (in loose terms, a firm’s economic environment and its accounting principles determine the exogenous parameters \( \omega, \gamma \)).

Combining the residual income valuation model in Equation (1) with the information dynamics in Equations (2a) and (2b) yields the following valuation function:

\[ P_t = y_t + \alpha_1 y_t^a + \alpha_2 v_t \]  (3)

Where: \( \alpha_1 = \frac{\omega}{(R_f - \omega)} \geq 0 \), \( \alpha_2 = \frac{R_f}{(R_f - \omega)(R_f - \gamma)} \).

This valuation function does not require explicit forecasts of future dividends, nor does it require additional assumptions about the computation of “terminal value”. The
information dynamics in Equations (2a) and (2b) along with the valuation function in Equation (3) embody the original empirical implications of Ohlson (1995).

Equation (3) represents the Ohlson valuation model, which has been attracting the attention of many researchers. The ability of this model to predict the stock price will be discussed after examining the other version of residual income, namely the Frankel and Lee model and EVA.


The valuation method that is used by Frankel and Lee (1998) uses a discounted residual income approach. They started by using a stock’s fundamental value, defined as the present value of its expected future dividends based on all currently available information:

\[
V_t^* = \sum_{i=1}^{\infty} \frac{E_t(D_{t+i})}{(1 + r_e)^i}
\]  

(1)

In this definition, \(V_t^*\) is the stock’s fundamental value at time \(t\),

\(E_t(D_{t+i})\) is the expected future dividend for period \(t+i\) conditional on information available at time \(t\),

\(r_e\) is the cost of equity capital based on the information set at time \(t\). This definition assumes a flat term-structure of discount rates.

Frankel and Lee (1998) state that it is easy to show that, as long as a firm’s earnings and book value are forecasted in a manner consistent with clean surplus accounting, Equation (1) can be rewritten as the reported book value, plus an infinite sum of discounted residual income:

\[
V_t^* = B_t + \sum_{i=1}^{\infty} \frac{E_t[NI_{t+i} - (r_e B_{t+i})]}{(1 + r_e)^i}
\]  

(2)
If we substitute $NI = ROE_{t+i} \times B_{t+i-1}$ in the above Equation (2) then:

$$V_i^* = B_i + \sum_{t=1}^{\infty} E_t \left( \frac{(ROE_{t+i} - r_e)B_{t+i-1}}{(1 + r_e)} \right)$$

(3)

Where $B_i$ is the book value at time $t$, $E_t[.]$ is expectation function based on information available at time $t$, $NI_{t+i}$ is the Net Income for period $t+i$, $r_e$ is the cost of equity capital and $ROE_{t+i}$ is the after-tax return on book equity for period $t+i$.

Note that this equation is equivalent to a dividend discount model, but expresses firm value in terms of accounting numbers. Therefore, it relies on the same theory and is subject to the same theoretical limitations as the dividend discount model. However, the model provides a framework for analysing the relation between accounting numbers and firm value and as will be explained later, relies on clean surplus relation (CSR).

Frankel and Lee (1998) pointed out that Equation (3) shows that equity value splits into two components, an accounting measure of the capital invested ($B_i$), and a measure of the present value of future residual income. If a firm earns future accounting income at a rate exactly equal to its cost of equity capital, then the present value of future residual income is zero, and $V_i = B_i$. In other words, firms that neither create nor destroy wealth relative to their accounting-based shareholders’ equity will be worth only their current book value. However, firms whose expected ROEs are higher (lower) than $r_e$ will have values greater (lesser) than their book values.

Frankel and Lee (1998) argue that Equation (3) presents a simple procedure for estimating a firm’s intrinsic value ($V_i^*$). The four parameters that are needed for the estimation are:

- The cost of equity capital ($r_e$)
- Future ROE forecasts ($FROE_t$)
- Current book value ($B_i$)
Dividend payout ratio \( (k) \)

The first three parameters' roles are readily seen in Equation (3). The last input, the dividend payout ratio \( (k) \), is used in conjunction with the \( (CSR) \) to derive future book values. In the following section, the specifics of the model estimation procedure will be discussed.

Cost of equity \( (r_e) \): In theory, \( r_e \) should be firm specific, reflecting the premium demanded by equity investors to invest in a firm or project of comparable risk. However, there is little consensus in practice on how this discount rate should be determined. Frankel and Lee (1998) use three different approaches: a constant discount rate and two industry-based discount rates derived by Fama and French (FF) (1995). The FF discount rates are based on a one-factor and a three-factor risk model.

Dividend payout ratio \( (k) \): The dividend payout ratio is the percentage of net income paid out in the form of dividends each period. Frankel and Lee (1998) obtain a firm-specific estimate of \( k \) by dividing the common stock dividends paid in the most recent year by net income before extraordinary items. For firms with negative earnings, they divide dividends by six percent of total assets to derive an estimated payout ratio. Six percent, they argue, reflects the average long run return-on-assets. They use this measure as a proxy for normal earnings when reported earnings are negative. This variable is used, in conjunction with \( B_t \), to derive forecasted book values

\[
B_{t+1} = B_t + NI_{t+1} - d_{t+1} = B_t + (1-k)NI_{t+1} = [1 + (1-k)ROE_{t+1}]B_t
\]

Since \( d_{t+1} = NI_{t+1} \times K \) and \( NI_{t+1} = ROE_{t+1} \times B_t \)

Analogously, all future book values can be expressed as functions of \( B_t, k, \) and \( (FROE_t) \). For example, the following formula can be written:

\[
B_{t+2} = [1 + (1-k)ROE_{t+1}] [1 + (1-k)ROE_{t+2}]B_t
\]

\( (FROE_t) \): The most important and difficult task in the discounted residual income valuation exercise is forecasting \( (FROE_t) \) (or, its equivalent, forecasting future
earnings). According to Frankel and Lee (1998) two alternatives, based on ex ante information, are: (1) use prior period earnings or (2) use analysts’ earnings forecasts. They use both methods and derive a value metric based on historical earnings ($V_h$) as well as a value metric based on consensus I/B/E/S analyst forecasts ($V_f$). Fairfield (1994) shows that, in large samples, the correlation between current year ROE, and next year’s ROE, is around 0.66, suggesting that the current period ROE, is a reasonable starting point for estimating ($FROE_i$). Prior studies, such as O’Brien (1988) and Brown et al (1987), show that analyst earnings forecasts are superior to time-series forecasts. However, the predictive superiority of an analyst-based value metric ($V_f$) over a historical-based value metric ($V_h$) is an open empirical question.

Forecast horizons and terminal value estimation: Equation (3) expresses a firm’s value in terms of an infinite series, but for practical purpose as they point out, the explicit forecast period must be finite. This limitation necessitates a terminal value estimate; an estimate of the value of the firm based on residual income earned after the explicit forecasting period. They add that one approach is to estimate the terminal value by first expanding Equation (3) to $T$ terms and then taking the next term in the expansion as perpetuity. For example, if the explicit forecast period ends after $T$ period, the terminal value is:

$$\frac{(ROE_1 - r_e)}{(1 + r_e)^T} B_t$$

This procedure is mathematically equivalent to a $T$-period discounted dividend model in which year $T + 1$ earnings is treated as a perpetuity. Therefore, the resulting value estimate depends critically on the particular earnings forecast used in the terminal value. In their study, Frankel and Lee (1998) take a simple approach using short-horizon earnings forecasts of up to three years. They state that in theory, $T$ should be set large enough for firms to reach their competitive equilibrium. However, their ability to forecast ($FROE_i$) diminishes quickly over time and forecasting errors are compounded in longer expansions. Therefore, they estimate three forms of $V_t$:

$$\hat{V}_t = B_t + \frac{(FROE_i - r_e)}{(1 + r_e)} B_t + \frac{(FROE_i - r_e)}{(1 + r_e)^2} B_t + \frac{(FROE_i - r_e)}{(1 + r_e)^3} B_t$$

(3.1)
Equation (3.1) represents a two-period expansion of the residual income model with the forecasted ROE for the current year \( FROE_t \) assumed to be earned in perpetuity. Equation (3.2) also represents a two-period expansion of the model, but they use a two-year-ahead forecasted ROE \( FROE_{t+1} \) in the perpetuity. Similarly, Equation (3.3) is a three-period model.

The right-hand side of each equation consists of ex ante observation. To estimate \( V_n \), they use the return on average book (for reason explained below) equity, 
\[
ROE_i = \frac{NI_i}{B_i + B_{i-1}}
\]
to proxy for all future \( (ROE_i) \) i.e., they substitute \( (ROE_i) \) for all the \( FROE_i \) in the above equations. To estimate \( V_f \), they derive future ROE, and book values from I/B/E/S consensus forecasts using a sequential procedure described below:

Their implementation of the model requires, three future ROE forecasts \( [FROE_t, FROE_{t+1}, FROE_{t+2}] \). They derive future ROE, from I/B/E/S consensus EPS estimates. Since year-end book values are dependent on current year \( ROE_t \), they use a sequential process to estimate future \( ROE_t \). The steps in the process are listed below. Year \( t \) refers to the year of portfolio formation.

Step 1: estimating \( FROE_t \) and \( B_t \). They require that all sample firms have a one-year-ahead I/B/E/S consensus EPS forecast, divided by the average book value per share during year \( t - 1 \). They argue that, use of the average, rather than year-end, book value reduces the chance of an extremely low denominator. Then they use \( FROE_t \) and the dividend payout ratio \( (k) \) to derive the ending book value for year \( t \). Notionally they have:
\[ FROE_i = \frac{FY_1}{(B_{i-1} + B_{i-2})/2}, \]

\[ B_i = B_{i-1}[1 + FROE_i(1-k)]. \]

Where \( FY_1 \) is the one-year ahead consensus forecast of EPS.

Step 2: Estimating \( FROE_{i+1} \) and \( B_{i+1} \). They also require that all sample firms have a two-year-ahead consensus forecast \([FY_2]\). They then compute \( FROE_{i+1} \) and \( B_{i+1} \) analogously:

\[ FROE_{i+1} = \frac{FY_2}{(B_{i+1} + B_{i+2})/2}, \quad B_{i+1} = B_i[1 + FROE_{i+1}(1-k)]. \]

Step 3: Estimating \( FROE_{i+2} \) and \( B_{i+2} \). Where a long-term earnings growth estimate \([Lt_g]\) is available, they compute \( FROE_{i+2} \) and \( B_{i+2} \) as follows:

\[ FROE_{i+2} = \frac{FY_2(1+Lt_g)}{(B_{i+1} + B_{i+2})/2}, \quad B_{i+2} = B_{i+1}[1 + FROE_{i+2}(1-k)]. \]

Where \([Lt_g]\) is not available, they use \( FROE_{i+1} \) to proxy for \( FROE_{i+2} \).
3.8 Comparison between the Models

3.8.1 Theoretical Equivalence between EVA and RI in Valuation

According to Lee (1996), both EVA and RI\textsuperscript{7} (i.e., the Ohlson and Frankel and Lee model) rely on the idea of 'Residual Income', defined as earnings in excess of an expected (benchmark) level of performance, and tied to capital employed. EVA for a given time period is:

\begin{equation}
\text{EVA} = \text{earnings}_t - r \times \text{capital}_{t-1}
\end{equation}

Equation (1) relates wealth creation to the amount of residual income generated. A company’s or division’s activities create wealth (generate positive EVA) if actual earnings exceed the expected dollar return on the capital employed. Stern Stewart computes EVA for all long-term investors, including shareholders and long-term debt holders. On the other hand, RI focuses only on equity investors. In the EVA approach:

\begin{align*}
\text{Earnings}_t &= \text{earnings before interest}_t = \text{EBI}_t, \\
r &= \text{weighted average cost of capital} = \text{(WACC)}, \\
\text{Capital}_t &= \text{total assets}_{t-1} = \text{TA}_{t-1}
\end{align*}

On the other hand, some texts refer to RI as NOPLAT\textsuperscript{7}, net operating profit less adjusted taxes (for instance, see Copeland et al, (2000)). According to Lee (1996), substitution of these definitions yields equation (2a):

\begin{align*}
\text{EVA}_t &= \text{EBI}_t - \text{WACC} \times \text{TA}_{t-1} \\
\text{EVA}_t &= \left(\frac{\text{EBI}_t}{\text{TA}_{t-1}} - \text{WACC}\right) \times \text{TA}_{t-1} \\
\text{EVA}_t &= \left(\frac{\text{ROA}_t}{\text{TA}_{t-1}} - \text{WACC}\right) \times \text{TA}_{t-1}
\end{align*}

ROA is the company’s return on assets. This equation shows that a firm or division is creating wealth for its long-term investors only if its ROA exceeds the cost of capital.

\textsuperscript{7} Although EVA is another form of RI model, to make the comparison easy we denote by RI the Ohlson (1995) and Frankel and Lee (1998) models.
(WACC). The amount of wealth created depends on the amount of capital employed (TA). Moreover, as noted by O’Hanlon and Peasnell (1998), the decomposition of EVA in Equation (2a) into profitability and scale components, highlights shortcomings of conventional accounting-based performance measures and the advantages of the residual income approach. Defining performance in terms of profitability has long been known to result in under-investment, since ROA can be maximised by rejecting positive-NPV projects that dilute ROA. Defining performance in terms of absolute amount of profit can have the opposite effect of encouraging over-investment (acceptance of negative NPV projects) due to failure to take proper account of profitability. By combining spread and scale factors as in Equation (2a), one class of problems created by the use of accounting performance measure can be overcome.

Lee (1996) points out that under RI (for example the Ohlson model) EVA can be defined as economic value added for equity-holders, so the components of residual income can be defined as follows:

\[
\text{Earnings}_t = \text{net income}_t = \text{NI}_t \\
\text{r} = \text{cost of capital (r$_e$)} \\
\text{Capital}_{t-1} = \text{total shareholder’s equity}_{t-1} = \text{book value}_{t-1} = \text{SE}_{t-1} = \text{B}_{t-1}
\]

Substituting these expressions yields Equation 2b:

\[
\text{EVA} = \text{NI}_t \cdot r_e \cdot B_{t-1} \\
= \left( \frac{\text{NI}_t}{\text{B}_{t-1}} - r_e \right) \cdot B_{t-1} \\
= (\text{ROE}_t - r_e) \cdot B_{t-1}
\]

This equation shows a firm is only creating wealth for its shareholders if it earns a return on equity (ROE) in excess of the cost of equity ($r_e$). Moreover, the amount of actual wealth created depends on the amount of equity capital employed. Lee (1996) points out that EVA is a powerful valuation tool when it is extended to multiple periods. Since a firm derives value from both invested capital and future activities, accordingly, the following Equation (3) can be written as follows:

\[
\text{Firm value}_t = \text{capital}_t + \text{present value of all future EVAs} \quad (3)
\]

(Stewart et al 1995)
Equation (3) according to Lee (1996) has engaged researchers from the academic community because it is based directly on the accounting numbers that can be seen on, or forecasted from the financial statements, and holds for any accounting system that satisfies the clean surplus relation. The first term on the right-hand side of Equation (3), capital invested at time t, comes from the balance sheet. The second term, the present value of future EVAs, comes from both the forecasted balance sheet and the income statement and can be viewed as the present value of expected residual earnings in the future. Lee (1996) on his research focuses on the equity-holders, and expressing the variables on par-share basis:

\[
\text{Capital}_t = B_t = \text{book value at time } t
\]
\[
\text{Firm value}_t = '\text{Synthetic Price}' \text{ at time } t = P_t^* 
\]

He uses the term 'Synthetic Price' to denote a firm's intrinsic value per share, based on fundamental analysis. Under this notation, Equation 3 becomes Equation 4:

\[
P_t^* = B_t + \sum_{i=1}^{m} EVA_{it}
\]
\[
= B_t + \left( \frac{ROE_{t+1} - r_e}{1 + r_e} \right) B_t + \left( \frac{ROE_{t+2} - r_e}{(1 + r_e)^2} \right) B_{t+1} + \left( \frac{ROE_{t+3}}{(1 + r_e)^3} \right) B_{t+2} 
\]

Dividing through by \( B_t \), we get Equation 5:

\[
\frac{P_t^*}{B_t} = 1 + \sum_{i=1}^{m} \left( \frac{ROE_{t+i} - r_e}{(1 + r_e)} \right) \frac{B_{t+i}}{B_t} 
\]

Equation (5) is one of the RI formulas appearing in the accounting literature (this formula is shown under the Frankel and Lee (1998) valuation model, based on Ohlson's (1995) model). Equation (5) provides several interesting insight: The price-to-book ratio is expressed in terms of future abnormal ROEs and growth in book value. How the RI model relates to traditional valuation models such as the DDM and DCF is made known in the next section.

### 3.8.2 Theoretical Equivalence between the RI and DCF Models

Although the DCF model has many variants, FCF in the most commonly applied version is defined as the cash flow available for distribution to both debt and equity
holders, and the discount rate is the WACC. This model estimates the value of the sum of the debt and equity of the firm; the market value of the firm’s debt net of the firm’s excess cash must be subtracted from the total value of the firm to obtain the value of equity.

Herz et al (2001) report that because of the inherent difficulty of projecting FCF indefinitely into the future, users of DCF models typically forecast FCF through a specified terminal date, and estimate the terminal value of the net equity plus debt, VT + net debt T, separately:

\[
V_0 + \text{net debt}_0 = \sum_{t=1}^{T} \left( \frac{E_0(FCF_t)}{1+r_{wacc}^t} \right) + \frac{E_0(V_T + \text{net debt}_T)}{1+r_{wacc}^T}
\]

\( r_{wacc} \) represents the WACC that is used as the discount rate. The standard approach to estimating the terminal value of the debt and equity assumes the firm reaches a steady state by the terminal date, and computes the terminal values as a FCF perpetuity growing at a constant g, \( E_0 \left( \frac{FCF_t(1+g)}{r_{wacc}^t - g} \right) \). The key ingredients of the DCF model are forecasts of the FCF and an estimate of \( r_{wacc} \), the WACC depends on the capital structure of the firm and changes whenever the capital structure of the firm changes. Users often assume a constant WACC based on the firm’s target debt to equity ratio (as is the case in Copeland et al (2000, 1994)).

However, the RI approach, introduced by Edwards and Bell (1961) and subsequently further developed by Peasnell (1982) and Ohlson (1995), derives from the DDM with the assumption of clean surplus accounting; RI is a variation of the EVA approach (as we saw in the previous section). RI is accounting net income less a charge for equity capital, equal to the cost of equity capital times the beginning of period book value of equity:

\[
RI_t = NI_t - r_tBV_{t-1}
\]

In the RI model, the market value of equity is estimated as the sum of the firm’s current book value of equity, BV0, plus the present value of the expected future RI.

---

8 Excess cash is defined as the cash or equivalents held by the firm in excess of the amount needed for working capital and investment purposes (Herz et al, 2001).
As with the DCF model, users of the RI model typically forecast RI only through a specified future terminal date, with the difference between the market and book value of equity at the terminal value date, $V_T - BV_T$, estimated separately:

$$V_o = BV_o + \frac{\sum E_o (RI_T)}{(1 + r_e)^T} + \frac{E_o (V_T - BV_T)}{(1 + r_e)^T}$$

Users of this model commonly assume that the firm reaches a steady state by the terminal date. They compute the terminal value (TV) as RI perpetuity growing at a constant growth rate $g$, so $TV = \frac{E_o [RI_T (1 + g)]}{r_e - g}$. The key components for applying the RI are forecasts of BV and accounting net income, and an estimate of $r_e$. In order to forecast book values, net dividends also must be forecast. However, Ohlson (1995) made a restriction on dividends by using linear information parameters as a third assumption and in that case; there is no need to forecast dividends.

Herz et al (2001) point out that because of their theoretical equivalence, the DDM, the DCF, and the RI all provide the same valuation when the flows are projected consistently to infinity and comparable discount rates applied. However, horizons over which the flows can be reasonably projected are limited in practice, and discount rates are estimated with error. These practical considerations cause some academicians and practitioners to prefer one valuation model to another. Indeed, as noted by Plenborg (2002), if the valuation approaches are improperly employed, as with the growth rate used in the terminal value calculation, the models yield different firm value estimates.

### 3.8.3 Differences between the RI and DCF Models

As illustrated, the DDM is generally considered to be the core theoretical model and is the subject of less empirical testing than the DCF method or RI. Both the DCF and RI models have many variants with adherents who perceive practical advantages for one specification over another. Although, as noted by Herz et al (2001), DCF is more popular both in the M.B.A classroom and on Wall Street, RI is gaining acceptance in
both settings. As shown by Damondara (1994), the growth rate does not have to be identical in the two valuation approaches due to the effect of leverage.

The claimed conceptual advantages of the DCF method are based on its finance roots that emphasise cash flows (Brealey and Myers, 2000). Practical valuation “handbooks” such as Copeland et al (1994) favour DCF rather than accounting earnings. Lee (1996) argues that valuation models based on discounted future earnings and cash flows also have shortcomings. They typically ignore much of the information contained in the balance sheet. This is the essence of the DCF method. By ascribing all of a firm’s value to its future earnings (cash flow) stream DCF valuations ignore the value-relevant information in balance sheets. In effect, the DCF method pushes the portion of firm value in the balance sheet into future projections of cash flows (or earnings) this causes a much greater proportion of the firm value to appear in later periods of the forecast. As a result, DCF valuations tend to be plagued by significant practical problems associated with terminal value estimations. These terminal values are higher and more volatile than they need to be because a large portion of the projected cash flow pertains to the current capital base.

Penman and Sougiannis (1998), among others, explain the claimed practical advantages of the RI; in their view, a shortcoming of the DCF model is the need to subtract long-term capital investment from operating cash flows to compute FCF. For growing firms, negative FCF often results for many years (note that these also could be the same problem for EVA). RI adherents maintain that accrual accounting eliminates the distorting effect of capital investment expenditure by placing it on the balance sheet as an asset. However, Lee (1996) argues that the RI model naturally depends on reported accounting numbers. A commonly raised concern is how vulnerable these numbers are to managerial misrepresentation or other biases arising from the accounting system, such as conservatism.

In theory, conservatism should not affect RI valuations. Ohlson shows that, in an infinite-horizon valuation, the mechanics of double-entry bookkeeping under clean surplus accounting automatically allow for conservatism. However, Feltham and Ohlson (1995) show that in finite-horizon valuation the level of accounting conservatism affects the terminal value calculation. On the other hand, Penman and Sougiannis (1998) demonstrate that the RI’s use of accrual accounting allows for
more reasonable valuations than the DCF model from forecasted payouts over relatively short horizons. Academic accountants, such as Bernard (1995), advocate the RI method because of its direct ties to earnings and book values, central concepts in accrual accounting, whereas the DCF model is founded only in finance theory. Another claimed RI advantage, as Herz et al (2001) argue, is that the terminal value is measured as the difference between the market value and the book value of the firm’s equity, \( V_T - BV_T \). This difference is typically smaller than the sum of the firm’s market value and its net debt, \( V_T + \text{net debt}_T \), the terminal value in the DCF model. This is an advantage because the terminal value often constitutes a large percentage of the computed total value of the firm in both models and is the component estimated with the greatest uncertainty; minimising the influence of the terminal value reduces uncertainty in the valuation process.

Clubb (2002) points out that the assumptions required for equivalence of the valuation models are not incorporated into practitioner-oriented valuation methodologies in the financial analysis literature. This suggests that the theoretical equivalence of the models may not be easily captured in practical applications of the models and hence that there is potential merit in research that identifies the assumptions and procedures required for generating equivalence valuation. The main problem in generating a set of equivalent valuations is in the reconciliation of the cost of equity-based models with WACC-based models. Miles and Ezzell (1980) demonstrate that discounting free cash flow at the after tax WACC provides a correct firm valuation (in an MM capital market setting with corporate tax) only if the firm pursues an active debt management policy of maintaining a constant market value leverage ratio.

Lundholm and O’Keefe (2001) assume that the equity valuations based on the cost of equity are correct, and characterise deviations from this valuation resulting from the use of an estimated WACC, as valuation errors due to the use of inconsistent discount rates. This approach has the merit of drawing attention to the impact of inconsistent discount rate assumptions on valuations in the empirical studies, which they review. However, the use of the cost of equity-based valuation as the bench-mark against which to compare WACC based valuation, is much less satisfactory from practical valuation viewpoints. In particular, following Modigliani and Miller (1958, 1963), it has been the conventional wisdom in corporate finance that the value of equity should
be determined by valuing the operations of the business plus the value of any tax
shields arising from debt financing and deducting the value of the debt. For instance,
Copeland et al (1994) advocate thinking in terms of target capital structure for two
reasons; firstly, they argue that at any point a company’s capital structure may not
reflect the capital structure expected to prevail over the life of the business. Secondly,
using a target capital structure solves the problem of circularity involved estimating
the WACC. This circularity arises because the market value weights need elucidation
in order to determine the WACC, but the market value weights cannot be known
without identifying what the market value is in the first place, especially the market
value of equity. To determine the value of equity, which is the objective of the
valuation process itself, the expected free cash flow must be discounted at the WACC.
In essence, the WACC cannot be known without discerning the market value of
equity, and the market value of equity cannot be established without knowing the
WACC.

However, Penman (2001) advocates cost of equity rather than WACC. His argument
is that the cost of equity will vary over time reflecting changes in financial risk linked
to changes in leverage. While the WACC may also change as operations change, the
task of forecasting the discount rate is reduced and an advantage of this approach is
that dividends are irrelevant and financing is irrelevant. Penman and Sougiannis
(1998) and Francis et al (2000) apply a target capital structure and constant cost of
debt and equity, which is consistent with Copeland et al (1994). However, Levin and
Olsson (2000) show that the weights implied by the forecasted debt and equity should
be applied. Ideally, the cost of equity and debt should also be adjusted according to
the capital structure in order to reflect the underlying financial risk (Gregory, 1992).
In a related study, Levin and Olsson (2000) demonstrate that by using the DCF
approach, and disregarding the weights in the WACC formula changing when the debt
ratio (in the market value) changes, leads to bias in firm value estimates.

Plenborg (2002) argues that in the absence of taxes, and as known from the literature,
WACC is constant across different capital structures. This is expected as in those
circumstances WACC is independent of the capital structure. Accordingly, the DCF
approach yields firm value estimates that are theoretically correct. On the other hand,
when financial leverage increases, the cost of equity increases as well. Thus, if the
target capital structure deviates from the implied capital structure in the forecasted financial statements and the cost of equity is adjusted according to the target capital structure, the RI approach yields biased firm value estimates.

3.9 Empirical Research

Many Values Based Management advocates contend that an organisation’s primary objectives must be stated in terms of “economic value” measures, such as EVA, in order to align internal goals with the maximisation of shareholder value (e.g., Copeland et al (1994) and Stern et al (1995)). This contention is based on assertions that changes in economic value measures track changes in shareholder wealth more closely than traditional accounting measures for goal setting, capital budgeting, and compensation purpose (Stern et al, 1995).

Moreover, claims that economic value measures are superior to traditional accounting measures are not limited to consultants and the business press. Compared to such common performance measures as return on capital, return on equity, growth in earnings per share, and growth in cash flow, EVA has the highest statistical correlation with the creation of value for shareholders; EVA drives stock prices (Garvey and Milbourn 2000). Analytical studies by Anctil (1996), Rogerson (1997) and Reichelstein (1997) show how the use of residual income-based measures such as EVA can ensure goal congruence between the principal and agent. Evidence provided by Wallace (1997) suggests that managers compensation based on EVA (instead of earnings) take actions consistence with EVA-based incentives.

However, much of the support for the claimed superiority of economic value measures, as reported by Biddle et al (1997), is based on relatively unsophisticated studies examining the relation between market measures (e.g., market value or shareholder returns) and EVA. Simple univariate tests by Milunovich and Tseui (1996) and Lehn and Makhiya (1997) find market-value added more highly associated with EVA than with accounting returns, earnings per share, free cash flow or free cash growth. O’Byrne (1996) uses regression models to examine the association between
market value and two performance measures: EVA and Net Operating Profit after Tax (NOPAT). Both measures have similar explanatory power when no control variables are included in the regression models, but a modified EVA model has greater explanatory power when industry indicator variables and the logarithm of capital for each firm are included as additional explanatory variables. However, O’Byrne does not make similar adjustments to the NOPAT model, making it impossible to compare results using the different measures.

Furthermore, Chen and Dodd (1997) examine the explanatory power of accounting measures (earnings per share, return on assets and return on equity), residual income and various EVA related measures. Although the EVA measures outperform accounting earnings in explaining stock returns, the earnings measures provide significant incremental explanatory power above EVA. The authors also find the explanatory power of the EVA measures to be far lower than claimed by proponents. Biddle et al (1997) use contemporary capital markets research techniques to examine the power of accounting measures (earnings and operating profits) to explain stock market returns relative to EVA and five components of EVA (cash flow from operations, operating accruals, after-tax interest expenses, capital charge and accounting adjustments). They find that traditional accounting measures generally outperform EVA in explaining stock prices. While the EVA measures’ capital charges and adjustments for accounting 'distortions have some incremental explanatory power over traditional accounting measures, the contribution from these variables is not economically significant in their test.

Wallace (1997) examines relative performance changes in 40 adopters of residual income-based compensation measures such as EVA, and a matched-pairs control sample of firms where incentive compensation continues to be based on traditional accounting earnings (e.g. EPS, operating profit). The results indicate significant increases noted in residual income for the firms adopting residual income-based compensation relative to the control firms. Compared to the control firms, residual income firms decrease new investments and increase dispositions of assets, increase payouts to shareholders through share purchases and utilise assets more intensively, leading to significantly greater change in residual income. He also finds weak evidence that stock market participants respond favourably to the adoption of residual
income-based compensation plans as evidence by increased stock return. Wallace's (1997) study examines changes in performance rather than performance levels, and only examines performance changes over one year.

Hogan and Lewis (1999) extend his study by investigating performance changes over a four-year period, and by matching control firms on past performance to control for possible mean reversion in performance levels. They find that adopters of residual income measures are relatively poor performers prior to the compensation plans' implementation and that the improved stock returns and operating performance reported by Wallace (1997) may not be unique to economic value adopters. After introducing past profitability as an additional matching criterion, they find no significant differences in the stock prices or operating performance of their two groups, and conclude that economic value plans are no better in their ability to create shareholder wealth than traditional plans blending earnings-based bonuses and stock-based compensation.

Ittner and Larcker (2001) argue that perhaps the biggest limitation in the preceding studies is the use of publicly available data on EVA values and uses. Studies of EVA's predictive ability typically employ published EVA data estimated by the consulting firm Stern Stewart. However, these numbers are computed using public financial data, and contain relatively few of the accounting adjustments EVA proponents encourage companies to make to more closely approximate economic profits.

Considerable debate exists over the relative ability of different economic value measures, (EVA, Cash Flow Return on Investment, Cash Value Added or variant of these measures) to predict stock returns, with many consulting firms claiming that their economic value measures are far better indicators of value creation than EVA (Myers, 1996; The Economist, 1996). Evidence on the benefits from tying compensation to EVA is mixed. As discussed earlier, studies by Wallace (1997) and Hogan and Lewis (1999) reach conflicting conclusions regarding the performance of firms adopting residual income-based compensation plans (such as EVA) relative to the performance of control samples. In contrast, Wallace's (1998) survey of EVA users finds that firms using this measure for compensation purposes report greater awareness of the cost of capital reduced average accounts receivable age increased
sales revenues, and a longer accounts payable age than EVA users who do not use the measure for compensation. Given these mixed results, the benefits of EVA-based compensation plans remain an open issue.

The RI and the DCF have received considerable attention in the past decade. Despite the theoretical equivalence between the RI and DCF approaches, the finance literature has argued in favour of the DCF approach for firm valuation since it is unaffected by accounting methods (Copeland et al, 1994). However, as demonstrated by Ohlson (1995), the RI model is insensitive to different accounting methods if clean surplus accounting is applied. Penman and Sougiannis (1998) and Francis et al (2000) examined empirically the accuracy of the RI and DCF models. Both studies find that RI model yields more accurate firm value estimates than the DCF model. However, since both valuation models employ the same theoretical framework, a proper implementation, as reported by Plenborg (2002), would imply that both approaches yield similar firm value estimates. Moreover, Olsson (1998) argues that the introduction of simplifying assumptions occurs when during the implementation of the different valuation approaches. Since simplifying assumptions introduce bias in the firm value estimates, they are likely to affect firm value estimates based on the RI and DCF approaches differently.

Bernard (1995), employing only the first 4 years of forecast data, finds that the RI approach explains 68 percent of a firm's stock price, while the Discount Dividend Model (DDM) explains only 29 percent. Plenborg (1999) finds similar results when comparing the information content of earnings and cash flows. Using Danish data, Plenborg (1999) finds that four years of RI earnings explains 22 percent of the stock price variation in the same measurement period. In comparison, accumulated free cash flows explain less than 1 percent of stock price variation in the same four-year period.

9 RI approach means models that are introduced by Ohlson (1995) and relative's studies such as Frankel and Lee (1998). This model was termed the Edwards-Bell-Ohlson (EBO) by Bernard (1994) and has been called the residual earnings model (Frankel and Lee, 1998). The DCF model can be found in the Copeland et al (1994) finance textbook.

10 Note that the theoretical equivalence between these two approaches is presented in the previous section of this chapter.

11 The introduction of a simplifying assumption implies that the internal coherence between the forecasted financial statements and the valuation approach (including cost of capital) is not intact (Plenborg, 2002).
period. The results of both Bernard and Plenborg indicate that the required forecast period is shorter for the RI approach than for both the DDM and DCF approach.

Penman and Sougiannis (1998) and Francis et al (2000) compare the reliability of firm estimates based on the DDM, RI and DCF approaches respectively. Although both studies use US data, a primary difference between them is that the forecast data are determined differently. Francis et al employ Value Line’s forecast data while Penman and Sougiannis use realised data as estimates of historical forecasts. Although these two studies employ different sources of forecast data, both show that the RI approach yields less biased firm value estimates than the DDM and the DCF approaches. This result is insensitive to different methods for calculating the terminal value. However, according to Plenborg (2002) the RI approach did not perform particularly well when terminal value calculations are important. This is the case when the book value of equity is a bad indicator of firm value.

The Penman and Sougiannis (1998) and Francis et al (2000) findings suggest that the RI approach yields more accurate firm estimates than the DDM and the DCF approaches. However, their findings conflict with the finding in Plenborg (2002) that the RI and DCF approaches are both inherently based on the DDM and thus, from a theoretical perspective, should yield the same firm value estimates. Plenborg (2002) also finds that the three valuation approaches generate the same point estimate of firm value in practice, if the same assumptions are applied. This indicates that, as Plenborg (2002) argues, neither Penman and Sougiannis (1998) nor Francis et al (2000) have taken into consideration that the same assumptions must be applied. An examination of their test methods also indicates that this is the case. For example, the growth rates used to estimate the terminal value are arbitrarily set at 0 and 4 percent in both studies. Thus, the link between the forecasted financial statements and the input in the different valuation approaches is most likely inconsistent. Furthermore, according to Plenborg (2002), both studies seem to ignore that growth generally affects the free cash flow negatively. They adjust the growth rate without a corresponding adjustment of the free cash flow.

on average, 68 percent of the variation in share price. He advocates the RI model for its accuracy and for its reliance on earnings and book value predictions over relatively short time periods compared with the longer periods generally needed for the DCF model. On the other hand, Lee et al (1999) do not compare valuation models but use the RI to estimate the intrinsic value of the Dow over 1963-1996. They use security analysts' consensus earnings forecasts after 1979 when they became available and time-series projections of earnings before that. Their estimates of intrinsic value predict both the future value of the Dow and the future stock returns to the Dow. Based on these results, Lee et al (1999) advocate use of the RI over alternative valuation models.

Accordingly, if the valuation approaches are not properly employed (as with the growth rate used in the terminal value calculations), the approaches yield different firm value estimates. Thus, the studies of Penman and Sougiannis and Francis et al indicate that if the internal coherence between the three valuation approaches is violated, the RI approach should be preferred for firm valuation at the expense of DDM and DCF approaches. Levin and Olsson (2000) demonstrate that if the steady state condition is not reached when the terminal value is calculated, the RI approach yields more accurate firm value estimates than DCF. Penman and Sougiannis (1998) argue that an attractive valuation approach should be easy to use and understand and it should help the user to perform better firm value estimates. Thus, valuation approaches based on measures that show value creation rather than value distribution are easier to understand and interpret and consequently analytically attractive (Penman, 1992).

Plenborg (2002) demonstrates that simplifying assumptions affect firm value estimates differently. In some cases the RI approach yields more accurate firm value estimates, while in others the DCF approach yields more accurate estimates. His study also shows that each of the assumptions examined affects firm value estimates in a predictable manner. For example, applying the growth term in the forecasted financial statements, the RI approach yields more accurate firm value estimates than the DCF approach. His study also argues that since the framework for forecasting is based on accrual accounting and since budget control is generally based on accounting numbers rather than cash flow measures, it seems logical to estimate firm
values based on concepts known from accrual accounting and financial statement analysis. According to this reasoning, the RI approach seems to be an attractive alternative to the DCF approach.

Dechow and Kothari (1998) pointed out that previous empirical applications of the RI model ignored Ohlson’s information dynamics. In many cases, the resulting valuation model is similar to past applications of dividend-discounting models that capitalise current or forecasted earnings, but make no appeal to book value or residual income. He argues that it is important to note that these empirical models are just a restatement of the DDM, which in no way depends on the properties of accounting numbers other than through the clean surplus relation. In their study, they find that residual income follows a mean reverting process. Their pricing tests indicate that stock prices partially reflect the mean reversion in residual income. An important implication of this result is that book value conveys additional information over earnings in explaining contemporaneous stock prices. However, they also find that book value provides very little additional information about stock prices beyond that contained in analysts’ forecasts of next year’s earnings. They conclude that Ohlson’s formulation of the residual income valuation model provides a parsimonious framework for incorporating information in earnings, book value and earnings forecasts in empirical research.

3.10 Conclusion

EVA, DCF, DDM and RI are compared theoretically and empirically in order to highlight whether it is possible to infer a superior method among them to be used in this research. Although these models are preferable to ROA, they are measuring value rather than performance. However, the first difference can be used as a performance measure but with a correction for dividend. It is difficult to infer from prior literature whether one valuation approach is superior to the other. Because of

\[ \text{Example from these models: Frankel and Lee (1998), it can be seen from this model that the authors used the book value + present value of residual income without using the third assumption of Ohlson's (1995) model.} \]
their theoretical equivalence, the four approaches, DDM, DCF and RI (including the Ohlson model, the Frankel and Lee model and the EVA model), all provide the same valuations when the flows are projected consistently to infinity and comparable discount rates are applied. However, horizons over which the flows can be reasonably projected in practice are limited, and estimated discount rates are subject to error. These practical considerations cause some academics and practitioners to prefer one valuation model to another. The claimed conceptual advantages of the DCF model are based on its corporate finance roots that emphasise cash flows (Brealey and Myers, 2000). Practical valuation ‘handbooks’ such as Copeland et al (2000, 1994) maintain that cash flows outperform accounting earnings for valuation purposes and thus advocate the DCF model.

However, as reported by Herz et al (2001), the claimed practical advantages of the RI are explained by Penman and Sougiannis (1998). In their view, a shortcoming of the DCF model is the need to subtract long-term capital investment from operating cash flows to compute FCF. For growing firms, negative FCF often results for many years. RI adherents maintain that accrual accounting eliminates the distorting effect of capital investment expenditure by placing it on the balance sheet as an asset. Depreciation and amortisation then allocate this investment cost to expense over time; in principle matching it against the revenues that it generates. Penman and Sougiannis (1998) also demonstrate that the RI model’s use of accrual accounting allows for more reasonable valuations than the DCF model for forecasted payoffs over relatively short horizons. On the other hand, the DCF model’s reliance on FCF may require many more years of forecasts to attain steady state and positive FCF.

In short, these methods, namely, EVA, DCF, and EBO (the Ohlson and the Frankel and Lee models) have a link with capital structure through cost of capital. Moreover, the literature showed that the cost of capital is very crucial in estimating the value on each model. This link can be very helpful in examining the relation between capital structures and firms’ performance. As will be explained in Chapter 5, the main shortcoming of implementing the EVA model in this project is due to the difficulties in obtaining the required information. In addition, DCF model looks at valuation more explicitly than performance, while RI models can be used for performance more explicitly than DCF. This gives preference for the RI models since the concern of this
project is for performance rather than valuation. From the literature, we can see that
certain difficulties would be encountered in implementing the linear information
dynamics in Ohlson’s model. Consequently, this gives preference in this study to the
Frankel and Lee (1998) model for reasons explained in Chapter 5 – Testing the
Residual Income Model.
4.1 Introduction

This chapter presents the research design and methodology to be applied in this study. Section two gives a brief idea of the research objectives, followed by the hypotheses in section three. The models that will be applied to test and analyses the formulated hypotheses are presented in section four. Section five discusses the type of data followed by data and sampling description. The appropriate software packages chosen for the analysis will be identified in section six, followed by the conclusion in section seven.

4.2 Research Objectives

This section outlines the research objectives. The aim of this study is to test for a relationship between capital structure and firm performance. From the literature, it was found that agency theory provides a major role for debt in disciplining managers and forcing them not to invest in negative NPV projects. This disciplining role should be reflected in better firm performance (Jensen 1986). Thus, a theoretical relationship between capital structure and firm performance was established. On the other hand, measurement of financial performance represents a real challenge since implementing different performance measures would lead to different results. In this study, the Residual Income model, presented by Frankel and Lee (1998), will be modified and used as a performance measure for reasons explained in Chapter five. Previous
studies have focused on the earnings figures, and less attention has been paid to the informative value of the balance sheet items. Alternatively, researchers have used the market returns to examine the effects of capital structure on performance. However, market returns reflect market expectations about a firm’s future and thus may not be able to capture managers’ current performance.

Accordingly, the main objectives of this study are:

- Testing the full Residual Income model (Frankel and Lee, 1998 version) as a valuation model using non-financial companies from FTSE 500 and S&P 500.

- Testing the Residual Income Component model (derived from the full Frankel and Lee model) to measure firm performance. The Frankel and Lee model (1998) is a measure of value creation that can be applied as a performance measure by adjusting the first difference for dividends (see Chapter 5).

- Testing whether the use of debt can mitigate the over-investment problem.

- Testing whether the use of debt, where it serves as mitigation for the over-investment problem, can lead to a better performance using the above tested performance measures (objectives 1 & 2).

- Examine country differences between UK and US, if any.

4.3 The Hypotheses

The research objectives can be transcribed into the following hypotheses:

The Residual Income model (the F&L 1998 model) is expected to perform well as a valuation model against the market stock prices in our sample and therefore, it will be hypothesised:
H1₀: There is no relation between the Residual Income valuation and stock market prices.

H1₁: The defined Residual Income model can, to a significant degree, capture cross-sectional variation in market stock prices of the sample over a significant period of time.

If the above null hypothesis is rejected, then firms performance will be determined based on the F&L (1998) model in two ways:

Firstly, by looking at first difference in the valuation model; as the F&L model is a valuation model the difference between the value at time t+1 and t will represent performance from an accounting point of view after adding back the dividends (see Chapter 5 – Determination of Performance). In this case, this hypothesis will be tested against market stock price performance i.e. first difference. Accordingly, our second hypothesis will be as follows:

H2₀: There is no relation between first difference of the Residual Income valuation (F&L model) and the performance of the stock market prices (first difference).

H2₁: The first difference of the defined Residual Income model can, to a significant degree, capture the first difference of the cross sectional variation in market stock prices of the sample over a significant period of time.

Secondly, firm performance can be determined from the F&L valuation model by using the components of the residual income model excluding beginning book value and continuing values. The book value at the beginning and the continuing value at the end of the model are important for valuation but not for performance as such (see Chapter 5). The modified model (the residual income components) will also be tested and the following will be hypothesised:

H3₀: There is no relation between the Residual Income Components (the modified F&L model) and the performance of the stock market prices (first difference).
H3: The defined Residual Income Components can, to a significant degree, capture cross-sectional differences in the performance of the market stock prices of the sample over a significant period of time.

Hypotheses two and three will then be compared against each other and against the simple residual income measure\(^\text{13}\) in order to choose a performance measure. Although the simple residual income can be used as a performance measure, the above two performance measures, the first difference and the RIC (hypotheses 2 and 3) are expected to outperform the simple residual income because they include 2-year analysts’ earnings forecasts. Evidently, analysts’ earnings forecasts contain more value-relevant information than is reflected in a historical simple one-period residual income model.

Testing for firm performance, debt and over-investment will be carried out in several steps. Firstly, since gearing can theoretically mitigate the over-investment problem to a certain degree, the following will therefore be hypothesised:

\[ H_{40}: \text{There is no relationship between gearing and over-investment.} \]

\[ H_{41}: \text{There is a negative relationship between gearing and over-investment.} \]

Where gearing mitigates the over-investment problem, this should be related to a higher performance and the following would therefore be hypothesised:

\[ H_{50}: \text{There is no relationship between over-investment and firm performance.} \]

\[ H_{51}: \text{There is a negative relationship between over-investment and firm performance.} \]

Introducing the interaction term between gearing and the over-investment problem would capture the link between the fourth and the fifth hypotheses and the following would therefore be hypothesised:

\[ H_{60}: \text{The interaction between the presence of over-investment and level of gearing is not related to firm performance.} \]

---

\(^{13}\) The simple residual income model will be tested as well and compared with the above two performance, namely, the first difference and the RIC in the model.
H6: the interaction between the presence of over-investment and level of gearing is related to better firm performance.

The research is divided into several steps. The first step is to test the Residual Income model on its ability to capture, to a significant degree, the cross-sectional variation in stock market prices of FTSE 500 and S&P 500 companies over a significant period of time (Hypothesis one). Although, the residual income model has been tested by F&L (1998), however, the sample was from the US market only. In this study, we will test this model for both the US and the UK market.

The second step tests the first difference of the Residual Income model against first difference of market stock prices, and tests the modified Residual Income model’s ability to capture, to a significant degree, the cross-sectional stock market performance (first difference) of the same sample over a significant period of time (Hypothesis two and three). Then performance measure will be defined based on results of hypotheses one to three.

Hypothesis four tests whether gearing, can to a significant degree, mitigate the over-investment over the sample and period. Hypothesis five tests the relation between over-investment and performance, and the final step tests the relation between firm performance, leverage and over-investment (Hypothesis six).

4.4 Research Methods

4.4.1 Regression Analysis

In order to test the above hypotheses, a regression analysis will be used. Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables, which are identified as the explanatory or independent variables, with a view to estimating and or predicting the mean or average value of the former in terms of known or fixed values of the latter (Gujarati, 2002). The single regression equation is of the following format:

\[ y = \beta_0 + \beta_1 x + \epsilon \]

14 The developments of these hypotheses together with research design are explained in Chapter 5.
\[ Y_i = \alpha + \beta_1 X_{i1} \quad (1) \]

Panel data regression\(^{15}\) differs from a regular time series or cross-section in that it combines both in a double subscript on its variables, i.e. \( X_{it} \).

\[ Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_t X_{it} + u_{it} \quad (2) \]

Where the \( i \) subscript denotes the cross-sectional dimension whereas \( t \) denotes the time-series dimension. \( Y_i \) is the dependent variable and \( X_{i1}, X_{i2}, \ldots, X_{it} \) are the explanatory variables. \( \alpha \) is the constant and \( \beta_1, \ldots, \beta_t \) are the slopes of the explanatory variables. The error terms are assumed to have the properties:

\[
E(u_i) = 0 \\
\text{Var}(u_i) = \sigma^2 \\
\text{Cov}(u_i, u_j) = 0 \text{ for } i \neq j
\]

These relationships state that the error terms are assumed to have normal distribution with mean 0 and variance \( \sigma^2 \), and that error terms must be independent. Most of the panel data applications utilise a one-way error component model for the disturbances, with:

\[ u_{it} = \mu_i + \nu_{it} \]

Where \( \mu_i \) denotes the unobservable individual specific effect and \( \nu_{it} \) denotes the remainder disturbance. Note that \( \mu_i \) is not time variant and it accounts for any individual specific effect that is not included in the regression. The remainder disturbance \( \nu_{it} \) varies with individual and time and can be thought of as the usual disturbance in the regression.

\(^{15}\) Panel data analysis will be used in this study, for more information refer to section 4.5 of this chapter
4.4.1.1 Sample Regression Function

The sample regression function (SRF) is determined based on the Population Regression Function (PRF) represented in Equation (1) and is determined using the Ordinary Least Square (OLS) method. The OLS method defines the SRF in such a way that the sum of the sample residuals is as small as possible, as a result of which the SRF is as closely aligned as possible to the PRF. The sample Regression Line has the following properties:

- It passes through the sample means of Y and X;
- The mean value of the estimated Y is equal to the mean value of the actual Y;
- The mean value of the estimated residuals is zero;
- Residuals are un-correlated with the estimated Y;
- Residuals are un-correlated with the estimated X.

4.4.1.2 Quality of the Sample Regression Function

Since data vary from sample to sample, estimates will change, which results in the requirement to measure the reliability or the precision of the estimates. The precision of the regression model is determined by its standard error(s). The standard error is simply the standard deviation of the Y values about the estimated regression line and is often used as a summary of the goodness of fit of the estimated line. The goodness of fit is determined by R^2, which measures the proportion or percentage of the total variation in Y explained by the regression model. The T-test or Chi^2 test is a test of significance, by which sample results are used to verify the validity or invalidity of the rejection of the null hypothesis. It is the measure of the overall significance of the estimated regression, and it is also a test of the significance of R^2.
4.4.1.3 Fixed and Random Effects Models

As we have already mentioned, most of the panel data applications utilise a one-way error component model for the disturbances, with:

\[ u_{it} = \mu_i + v_{it} \]

Where \( \mu_i \) denotes the unobservable individual specific effect and \( v_{it} \) denotes the remainder disturbance. Note that \( \mu_i \) is not time variant and it accounts for any individual specific effect that is not included in the regression. The remainder disturbance \( v_{it} \) varies with individual and time and can be thought of as the usual disturbance in the regression. A simple way to take account of heterogeneity across individuals and/or through time is to use variable intercept models. The basic assumptions of such models are that, conditional on the observed explanatory variables, the effects of all omitted variables are driven by three types of variables: individual time-invariant, period individual-invariant, and individual time-variant variables.

The individual time-invariant variables are variables that are the same for a given set of cross-sectional unit through time, but vary across cross-sectional units. Examples of these are attributes of the individual-firm management, ability, sex etc. The period individual-invariant variables are variables that are the same for all cross-sectional units at a given point in time, but vary through time. Examples of these are prices, interest rates, and widespread optimism or pessimism. The individual time-variant variables are variables that vary across cross-sectional units at a given point in time and also exhibit variations through time. Examples of these variations are a firm’s sales, profits, and capital stock.

The variable intercept model assumes that the effects of numerous omitted variables are each individually unimportant but are collectively significant and possess the property of a random variable that is un-correlated with all other included and excluded variables. On the other hand, because the effects of the remaining omitted variables either stay constant through time for a given cross-sectional unit or are the same for all cross-sectional units at a given point in time, or a combination of both, they can be absorbed in the intercept term of a regression model as a means to
explicitly allow for individual and/or temporal heterogeneity contained in the temporal cross-sectional data. The variable intercept model can provide a useful specification for fitting regression models using panel data.

The fixed effects estimation method treats the unobservable specific effects $\mu_i$ as fixed constants. In this case $\mu_i$ are assumed to be fixed parameters and the remaining disturbances stochastic, with $\nu_{it}$ independent and identically distributed. The $X_{it}$ are assumed independent of the $\nu_{it}$ for all $i$ and $t$. The fixed effects model is an appropriate specification for focusing on specific sets of firms, and our inferences are restricted to the behaviour of these sets of firms.

On the other hand, random effect models treat the individual specific effects $\mu_i$ like $u_i$, i.e. as random variables. It is standard practice in regression analysis to assume that large numbers of factors affect the value of the dependent variable, but which have not been explicitly included as independent variables, can be explicitly appropriately summarised by a random disturbance. When numerous individual units are observed over time, it is sometimes assumed that some of the omitted variables will represent factors peculiar to both the individual units and the time periods for which observations are obtained, whereas other variables will reflect individual differences that tend to affect the observations for a given individual in more or less the same way over time. Still other variables may reflect factors peculiar to specific time periods, but affecting individual units more or less equally. The random effects model is an appropriate specification if we are drawing $n$ individuals randomly from a large population. In this case, $n$ is large and a fixed effects model would lead to an enormous loss of degree of freedom.

The issues of whether to treat unobserved heterogeneity as random or fixed is a question with no easy answer, and it has aroused significant interest among econometricians and has paramount importance in panel data modelling (Matyas and Sevestre, 1996). To decide on an appropriate structure for the analysis, namely random effects verses fixed effects; it appears that consideration should be given to: a)

---

16 Hausman (1978) specification test can be used to determine the appropriate method i.e. fixed or random effect. However, this kind of test is not available in the software packages that will be used in this study namely, E-Views and PC-give.
the objectives of the study, b) the context of the data, the manner in which they are
generated, and the environment from which they came. Thus, the fixed effects model
is an appropriate specification if we are focusing on specific sets of firms, and our
inferences are restricted to the behaviour of these sets of firms, while the random
effects model is an appropriate specification if we are drawing \( n \) individuals randomly
from a large population. Fixed effect models may be more appropriate for this study\(^{17} \).

### 4.4.1.4 Other Issues to be Considered

Several aspects connected with regression analysis have to be taken into consideration
when writing the final conclusions. Two of these issues are Heteroscedasticity, which
will be controlled by the software used in this study, and Multicollinearity, which
means the existence of a perfect or exact linear relationship among some or all-
explanatory variables of a regression model. Therefore, the quality of the regression
analysis will be based on a robustness test of the model.

### 4.4.2 Regression on Dummy Dependent Variable

A basic statistical method for social science is linear regression analysis, which
requires a continuous dependent variable and explanatory variables that are either
continuous or categorical. However, the classical regression model cannot be applied
in situations where the dependent variable refers to characteristics, attitudes,
behaviours and decisions that are measured in discrete, nominal, ordinal, or any non-
continuous way. In such situations the dependent variable takes a discrete number of
mutually exclusive and collectively exclusive values (Borooah, 2001).

In this study, and in hypothesis three, the dependent variable is a discrete variable,
which takes a value of one if there is an over-investment and zero otherwise. The
statistical methods that are used to analyse such data are known as qualitative choices
models and present a common characteristic, i.e. they all model the probability of an
event, namely, how likely the event is to occur (Liao, 1994). Examples of such
methods of statistical analysis are “binary data analysis”, “ordered analysis” or

\(^{17}\) Although it has been decided to use fixed effect models, in the analysis we use random effect models
as well and the results are comparable.
"discrete choice analysis", such as the logit and probit models. When the dependent variable is the outcome of more than two choices, then a multinomial logit model is used, and when there is an ordered choice in which the discrete variable can take the values of 0, 1 and 2, then the ordered model should be used. When the dependent variable is a count variable, e.g. number of accidents, number of visits in a theme park, it is assumed that the random component in Y resembles the Poisson distribution and thus the logarithmic model is used. As the dependent variable in this study takes two values only, i.e. either 0 or 1, a logit or probit model should be used.

When we study a random variable Y (dependent variable) using a linear model, we specify its expectation as a linear combination of K independent variables (X) as follows:

\[ E(Y) = \mu = \sum_{k=1}^{K} \beta_k X_k. \]  

This is the ordinary linear model, and it can be used when the normality of the distribution of the random component in Y (i.e. the part that cannot be explained by the independent variables) can be assumed. However, when the dependent variable is measured in a non-continuous way, the random component in Y does not follow the normal distribution, in which case the ordinary linear model in Equation (3) cannot be applied.

There are four main reasons why the classical regression (and the OLS estimator) cannot be applied to estimate models where the dependent variable is non-continuous (Gujarati, 2002). The first refers to the fact that the random component in Y does not follow the normal distribution. Another problem is that the random component is heteroscedastic and thus one needs to find appropriate transformations to make it homoscedastic. Yet, the most important handicap of using the OLS estimator is that there is no guarantee that the estimated probability that the event under examination will occur will lie within the limits of 0 and 1. For example, if the estimated probability \( P_i \) is negative, or greater than 1, it will have no practical meaning. Finally, by applying classical regression one would assume that the rate of change of probability per unit change in the value of the explanatory variable is constant and is given by the value of the slope.
However, when a logit or probit model is used, the estimated probabilities will always fall between 0 and 1. Moreover, the probability of an event occurring does not increase linearly with a unit increase in the value of the explanatory variable. Rather, the probability approaches zero (or one) at a slower and slower rate as the value of the explanatory variable gets smaller (or larger) respectively. This is a more realistic pattern of change in the probability compared to linear model. More specifically, if the dependent variable is a binary outcome (0/1), which is the case in this study, the random component in Y follows the binomial distribution and according to Liao (1994) the appropriate model to be used for estimation is either logit:

\[ E(Y) = \log \left( \frac{\mu}{1 - \mu} \right) \]  

(4)

or Probit:

\[ E(Y) = \phi^{-1}(\mu) \]  

(5)

Where \( \phi^{-1} \) is the inverse of the standard normal cumulative distribution function.

**4.4.2.1 Binary Logit/Probit Model**

As the simplest types of probability model, binary logit and probit models have only two categories in the response variable, i.e. 0 and 1. Binary models are usually denoted as latent variable specifications in which the response variable \( y^* \) is linearly related to a set of \( k \) explanatory variables, \( x_k = (x_1, x_2, \ldots, x_k) \) and is defined by the regression relationship:

\[ y^* = \sum_{k=1}^{k} \beta_k x_k + \varepsilon \]  

(5)

In practice, \( y^* \) is unobserved and \( \varepsilon \) is systematically distributed with zero mean and has its cumulative density function (CDF) defined as \( F(\varepsilon) \). The observed dependent variable is determined by whether \( y^* \) exceeds a threshold value:

\[ y = 1 \text{ if } y^* > 0 \text{ or } y = 0 \text{ otherwise.} \]
Then, the probability of observing a value of 1 is denoted by the following relationship:

\[
prob(y = 1) = prob\left(\sum_{k=1}^{k} \beta_k x_k + \varepsilon > 0\right) = prob\left(\varepsilon > -\sum_{k=1}^{k} \beta_k x_k\right) = 1 - F\left(-\sum_{k=1}^{k} \beta_k x_k\right) \tag{6}
\]

Where \(F\) is the CDF of \(\varepsilon\).

The functional form of \(F\) depends on the assumption made about the distribution of the random component \(\varepsilon\). If the random component is assumed to follow the logistic distribution, then the binary model is known as logit model, where, substituting \(L\) (signifying the logistic distribution) for \(F\) in Equation (6):

\[
prob(y = 1) = 1 - L\left(-\sum_{k=1}^{k} \beta_k x_k\right) = L\left(\sum_{k=1}^{k} \beta_k x_k\right) = \frac{\exp\left(\sum_{k=1}^{k} \beta_k x_k\right)}{1 + \exp\left(\sum_{k=1}^{k} \beta_k x_k\right)} \tag{7}
\]

If the random component is assumed to follow the standard normal distribution, then the binary model is known as the probit model and \(Prob(y=1)\) is given by the following equation:

\[
prob(y = 1) = 1 - F\left(-\sum_{k=1}^{k} \beta_k x_k\right) = F\left(\sum_{k=1}^{k} \beta_k x_k\right) = \phi\left(\sum_{k=1}^{k} \beta_k x_k\right) \tag{8}
\]

4.4.2.2 Should the Logit or the Probit Model be Chosen?

It can be seen from the above formulas that there is little difference between the two CDFs, i.e. the logit and the probit model. Normally, a logit model has flatter tails compared to probit. That is, the probability \(P_i\) approaches 1 or 0 at a slower rate in logit models than in probit. Given the similarities between them, either model will give very comparable conclusions in most applications\(^{18}\). It is, in fact, easy to go from one set of estimates to the other. If a probit estimate is multiplied by a factor of approximately 1.6, then the corresponding logit estimate can be obtained.

\(^{18}\) In the analysis, both logit and model have been used and the results are comparable.
There are situations, though, where estimates from logit and probit models may differ substantially, and in such cases care must be taken in choosing the more appropriate one. These are cases with an extremely large number of observations and with heavy concentrations of observations in the tails of the distribution. Logit models are more appropriate for distributions with heavier tails. That is, if movement towards probability of 0 or 1, after certain values of the regressors have reached, occurs quickly, then a probit model provides a better approximation to the data generation process. Otherwise, the logit model is preferred and more commonly used (Seddighi, Lawler and Katos, 2000).

4.5 Panel Data

Three types of data may be available for empirical analysis: time series, cross-sectional and pooled (combination of time series and cross-sectional) data. Panel data analysis is at the watershed of time series and cross-section econometrics. While the identification of time series parameters traditionally relies on notions of stationary, pre-determinedness and uncorrected shocks, cross-sectional parameters appealed to exogenous instrument variables and random sampling for identification. By combining the time series and cross-sectional dimensions, panel data sets have enriched the set of possible identification arrangements, and forced economists to think more carefully about the nature and sources of identification of parameters of potential interest.

Hsiao (2003), Klevmarken (1989), Solon (1989) and Baltagi (2003) list several benefits from using panel data. These include the following:

- Panel data give the researchers a large number of data points, increasing the degree of freedom, variability and efficiency, and reducing collinearity among explanatory variables. The large number of data points is very important when using financial accounting data, which are published only annually.
Panel data allow a researcher to analyse a number of important economic questions that can not be addressed using either pure cross-sectional or time series data sets. Panel data are better able to study the dynamics of adjustments and are able to identify and measure effects that are simply not detectable in pure cross-sectional and time series data. They also provide a dynamic picture of the samples' financing behaviour.

Panel data better control individual heterogeneity. Panel data allow for the fact that individuals, firms, countries are heterogeneous. Time series and cross-sectional data not controlling for this heterogeneity run the risk of yielding biased results.

Panel data models allow us to construct and test more complicated behaviour models than do purely cross-sectional or time series data, and also fewer restrictions can be imposed in panels than in a purely time series study.

Panel data are usually gathered for micro-level units, like individuals, firms, and households. Many variables can be more accurately measured at the micro level and biases resulting from aggregation over firms or individuals are eliminated.

Accordingly, panel data analysis will be used in this study to empirically examine the hypotheses formulated before. Panel data analysis is very popular in product placement and marketing research, as well as in studies concerned with the labour force. Econometric analysis in finance is mainly carried out using time-series or cross-sectional data. The availability of long time series or disaggregated price data partially explains this phenomenon. None of the empirical work for testing the Residual Income model using panel data has been done. In fact, all of the previous studies that have tested this model have either used cross-sectional or time-series models. Therefore, it will be of interest to examine the behaviour of the Residual Income model using the panel data model. Conversely, although many new empirical studies use panel data to examine for capital structure, this is still limited in scope.

However, one should be careful in analysing panel data and paying attention to the following limitations:
• Design and data collection problem: these include problems of coverage i.e. incomplete account of the population of interest.

• Distortions or measurement errors: measurement errors may arise because of inappropriate information, misrecording of data.

• Selectivity problems: in the form of self-selectivity, non-response or attrition.

• Short time series dimension: typical panels include annual data covering a short span of time for each individual. Increasing the number of years is not without cost either. This increases the chances of attrition and increases the computational difficulty for limited dependent variable panel data models.

4.6 Data and Sample Description

Most of the data used in this study were gathered from secondary sources. The main source of information has been the Datastream database, which contains published accounts data as well as stock prices. In a few cases, copies of companies’ annual reports were used in order to complete the set of data or to make certain clarifications. The original sample consists of both FTSE 500 and S&P 500 comprising both US and UK markets to allow a comparison between these two markets. In addition, companies from both markets were grouped based on the industry to which they belong. This sorting of companies will allow us to make a comparison between industries and to analyse the differences. The sample including industry classification and based on the above classification is presented below in two separate tables representing both FTSE 500 and S&P 500 respectively as follows:
Table 4.1: Description of FTSE 500 and S&P 500

<table>
<thead>
<tr>
<th>Industry</th>
<th>FTSE500</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Services</td>
<td>64</td>
<td>106</td>
</tr>
<tr>
<td>Utility</td>
<td>28</td>
<td>66</td>
</tr>
<tr>
<td>Business Support (Services)</td>
<td>45</td>
<td>11</td>
</tr>
<tr>
<td>Media &amp; Publication</td>
<td>38</td>
<td>26</td>
</tr>
<tr>
<td>Engineering</td>
<td>41</td>
<td>16</td>
</tr>
<tr>
<td>Retail</td>
<td>47</td>
<td>40</td>
</tr>
<tr>
<td>Construction</td>
<td>38</td>
<td>5</td>
</tr>
<tr>
<td>Transportation</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>Real Estate Development</td>
<td>35</td>
<td>12</td>
</tr>
<tr>
<td>Chemical and Pharmaceutical</td>
<td>28</td>
<td>44</td>
</tr>
<tr>
<td>Computer</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>Electronic and Telecom Equip.</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>Hotels and Food</td>
<td>48</td>
<td>31</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>Health Care &amp; equipment</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td><strong>Total Firms</strong></td>
<td><strong>500</strong></td>
<td><strong>500</strong></td>
</tr>
</tbody>
</table>

Financial companies have specific characteristics in their capital structure that may be affected by regulatory requirements such as capital adequacy. In addition, debt-like liabilities of financial firms such as banks and insurance companies are not strictly comparable to debt issued by non-financial firms (Rajan and Zingales, 1995). Due to the above-mentioned reasons, and the desirability of a homogenous sample, it has been decided to exclude financial firms from both FTSE 500 and S&P 500. Moreover, by examining the sample it has been observed that real estate development companies have some special characteristics and legislation that are different from the non-financial companies as well. For instance, such companies do not report sales figure in their trial balances and they are therefore excluded. As a result, FTSE 500 has
decreased from 500 to 401 companies while S&P 500 has decreased from 500 to 382 companies, giving a total of 783 companies.

Only companies with a full set of data for the 15-years (1988-2002) period were selected. Furthermore, companies were required to have a one-year-ahead and a two-years-ahead earnings-per-share (EPS) forecast from I/B/E/S in order to test the Frankel and Lee (F&L) (1998) model. Further, it was realised that many companies within the sample have different fiscal-year-ends. In order to have a comparable data within the companies, the sample was constrained to companies with fiscal-year-ends between September 30th and December 31st as well as between January 31st and March 31st. The number of companies was thus reduced from 783 to 317 comprising 185 companies from the US market and 132 companies from the UK.

The initial number of firms was reduced so drastically because many companies had only a few years’ data or have no reported EPS forecast. Some of them had “died”, or were young companies, or were taken over. Some others were excluded because of their fiscal-year-ends. A few firms had negative book values and these firms were also excluded because according to F&L (1998) model, return on equity (ROE) for these firms cannot be interpreted in economic terms. Finally, it was clear that the F&L (1998) model does not work when the dividend pay-out ratio (K) exceeds 100% (see definition of the variables on Chapter 5 section 5.4). Such firms with k above 100% were also eliminated.

The above common sense filters ensure the subsequent results are not driven by outliers. The tests are carried out on the overall sample of 185 US firms and 132 UK firms as well as on the specific portfolios created in order to see if there were any differences among them. These differences are expected because different firms belong to different industries as well as different markets and therefore may face different business risk. A careful investigation of firms’ business line was made by looking at all the SIC classifications to which these firms belonged and ensuring that the grouped firms were as similar as possible as well as ensuring that the portfolios were large enough in order for them to contain the necessary number of observations for the statistical tests to be robust¹⁹.

¹⁹ The seven major sectors used by FT Actuaries/Goldman Sachs were also used in this study. See Richard Roll, the Journal of finance, Vol. 47(1), (Mar., 1992), 3-41.
The year 1988 was selected as a cut-off year because 9 years of data were deemed necessary in order to draw sound statistical conclusions from the tests described later. This takes into consideration the fact that some of the variables are calculated by using two years’ previous accounts; future return on equity at time \( t \) is calculated by using book value at time \( t - 1 \) and \( t - 2 \) respectively. Furthermore, some of the variables are also calculated as averages or standard deviations of three years. In addition, one more year was needed in order to obtain a variable (performance), which is calculated as the difference between two years. Therefore, an initial 15 year period of raw data was needed in order to produce a 9 year period for both dependent and independent variables.

4.7 Statistical Packages

Several software packages could be used to carry out the empirical analysis of this study. SPSS, E-Views, LIMDEP and PC-give are among the most commonly used. Although SPSS is a very user-friendly statistical package, which is widely used, it could not be used for the current analysis, since it accepts a maximum of 10 dependent and independent variables (combined). Moreover, SPSS has limited functions for panel data analysis, which will be applied in this study. Because of the previous mentioned restrictions, SPSS would not be used in this study. LIMDEP was another option; however, due to unavailability of the software, it could not be considered either.

‘E-Views’ statistical package provides all required tools to deal with regression analysis. For instance, ‘E-views’ does not weight observations in pooled estimation by default, but there is the option of estimating weighted versions of the specifications. E-views can control for heteroscedasticity through White heteroscedasticity consistent covariance. ‘E-Views’ allows for options such as fixed and random effects. However, ‘E-Views’ has a limited function to deal with discrete choice models. These models have to be performed either using cross-sectional or time-series analysis. On the other hand, ‘PC-give’ has almost the same features of ‘E-Views’ such as controlling for heteroscedasticity and allowing for fixed and random effects, in addition to analysing
discrete models using panel data. Therefore, ‘E-Views’ and ‘PC-give’ will be used in this study.

4.8 Conclusions

This chapter describes the type of research design and methodology used in the following empirical research. The objectives of this study were highlighted in the first section. The main objectives were to determine firm performance and then to link this to capital structure through hypotheses based on agency theory and the over-investment problem. These objectives were then transcribed into five main hypotheses, which will be tested in the following empirical research. Two main ways were then discussed to analyse and test the formulated hypotheses. The first was OLS regression analysis, which will be used to analyse the results where the dependent variable is a continuous variable. The second was binary discrete choice (logit and probit) models which will be used to analyse the results where the dependent variable is a discrete choice such as 0 or 1. The main properties and features of each model were then highlighted.

This chapter discussed the advantages of using panel data in this study as well as the problems that might arise as the result of pooling together time series and cross-sectional data. The chapter describes the data and sample description that will be used in the following empirical research. The data used in this research come from secondary resources, more specifically from the Datastream database. The study uses 15 years of accounting data from 1988 to 2002, during which the databases provided full set of accounts. FTSE 500 and S&P 500 are the sources used in the sample of this study. After filtering the data, the remaining samples consist of 184 companies from US and 132 companies from UK, comprising a total of 317 companies. Finally, this chapter discusses the use of both ‘E-Views’ and ‘PC-give’ to carry out the empirical analysis of this study.
Chapter Five

Testing the Residual Income Model

5.1 Testing the Residual Income Model

This chapter examines the ability of the Residual Income model (RI) to capture and explain, to a significant degree, cross sectional variation in market stock prices of the selected sample over a significant period of time. This test provides, in addition to how performance will be determined based on a valuation model, an answer to the question of how best to measure performance. Such a question is important because accounting and budgeting systems, performance measurement systems, transfer pricing systems, and decision support systems affect how people and organisations interact.

It can be seen from the literature that the continually growing criticism of traditional performance measures, which ignore the cost of capital, motivate dysfunctional behaviour causing managers to pay attention to the wrong things. Back in 1979, Lev and Sunder commented that in using such traditional measures, i.e. ratio analysis, it appears that the extensive use of financial ratios by both practitioners and researchers is often motivated by tradition and convenience rather than resulting from theoretical considerations or from careful statistical analysis. For this purpose, it has been decided to exclude such performance measures and look at models that take into consideration the spread between return on capital and the cost of capital.

Consulting firms are developing and marketing alternative financial measures of performance such as economic value added, cash value added, shareholder value, etc. They claim that they provide “superior” measures of performance and better incentives in motivating managers to take the right actions. It has been found that the most attractive model among them is the economic value added since it is based on the strong economic argument that the “profit” should be calculated after the cost of capital is deducted. However, empirical research carried out shows that EVA may be
a good proxy for economic profits, but may not outperform the current realisations of other performance measures (Biddle et al, 1997).

Most importantly, implementing the model requires adjusting a company’s balance sheet and income statement. Practically, the points that might arise are:

- It is difficult to make all the recommended adjustments, and to ensure that these adjustments reflect the economic value in order to implement this model. Estimated accounting adjustments may contain measurement error relative to what the market is using for valuing firms.

- Data required to compute EVA are not easily estimated from secondary sources, since the data required are often not directly available on the balance sheet and income statements or notes to the accounts, and it is difficult, if not impossible, to obtain inside information.

This suggests that such a measurement might be useful for consultancy companies when they have easy access to inside information and are therefore able to make the recommended adjustments; this is not the case here. Therefore, EVA will not be used as a performance measure in this project.

It is clear from the literature that RI models such as those of Ohlson (1995) and Frankel and Lee (1998), the Economic Profit model (EP) presented in Copeland et al (1994, 2000) and the Discounted Cash Flow model (DCF) presented in many practical valuation handbooks such as Copeland et al (2000, 1994) are the most competitive models in many of the recent academic books and research papers. Therefore, the first task of this project is to choose one model among RI, EP and DCF. Although each model can be used as a performance measure, DCF, as discussed in the literature review, looks at the valuation of firms more explicitly than performance, while RI and EP can be used for performance more explicitly than DCF. For example, RI and EP models have an advantage over the DCF model in a way that they are useful measure for understanding a company’s performance in any single year, while discounted cash flow is not. Since the purpose of this project is to look at the performance measurements rather than valuation, it seems logical to use RI and EP models, in accordance with their function to evaluate periodic performance, rather than DCF. Therefore, DCF will not be used in this project as a performance measure.
On the other hand, as can be seen from the literature, an obvious limitation concerns the absence of a leverage concept in the Ohlson (1995) model. His model has been based on risk neutrality, and this does not permit the required return to reflect any compensation for the inherent risk in equity securities or firm specific risk. In addition, it should be noted that certain difficulties would be encountered in implementing the linear information dynamics in the Ohlson (1995) model, which frames the stochastic time series behaviour of abnormal earnings. The difficulty is in setting parameter values such as $v_t$ which represent other information in the model. Obversely, Frankel and Lee (1998) operationalised the residual income model in which the cost of capital is risk-adjusted by using an overall assessment of the perceived risk of the investment as a whole. This leads to a more easily implemented model and therefore, leads to preference for the operationalised RI in Frankel and Lee (1998) model over the Ohlson (1995) model.

The EP model, from both a mathematical and theoretical point of view, looks very similar to RI (Frankel and Lee version)\(^{20}\). The two main differences between them are:

- The forecasting over time horizon (period). RI (Frankel and Lee version) has a shorter explicit forecast period than EP (Copeland \textit{et al}, 2000, 1994).

- Using the cost of equity in RI (Frankel and Lee version) because it measures the value of firm’s equity, while the EP (Copeland \textit{et al} (2000) use the WACC as a cost of capital because it measure the firm’s value (equity + debt).

The latter point, i.e. cost of equity or cost of capital in the form of WACC, seems to be less important, as Lee \textit{et al} (1999) point out: the RI model can be adjusted to reflect cost of capital (WACC) rather than just the required return on equity:

$$V_t = B_t + \sum_{i=1}^{n} \frac{(Earnings \ before \ Interest) - (Invested \ Capital \times WACC)}{(1 + WACC)^i}.$$ 

This valuation formula looks the same as EP (Copeland \textit{et al} (2000): 287) where:

$$V_t = Invested \ Capital + CV \ of \ economic \ profit.$$ 

\(^{20}\) RI and EP as can be seen from the literature are just two names for the same concept.
Accordingly, the main difference between these two models is the choice of forecast horizon. Copeland et al (2000) stated that the explicit forecast should be long enough so that the business will have reached a steady state of operations by the end of the period. Similarly, Frankel and Lee (1998) pointed out that in theory the explicit forecast period should be set long enough for firms to reach their competitive equilibrium. However, Frankel and Lee (1998) argued that the ability to forecast future ROEs diminishes quickly over time due to the difficulties in estimating future ROE for more than three years, and forecasting errors are therefore compounded in longer expansions. Henceforth, they estimated only three years of RI, which is also consistent with the consensus forecasts, which give projections for 3 years.

From the literature, it can be seen that Bernard (1995), Penman and Sougiannis (1998), Lee et al (1999) and Frankel and Lee (1998) have tested the RI model and have concluded that it yields to smaller valuation errors, as measured against stock price and over short horizons. Moreover, F&L (1998) point out that the estimated intrinsic values from the RI explain more than 70% of the variation in stock prices over the period of 1975-1993. In summary, many empirical studies conclude from large sample results that the RI, even with a short time horizon, explains stock prices and stock returns better than the alternative valuation models. It also can be concluded that short time horizons eliminate the possibility of large errors that are usually generated by long time horizons, since the larger the time horizon, the bigger the estimated errors.

In general, it can be concluded that RI, modified by Frankel and Lee (1998) among others, provides the preferred model. It can easily be applied using accounting data, requires estimates for a smaller number of variables, and as suggested by the literature, explains a significant percentage of cross sectional variation in market stock prices. This good measure can also be achieved with a short time horizon, which is appropriate for the sample that will be used in this thesis\textsuperscript{21}. Nevertheless, with such a large sample, forecasting for long time horizon (an average of 7 years)\textsuperscript{22}, as suggested by Copeland et al (2000) in order to implement the EP model, presents difficulties.

\textsuperscript{21} For more information about the sample size, please refer to data sampling section.

\textsuperscript{22} Copeland et al (2000) did not mention a specific time horizon; they said it should be long enough in order to insure that the company reaches a steady situation. However, from their examples, it seems that they favourite the time horizon to be a minimum of 7 years.
Therefore, the F&L (1998) model will be used in this study to determine firms’ performance.

5.2 Definition of the Variables

The dependent variable in the above model (P) is the closing market price. The current price on Datastream’s equity program is the latest price available from the appropriate market in primary units of currency. P was calculated by taking the average of eight weeks (four prior to the publication of the firm’s annual profit and 4 weeks after) in order to avoid as much as possible the effect of announcing a firm’s annual result on its stock price.

The independent variable V as stated above is the RI model and is defined as in Frankel and Lee (1998) as well as in Lee, Myers and Swaminathan (1999) as follows:

\[ \hat{V}_i = B_i + \left( \frac{FROE_i - r_e}{1 + r_e} \right) B_i + \left( \frac{FROE_{i+1} - r_e}{1 + r_e} \right) B_{i+1} + \left( \frac{FROE_{i+2} - r_e}{1 + r_e} \right) B_{i+2} \]  \tag{1}

B is the book value per share (Datastream Item 1308) calculated on an issued basis, using that portion of share capital and reserves (excluding preference capital) minus intangibles attributable to the issue, divided by the year-end number of shares in that issue. It is adjusted for subsequent rights and scrip issues.

FROE is the future return on equity per share calculated as follows:

\[ FROE_i = \frac{FY1}{(B_{i-1} + B_{i-2})/2} \]

(The use of the average, rather than year-end, book value, as Frankel and Lee (1998) argue, reduces the chance of an extremely low denominator).
Where FY1 is the earnings-per-share forecasts one-year-ahead derived from the I/B/E/S mean\textsuperscript{23} (also called consensus) forecast\textsuperscript{24}.

\( B_{t-1} \) and \( B_{t-2} \) are the book values at time \( t-1 \) and \( t-2 \) respectively, defined as stated above, where book value at time \( t \) is calculated as follows:

\[ B_t = B_{t-1} \left[ 1 + FROE_t \left( 1 - K \right) \right]. \]

\( K \) in the above equation is the dividend payout ratio. The dividend payout ratio is the percentage of net income paid out in the form of dividends each year. \( K \) is estimated by dividing the Dividends paid by Earnings (Earned for ordinary). Dividend paid (Datastream Item 434) is defined as the total amount of cash dividends listed for common shares. Dividend per share is defined as dividend per share declared during the fiscal year, including extra dividends declared during the year.

Earned for ordinary (Datastream item 625) is defined as the net profit arrived at after deducting tax, minority interest and preference dividends, but before any post-tax as reported extraordinary items, allocations to reserves other than untaxed reserves and post tax disclosed extraordinary items.

For firms with negative earnings, \( K \) was estimated by dividing the common stock dividends by six percent of total assets to derive an estimated payout ratio\textsuperscript{25}.

In estimating \( FROE_{t+1} \) and \( B_{t+1} \) all companies in the sample were required to have a two-year-ahead consensus forecast (FY2). Then \( FROE_{t+1} \) and \( B_{t+1} \) were computed analogously:

\[ FROE_{t+1} = \frac{FY2}{(B_t + B_{t-1})/2} \]

\textsuperscript{23} Using median rather than mean forecasts is unlikely to affect results because the distribution of forecasted growth is quite symmetric.

\textsuperscript{24} We use analyst's forecast of ROE rather than the historic or actual ROE because the forecast ROE includeds analysts' earnings forecasts and evidently, analysts' earnings forecasts contain more value-relevant information then is reflected in historical simple residual income model.

\textsuperscript{25} Six percent reflects the average long-run return-on-assets (see appendix A); as in F&L this measure was used to proxy for normal earnings when reported earnings are negative.
Estimating $FROE_{t+2}$ and $B_{t+2}$ where a long-term earnings growth estimate (LTG) is available, then $FROE_{t+2}$ and $B_{t+2}$ were computed as follows:

$$FROE_{t+2} = \frac{FY2(1+LTG)}{(B_{t+1} + B_{t})}$$

$$B_{t+2} = B_{t+1}[1 + FROE_{t+2}(1 - K)]$$

Where LTG is not available, $FROE_{t+1}$ was used to proxy for $FROE_{t+2}$.

$r_e$ is the cost of equity. In theory, cost of equity should be firm specific, reflecting the premium demanded by equity investors to invest in a firm or project of comparable risk. In practice, however, there is little consensus on how this discount rate should be determined. For this study, the Capital Assets Pricing Model (CAPM)$^{26}$ will be used to estimate $r_e$. The definitions of the variables are summarised below:

---

$^{26}$ There are other methods for calculating the cost of equity such as market model. However, there is little consensus on which one is better. Therefore, the question of which of these models is better is still an open one.
### Table 5.1: Definition of the Variables as in Data-stream

<table>
<thead>
<tr>
<th>Item</th>
<th>Not.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>190</td>
<td>(DPS)</td>
<td>Dividend per share. This item is used to determine the dividends pay out ratio and the related effect on the future book value.</td>
</tr>
<tr>
<td>305</td>
<td>(TBV)</td>
<td>Equity Capital and Reserves. This item is used in combination with number of shares (NOSH) to determine book value per share in the event that item 1308 is not available.</td>
</tr>
<tr>
<td>392</td>
<td>(TA)</td>
<td>Total Assets.</td>
</tr>
<tr>
<td>625</td>
<td>(NI)</td>
<td>Earned for Ordinary. This is the net profit arrived at after deducting tax, minority interest and preference dividends, but before any post-tax as reported extraordinary items, allocation to reserves other than untaxed reserves and post tax disclosed extraordinary items.</td>
</tr>
<tr>
<td>1308</td>
<td>(B)</td>
<td>Book value per share. Calculated on an issued basis, using that portion of share capital and reserves (excluding preference capital) minus intangibles attributable to the issue, divided by the year-end number of shares in that issue. It is adjusted for subsequent rights and scrip issues.</td>
</tr>
<tr>
<td>NOSH</td>
<td>NOSH</td>
<td>Number of shares</td>
</tr>
<tr>
<td>F1MN (FY1)</td>
<td>F1MN (FY1)</td>
<td>FY1. The earnings-per-share forecasts one-year-ahead taken from I/B/E/S.</td>
</tr>
<tr>
<td>F2MN (FY2)</td>
<td>F2MN (FY2)</td>
<td>FY2. The earnings-per-share forecasts two-year-ahead taken from I/B/E/S.</td>
</tr>
<tr>
<td>LTMN (LTG)</td>
<td>LTMN (LTG)</td>
<td>Long-term growth (LTG). I/B/E/S consensus long-term earnings growth estimate</td>
</tr>
<tr>
<td>Beta (β)</td>
<td>Beta (β)</td>
<td>Annual average of monthly betas provided by Datastream. Beta is calculated based on monthly observations extending over 5 years i.e. 60 months and for each of the proceeding 60 months return on security (Rj) is calculated for every security and regressed against market rates (Rm).</td>
</tr>
<tr>
<td>Rf</td>
<td>Rf</td>
<td>Risk free rate. Government bond rates for both US and UK were used over the sample period for US and UK sample</td>
</tr>
<tr>
<td>Rm</td>
<td>Rm</td>
<td>Market rates. The return for FTSE 500 and S&amp;P 500 obtained from Datastream were used as a proxy for market return for UK and US sample respectively.</td>
</tr>
<tr>
<td>Price</td>
<td>(p)</td>
<td>Market stock price. Calculated as an average of 8 weeks around year-end.</td>
</tr>
</tbody>
</table>

Table 5.1: DataStream data items.
5.3 The Hypothesis and the Regression

Based on the literature and the above discussion, the Residual Income valuation model (the F&L (1998) model) is expected to perform well against the market stock prices in our sample and therefore, it will be hypothesised:

\[ H_0 : \text{There is no relation between the Residual Income valuation and stock market prices.} \]

\[ H_1 : \text{The defined Residual Income model can, to a significant degree, capture cross-sectional variation in market stock prices of the sample over a significant period of time.} \]

\[ H_0 \] will be accepted if the results of the data analysis for both UK and US data lead to the conclusion that there is no relationship between Residual Income and the market stock prices.

The model to be tested is:

\[ P_{it} = \alpha_t + \beta V_{it} + u_{it} \quad (1.1) \]

Where \( P \) is the market stock price and \( V \) is the valuation using the RI model (the F&L model). The definitions of these variables are presented in the following section.

Data are analysed using panel data as well as cross-sectionally. 'PC-Give' is used as statistical software. The results of the both panel and cross-sectional data for both UK and US markets are presented below.

5.4.1 Panel Data Analysis

The sample is initially analysed by grouping all data from FTSE 500 in UK market and all data from S&P 500 in US market separately. Then two separate regressions
for both UK and US data were carried out using the criteria presented in the following table:

**Table 5.2: Criteria Used in Analysing Panel Data**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>P</td>
<td>Market Price</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>V</td>
<td>Value calculated using Residual Income see definitions of the variables on this chapter.</td>
</tr>
<tr>
<td>Intercept</td>
<td>Fixed</td>
<td>The basic value (intercept) for each company depends on organisation-specific issues such as balance sheet size. As a result, the intercept is defined as fixed, whereby for each individual pool member the intercept is estimated.</td>
</tr>
<tr>
<td>White heteroskedasticity Covariance</td>
<td>Yes</td>
<td>Controls covariance matrix on the error term for heteroskedasticity.</td>
</tr>
</tbody>
</table>

Table 5.3 shown below provides the statistical properties of the dependent and independent variables:
Table 5.3: UK and US Descriptive Statistics in “Pence” and in “Dollars” Respectively, “P” and “V”

<table>
<thead>
<tr>
<th></th>
<th>UK descriptive data</th>
<th></th>
<th>US descriptive data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P (Pence)</td>
<td>V (Pence)</td>
<td>P ($)</td>
<td>V ($)</td>
</tr>
<tr>
<td>Mean</td>
<td>292.0012</td>
<td>216.0230</td>
<td>27.87653</td>
<td>16.50428</td>
</tr>
<tr>
<td>Median</td>
<td>238.9323</td>
<td>165.9100</td>
<td>24.98500</td>
<td>11.97354</td>
</tr>
<tr>
<td>Maximum</td>
<td>1383.5000</td>
<td>1785.11</td>
<td>126.57500</td>
<td>241.3767</td>
</tr>
<tr>
<td>Minimum</td>
<td>6.650000</td>
<td>1.789292</td>
<td>0.190000</td>
<td>0.367880</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>214.4708</td>
<td>191.0904</td>
<td>17.96594</td>
<td>16.85531</td>
</tr>
</tbody>
</table>

P is the market stock price. V is the value per share calculated using the F&L (1998) RI model.

Table 5.4: Correlation Matrix between V and P

<table>
<thead>
<tr>
<th></th>
<th>UK Correlation Matrix</th>
<th>US Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>.715*</td>
<td>V</td>
</tr>
</tbody>
</table>

. Correlation is significant at 1% level (2-tailed).

P is the market stock price. V is the value per share calculated using the F&L (1998) RI model.

The results of the pooled data fixed effect cross section regression, using the above-described criteria for both UK and US, are shown in Table 5.5 below:

Table 5.5: UK and US results

\[
P_t = \alpha_i + \beta V_t + u_t \tag{1.1}
\]

<table>
<thead>
<tr>
<th></th>
<th>UK Results</th>
<th>US Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.551697</td>
<td>0.489471</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.05706</td>
<td>0.017357</td>
</tr>
<tr>
<td>t-Statistic</td>
<td>9.67</td>
<td>28.20</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.779622</td>
<td>0.766467</td>
</tr>
<tr>
<td>Chi^2-Statistic</td>
<td>93.48</td>
<td>28.2034</td>
</tr>
<tr>
<td>Prob (Chi^2-Statistic)</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

The above table shows a high correlation between the market prices and the values estimated using the residual income model, with R-squared (goodness of fit) 77.16% and 76.46% for both UK and US respectively and significance level based on Chi^2-

---

27 We use non-parametric test as well, specifically, we use Spearman's correlation as this test is used by Frankel and Lee (1998) in order to compare our results with their results. For more information please refer to appendix 5.A. We also used random effect models and the results are comparable.
statistic of 99% for both UK and US respectively. This implies that, for the UK market, more than 77% of the cross sectional variation in stock prices can be explained by the value calculated using the residual income model. For the US market, it means that more than 76% of the cross sectional variation in stock prices can be explained by the value calculated using the residual income model. The above results are in line with F&L (1998), who find that the model explains more than 70% of the cross-sectional variation in stock prices.

As a consequence, the above results of panel data analysis are statistically significant enough to reject the null hypothesis, which states that there is no relationship between residual income (F&L, 1998 version) and the stock market prices. This leads to accept the alternative hypothesis, which states that the residual income model (F&L, 1998, version) can to a significant degree, capture cross-sectional differences in stock market performance of companies used in this sample and over a significant period of time. In addition, the results also reveal that there is no significant difference in this respect between the UK and the US markets.

5.4.2 Pooling the Data Cross-Sectionally

In this section, the data will be analysed cross-sectionally only. This will enable us to look at the behaviour of the model year by year, as the same sample size will be used in all different years. The results on a year-to-year basis, using the Least Squares Method, for both UK and US data are presenting in the following tables:

---

28 We also transform the original data logarithmically and the results are almost the same. For more information, please refer to table 5.6 in appendix 5.B.

29 We also compared the F&L model with a naïve book value model i.e. how much of the variation in market stock price does book value explain. The results are very similar to the F&L results and are presented in appendix 5.C.
It is clear that R-squared is higher for panel data than for cross-sectional data. The reason is because panel data captures both time-series and cross-sectional influences which the cross-sectional analysis is unable. On the other hand, although the value of R-squared differs between the years, these differences, except for the last two years in US data, are not very significant and can largely be explained by market dynamics and overreaction to shocks in these particular years. The results are controlled for heteroscedasticity using the White Consistent Coefficient Covariance. On the other hand, unlike panel data, cross-sectional results have revealed some differences between the UK and US markets. Therefore, in order to examine the differences in R-squared between UK and US in the above table, the model will be tested for each industry i.e. a regression will be carried out for each portfolio in both the UK and US samples.

Table 5.8 below provides R-Squared for each portfolio for UK companies, with P1 standing for Chemical Industry, P2 Engineering, P3 Food and Hotel, P4 Media, P5 Retail, P6 Transportation, P8 Electronic, P9 Services, and P10 construction industry.
Note that the number of companies in P7, Utility industry, is only 5 and therefore this industry has been excluded.

Table: 5.8 Cross-Sectional Results R-squared for UK Portfolio.

<table>
<thead>
<tr>
<th>Year</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.66</td>
<td>0.92</td>
<td>0.82</td>
<td>0.95</td>
<td>0.82</td>
<td>0.60</td>
<td>0.69</td>
<td>0.74</td>
<td>0.52</td>
</tr>
<tr>
<td>1992</td>
<td>0.63</td>
<td>0.26</td>
<td>0.86</td>
<td>0.91</td>
<td>0.76</td>
<td>0.58</td>
<td>0.64</td>
<td>0.88</td>
<td>0.73</td>
</tr>
<tr>
<td>1993</td>
<td>0.41</td>
<td>0.11</td>
<td>0.71</td>
<td>0.64</td>
<td>0.82</td>
<td>0.77</td>
<td>0.68</td>
<td>0.80</td>
<td>0.68</td>
</tr>
<tr>
<td>1994</td>
<td>0.91</td>
<td>0.27</td>
<td>0.64</td>
<td>0.77</td>
<td>0.87</td>
<td>0.83</td>
<td>0.83</td>
<td>0.91</td>
<td>0.78</td>
</tr>
<tr>
<td>1995</td>
<td>0.93</td>
<td>0.76</td>
<td>0.86</td>
<td>0.39</td>
<td>0.89</td>
<td>0.51</td>
<td>0.84</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>1996</td>
<td>0.90</td>
<td>0.72</td>
<td>0.78</td>
<td>0.56</td>
<td>0.83</td>
<td>0.68</td>
<td>0.82</td>
<td>0.70</td>
<td>0.92</td>
</tr>
<tr>
<td>1997</td>
<td>0.81</td>
<td>0.63</td>
<td>0.76</td>
<td>0.33</td>
<td>0.86</td>
<td>0.63</td>
<td>0.84</td>
<td>0.54</td>
<td>0.92</td>
</tr>
<tr>
<td>1998</td>
<td>0.82</td>
<td>0.50</td>
<td>0.52</td>
<td>0.28</td>
<td>0.88</td>
<td>0.49</td>
<td>0.57</td>
<td>0.53</td>
<td>0.91</td>
</tr>
<tr>
<td>1999</td>
<td>0.60</td>
<td>0.41</td>
<td>0.76</td>
<td>0.82</td>
<td>0.72</td>
<td>0.38</td>
<td>0.53</td>
<td>0.34</td>
<td>0.85</td>
</tr>
<tr>
<td>2000</td>
<td>0.63</td>
<td>0.51</td>
<td>0.55</td>
<td>0.34</td>
<td>0.87</td>
<td>0.78</td>
<td>0.26</td>
<td>0.44</td>
<td>0.73</td>
</tr>
<tr>
<td>2001</td>
<td>0.42</td>
<td>0.61</td>
<td>0.62</td>
<td>0.51</td>
<td>0.81</td>
<td>0.71</td>
<td>0.46</td>
<td>0.12</td>
<td>0.60</td>
</tr>
<tr>
<td>2002</td>
<td>0.97</td>
<td>0.83</td>
<td>0.72</td>
<td>0.43</td>
<td>0.86</td>
<td>0.78</td>
<td>0.81</td>
<td>0.30</td>
<td>0.71</td>
</tr>
</tbody>
</table>

The above results are controlled for heteroskedasticity using the White Consistent Coefficient Covariance.

Table 5.9 below provides R-Squared for each portfolio for US companies, with P1 standing for Chemical Industry, P2 Engineering, P3 Food and Hotel, P4 Media, P5 Retail, P6 Transportation, P7 Utility, P8 Computer, P9 Miscellaneous, and P10 Telecom Equipment and P11 for Pharmaceutical industry.
Table: 5.9 Cross-Sectional Results R-squared for US Portfolio.

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.37</td>
<td>0.58</td>
<td>0.95</td>
<td>0.93</td>
<td>0.51</td>
<td>0.59</td>
<td>0.49</td>
<td>0.96</td>
<td>0.57</td>
</tr>
<tr>
<td>1992</td>
<td>0.52</td>
<td>0.45</td>
<td>0.93</td>
<td>0.95</td>
<td>0.55</td>
<td>0.66</td>
<td>0.42</td>
<td>0.50</td>
<td>0.54</td>
</tr>
<tr>
<td>1993</td>
<td>0.69</td>
<td>0.38</td>
<td>0.82</td>
<td>0.97</td>
<td>0.54</td>
<td>0.66</td>
<td>0.57</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>1994</td>
<td>0.80</td>
<td>0.15</td>
<td>0.81</td>
<td>0.99</td>
<td>0.56</td>
<td>0.63</td>
<td>0.45</td>
<td>0.79</td>
<td>0.85</td>
</tr>
<tr>
<td>1995</td>
<td>0.56</td>
<td>0.30</td>
<td>0.86</td>
<td>0.96</td>
<td>0.83</td>
<td>0.53</td>
<td>0.36</td>
<td>0.62</td>
<td>0.86</td>
</tr>
<tr>
<td>1996</td>
<td>0.51</td>
<td>0.29</td>
<td>0.61</td>
<td>0.99</td>
<td>0.88</td>
<td>0.62</td>
<td>0.23</td>
<td>0.87</td>
<td>0.73</td>
</tr>
<tr>
<td>1997</td>
<td>0.66</td>
<td>0.06</td>
<td>0.62</td>
<td>0.99</td>
<td>0.65</td>
<td>0.97</td>
<td>0.48</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td>1998</td>
<td>0.37</td>
<td>0.26</td>
<td>0.31</td>
<td>0.95</td>
<td>0.14</td>
<td>0.64</td>
<td>0.42</td>
<td>0.84</td>
<td>0.68</td>
</tr>
<tr>
<td>1999</td>
<td>0.44</td>
<td>0.37</td>
<td>0.12</td>
<td>0.92</td>
<td>0.21</td>
<td>0.44</td>
<td>0.06</td>
<td>0.31</td>
<td>0.79</td>
</tr>
<tr>
<td>2000</td>
<td>0.33</td>
<td>0.12</td>
<td>0.34</td>
<td>0.89</td>
<td>0.70</td>
<td>0.60</td>
<td>0.14</td>
<td>0.68</td>
<td>0.79</td>
</tr>
<tr>
<td>2001</td>
<td>0.62</td>
<td>0.27</td>
<td>0.40</td>
<td>0.96</td>
<td>0.28</td>
<td>0.86</td>
<td>0.05</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>2002</td>
<td>0.50</td>
<td>0.08</td>
<td>0.40</td>
<td>0.98</td>
<td>0.43</td>
<td>0.89</td>
<td>0.07</td>
<td>0.77</td>
<td>0.85</td>
</tr>
</tbody>
</table>

The above results are controlled for heteroskedasticity using the White Consistent Coefficient Covariance.

Examining the above two tables, P5 (UK retail industry) data has the highest significance with average R-squared equal to 0.83. The utility industry in UK data was excluded since it cannot be taken into consideration, as the number of companies in the portfolio was very small (5 companies only). The total average R-squared for UK portfolio for 12 years is equal to 0.64. On the other hand, P4 i.e. Media industry is the highest significant industry in the US data with average R-squared equal to 0.96. P11 (Pharmaceutical industry) together with P10 were excluded (added later to miscellaneous) as the number of companies in these portfolios were also very small. The total average R-squared for US portfolio for 12 years is equal to 0.60.

It can be concluded that the residual income model performs very well in explaining the variation in market stock prices for all UK portfolios except for Utility where there were insufficient observations in order to make a decision. The main weakness of the residual income model for the US portfolio was in the Pharmaceutical industry (later added to miscellaneous). The reason why the residual income model performs poorly in the Pharmaceutical industry might be that the number of observations is very small. In addition, such companies usually invest heavily in R&D, which is classified as an intangible asset, and it could be difficult for those intangibles to be
valued by the market. In addition, many take-overs in the Pharmaceutical industry have taken place in the past few years.

5.4.3 Conclusions from Both Panel and Cross-Sectional Data

The results of both panel and cross-sectional data analysis are statistically significant, yet the UK data set provides more significant results than US in cross-sectional results. This could be due to the fact that the residual income model performs poorly for the pharmaceutical industry in the US sample, while there was no pharmaceutical industry in UK portfolios. From the US portfolios’ results after the pharmaceutical industry was removed R-squared gets approximates to the UK figure. Moreover, these differences also could be due to the characteristics of each UK and US market. For instance the Utility industry represents a very high proportion in US data (41 companies) while there were only 5 companies in the UK data set. In addition, cross-sectional data analysis has shown that R-squared for US data set has dropped to a low figure of 0.135261 (table 5.7) in year 2001. This might be attributed to the accidents that happened in US on 11 September 2001 when most of the stock prices went down.

In general, the above results from both panel and cross-sectional data analysis are statistically significant and therefore the null hypothesis, which states that there is no relation between residual income and the stock market prices, is rejected. This will lead us to accept the alternative hypothesis that states that the residual income model can, to a significant level; explain the cross-sectional variation in stock prices of the sample firms and over a significant period of time.

5.5 Determination of Performance

In the previous section, the test results of panel data and cross-sectional analysis show that the majority of the cross-sectional variations in the levels of stock market prices of both UK and US firms can be explained by the value calculated using the Residual
Income model specified in Equation (1). However, despite the fact that the RI model can be employed for performance measurement, it is still a measure of value creation over the time interval concerned (12-years) and as stated previously, the model estimates the value of equity at a particular point in time, but does not directly indicate the value created during a certain period. For this reason, this model can be operationalised as a performance measure in two ways:

- From an accounting point of view, the first difference in Equation (1) page 110, namely the first difference between $V_t$ and $V_{t-1}$ divided by $V_t$ and adjusted for dividends, represents firm performance. Therefore, the determination of firms performance using Equation (1) will lead to the following Equation (2):

$$\text{Performance}_t = V_t - V_{t-1} + \text{Dividend}_t \quad (2)$$

The reason we add back the dividend is that it is part of the value created during the period; it is a component of performance that is not included in $V_t$.

- The value of equity at a particular point in time ($V_t$) is largely dependent on the value of equity in the previous period ($V_{t-1}$). The actual value created in a certain period of time consists of earnings (including dividend) over and above the cost of capital employed, which is the Residual Income Component (RIC) of the Residual income model as defined in Equation (1) page 110 and is represented in Equation (3):

$$\text{RIC}_{(t,t+1)} = \left( \frac{\text{FROE}_t - r_c}{1 + r_c} \right) B_t + \left( \frac{\text{FROE}_{t+1} - r_c}{(1 + r_c)^2} \right) B_{t+1} \quad (3)$$

The RIC from the accounting point of view represents a performance measure. The concern of this study is performance rather than valuation and it can be argued that the performance measure can manifest itself in the residual income component rather than in book value and continuing value. These are important in valuing companies and

---

30 As explained in Chapter four, although the simple residual income, which will also be tested and compared with the above two performance measures in appendix 5.13, can be used as a performance measure. However, the above two performance measures (Equation 3 and 4) are expected to outperform the simple residual income as they include analysts’ earnings forecasts; evidently, analysts’ earnings forecasts contain more value-relevant information than is reflected in simple historical residual income model.
not in calculating their performance. Furthermore, we scaled Equation 3 by total assets per share to account for firm size.

Accordingly, two regression analyses will be carried out to test for Equation (2) and Equation (3) in the next section. The results from both regressions will be compared in order to identify the performance measurement that will be used in the next chapter.

### 5.5.1 The Hypotheses and the Regressions

If the RIC model presented in Equation (3) scaled by total assets per share is validated, then it will be used instead of the full one presented in Equation (2) to measure the performance as a dependent variable in next section. Accordingly, the following will be hypothesised:

- **H02**: There is no relation between first difference of the Residual Income valuation (F&L model) and the performance of the stock market prices (first difference) corrected for dividends.

- **H12**: The first difference of the defined Residual Income model can, to a significant degree, capture the first difference of the cross-sectional variation in market stock prices of the sample over a significant period of time.

- **H03**: There is no relation between the Residual Income Components (the modified R&L model) and the performance of the stock market prices (first difference) corrected for dividends.

- **H13**: The defined Residual Income Components can, to a significant degree, capture cross-sectional differences in the performance of the market stock prices of the sample over a significant period of time.

The above two hypotheses will be tested and compared in order to choose a performance measure from one of them. These two hypotheses will be tested using the two following models respectively:

\[
\Delta P_t = \alpha + \beta CV_t + u_t \tag{2.1}
\]
\[ AP_t = \alpha_t + \beta RIC_t + u_t \quad (3.1) \]

Equation (2.1) is to test H2, which is based on Equation (2) while Equation (3.1) is to test H3, which is based on Equation (3). In Equation (2.1), \( \Delta P \) is the market stock prices performance calculated as the first difference i.e. \( \Delta P_t = P_t - P_{t-1} + D_t \) where \( D_t \) is the dividend at time \( t \). \( CV \) is the performance measure calculated based on first difference of the RI i.e. \( CV_t = CV_t - CV_{t-1} + D_t \). Both \( CV \) and \( P_t \) were divided by the base year to account for size and percentage; \( D \) is included in both equations for reasons explained above. However, in Equation (3.1), \( \Delta P \) is the same as in Equation (3.1); however, it is for two periods of time i.e. \( P_t, P_{t+1} \) and RIC is the residual Income component as represented in Equation (4) and scaled by total assets per share. All the above variables are defined as before.

5.5.2 Results

<table>
<thead>
<tr>
<th>Table 5.10: Descriptive Statistics for Performance Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>( \Delta P ) / Pence</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
</tbody>
</table>

The above table reveals that the mean of the performance measured by the RIC in both the US and the UK markets is very close. The standard deviation is very close as well. On the other hand, the above table reveals that \( P \) in the US market is higher than the one in the UK market. The criteria in Table 5.2 used in the previous results will be reused here. The results of the panel data fixed effect cross section regression from Equation (2.1) for both UK and US are shown in the tables below:
Table 5.11: Results from Equation (2.1) First Difference:

\[ \Delta P_t = \alpha_i + \beta CV_t + u_t \] (2.1)

<table>
<thead>
<tr>
<th></th>
<th>UK Results</th>
<th>US Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.014</td>
<td>0.011</td>
</tr>
<tr>
<td>t-Statistic</td>
<td>20.61</td>
<td>21.13</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>T-Statistic</td>
<td>2.98</td>
<td>3.75</td>
</tr>
<tr>
<td>Prob (T-Statistic)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The results of the panel data fixed effect cross section regression from Equation (3.1) for both UK and US are shown in the tables below:

Table 5.12: Results from Equation (3.1) RIC:

\[ \Delta P_t = \alpha_i + \beta RIC_t + u_t \] (3.1)

<table>
<thead>
<tr>
<th></th>
<th>UK Results</th>
<th>US Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.54</td>
<td>0.51</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.103</td>
<td>0.097</td>
</tr>
<tr>
<td>t-Statistic</td>
<td>5.26</td>
<td>5.19</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.22</td>
<td>0.29</td>
</tr>
<tr>
<td>T-Statistic</td>
<td>3.01</td>
<td>4.06</td>
</tr>
<tr>
<td>Prob (T-Statistic)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 5.11 shows the correlation between the market stock prices performance and the performance estimated using Equation 2.1. R-squared (goodness of fit) is 22% and 26% for both the UK and US markets respectively. These results are also significant at T-statistic of 99% significance level for both UK and US respectively. This implies that, for the UK market, more than 22% of the cross sectional variation in stock prices performance can be explained by the performance calculated using the RI first difference. For the US market, it means that more than 26% of the cross sectional variation in stock price performance can be explained by the performance measurement calculated using the RI first difference. These results lead us to reject the null hypothesis and to accept the alternative.
On the other hand, Table 5.12 shows also a correlation between the market stock price performance and the performance estimated using Equation (3.1). R-squared (goodness of fit) is 22% and 29% for both the UK and US markets respectively. These results are also significant at T-statistic of 99% significance level for both UK and US respectively. This implies that, for the UK market, more than 22% of the cross sectional variation in stock prices performance can be explained by the performance calculated using the RIC in Equation (3.1). For the US market, it means that more than 29% of the cross sectional variation in stock price performance can be explained by the performance calculated using the RIC in Equation (4) scaled by total assets per share. These results lead us to reject the null hypothesis and to accept the alternative.

Although the above results are statistically significant, allowing us to conclude that the majority of the changes in the market stock prices can be explained by either Equation 2.1 or 3.1. However, the results also reveal significant differences between the full RI valuation model (Equation (2)) and the two-year performance model i.e. Equation 2.1 and 3.1. This can be attributed to the fact that the full model explains the cross sectional variation of the levels in stock prices rather than the first differences; this does not mean that either the RIC or the first difference model can explain the cross sectional variation of the levels in stock prices the same as the full RI valuation model does.

### 5.5.3 Comparison between the two Performance Models

<table>
<thead>
<tr>
<th></th>
<th>Equation 2.1 based on the Residual Income first difference</th>
<th>Equation 3.1 based on the Residual Income Component Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK Results</td>
<td>US Results</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>2.98</td>
<td>3.75</td>
</tr>
<tr>
<td>Prob (F-Statistic)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

31 We also test for the simple residual income model in appendix 5.D. However, as expected, the above two performance measures reveal better and significant results.
The above table reveals that both models are statistically significant and they lead us to reject the null hypothesis. In addition, $R^2$ from both models are very close. This means that both the RIC and the first difference models can be used in explaining performance measure for the US and the UK markets.

5.6 Conclusions

The first purpose of this chapter was to test the hypothesis that states that the defined Residual Income model can, to a significant degree, capture cross-sectional variation in stock market price of the sample over a significant period of time. The results from both panel and cross-sectional data support the above stated hypothesis. This leads to the conclusion that the residual income model represented in Equation (1) captures cross-sectional variation in stock market prices and therefore, it can be used to determine firm performance in the next chapter. However, as stated previously in this chapter, Equation (1) measures equity value and not firm performance. Therefore, two ways were proposed in order to determine firm performance. The first is represented in Equation (2.1). From an accounting point of view, the first difference in Equation (1) namely first difference between $V_t$ and $V_{t-1}$ (adding back dividend at time $t$) represents firm performance. The second is represented in Equation (3.1) and it is based on the argument that the actual value created in a certain period of time consists of earnings (including dividend) over and above the cost of capital employed, which is the Residual Income Component (RIC) of the Residual income model as defined in Equation (1). Furthermore, these two performance models i.e. Equation (3) and Equation (4) were empirically tested and compared. The results revealed that the RIC model represented in Equation (3.1) is very close the first difference in equation (2.1) in capturing stock price returns performance. This leads to the conclusion that either Equation (2.1) or Equation (3.1) can be used as a performance measure in the next chapter to test the relationship between firm performance, gearing and over-investment problem.
Chapter Six

The Relation between Over-investment, Gearing and Firm Performance

6.1 Introduction

Modigliani and Miller (1958) demonstrated that in a world of perfect capital markets, investment, financing and dividend decisions are independent. However, during the last decades, the empirical evidence has shown that these decisions are in practice interdependent. Accordingly, some theories that explain this evidence have been developed. These theories are based on capital-market imperfections; in particular, with respect to investment decisions, the existence of asymmetric information between the main stakeholders is the foremost imperfection.

The role of the asymmetry of information in investment decisions has its primary basis in the theoretical work of Jensen and Meckling (1976), Myers (1977) and Myers and Majluf (1984). The first two papers emphasise the consequences of the existence of post-contract asymmetric information between shareholders and bondholders, while the paper of Myers and Majluf (1984) emphasises the role of the pre-contract asymmetric information between current and prospective shareholders. All the above-mentioned papers show that information asymmetries may lead to some profitable investment projects not being undertaken.

A second foundation in the study of inefficient investment decisions is the work of Jensen (1986). His work, starting from the hypothesis of the existence of asymmetric information between managers and shareholders, introduces the so-called problem of over-investment as a basic argument of his free cash flow theory. According to this
theory, there can be negative NPV investment projects that end up being undertaken, which is known as the over-investment problem, and will be tested empirically in this chapter. More precisely, we test whether a certain level of gearing can, to a certain degree, mitigate the over-investment problem and thus increase firm performance.

The remainder of this chapter is organised as follows: section two summarises the existing literature, while section three develops the hypotheses tested in this chapter. Section four illustrates the model that will be used in this study to capture the over-investment problem, followed by the definition of the variables in section five. Section six presents descriptive statistics and discusses the regression results. The testing of the hypotheses will be carried out through a three-stage process. In the first stage, the effect of gearing on over-investment will be tested, using a logit model since the dependent variable is dichotomous. In this stage, the results of using book value of gearing, market value of gearing, industry-adjusted gearing and classification of gearing (high, middle and low gearing) will be presented as well as the results from both linear and non-linear logit models, followed by a conclusion on the relationship between gearing and investment decisions. The second stage will test the relationship between over-investment and performance. The third stage will test at the relationship between over-investment, gearing and performance. The final stage links the two previous stages together and draws conclusions from the three stage results. Section seven summarises the main findings of this chapter and their implications.

6.2 The Relationship between Firm Performance, Debt and Over-investment

Agency theory argues that an appropriate ownership structure is a key factor in motivating managers’ investment decisions consistent with shareholder objectives. While the above control variable addresses the role played by the firm’s ownership

---

32 Initially ownership structure was to be included in this study; however, we were unable to get such information as Datastream does not provide ownership information, so it was decided to omit this variable.
structure in aligning the incentives of managers and shareholders, it does not consider the role of management's discretion in investment decisions. Agency theory contends that a firm's investment decisions are also influenced by the size of the internally available funds. According to the theory, given abundant resources for investment activity managers have an incentive to grow the firm beyond its optimal size. Managers derive benefits from investment (the larger the organisation, the greater the economic and political power of the top management team), so that their utility is increasing with the level of investment. Managers are empire-builders and continue to choose investment projects even after all positive net present value projects have been taken up.

The idea that a firm might systematically over-invest originated with Jensen (1986), who argues that shareholders must find a mechanism to induce managers to disgorge excess cash flow\(^{33}\) rather than to over-invest. Jensen (1986) gives the following example from the oil industry in 1970s. Given increases in oil prices and slow industry growth, managers used internal funds to diversify outside the industry in some of the worst performing acquisitions of the 1970s. By contrast, firms with smaller available cash flows are less likely to engage in value-destroying venture activity due to the scrutiny of external financial markets (Jensen 1986). Moreover, as observed by Agency Theory, ownership and control have become separated in large corporations because of the dilution in equity positions. This situation has provided an opportunity for professional managers, as those in control, to act in their own interests. The central issue for agency theory is how to resolve the conflict between owners and managers over the control of corporate resources (Jensen, 1986) by using contracts that seek to allocate decision rights and incentives.

Managers have a number of incentives to pursue growth-oriented strategic options. The larger the organisation, the greater the economic and political power of the top management team, and the greater the ability of the organisation to marshal the resources necessary to deal effectively with its competitive and social environment (Morck et al, 1989). As a generalisation, it can be said that growth from investment does lead to increases in the shareholders' wealth. However, the concern is that too many of the activities associated with that growth in investment and hence with

\(^{33}\) Excess Cash Flow, as this term is used by Jensen (1986), is the cash flow that remains after all positive NPV projects are undertaken.
increasing the size of organisations is motivated not by a desire for maximising shareholders' wealth, but by opportunities for the self-aggrandisement of management (Jensen and Ruback, 1983). The problem, according to Jensen (1986), is how to control managers contractually so that they will return profits or excess cash flow to investors rather than invest funds in projects with returns below the cost of capital.

The contractual device suggested by Agency Theory to accomplish the transfer of wealth from the organisation to the investors is debt creation (Jensen, 1986). Debt provides a means of bonding managers' promises to pay out future cash flows rather than investing in wealth-destroying ventures. It also provides the means for controlling opportunistic behaviour by reducing the cash flow available for discretionary spending. Top managers' attention is then clearly focussed on those activities necessary to ensure that debt payments are made. Companies failing to make interest and principal payments can be declared insolvent. This use of debt as a disciplinary tool makes survival the central issue for all concerned.

The over-investment idea has gained much support in the literature. For instance, Jong and Veld (2001) find that Dutch managers avoid using debt (they want to avoid the disciplining role of debt) because debt disciplines managers and prevents them from over-investing. However, the market reaction shows that this over-investment behaviour is recognised as they found a negative relationship between firms with FCF (defined as operating income minus taxes, interest an dividends, divided by the market value of equity) and low growth opportunities (measured by Tobin's Q) and stock market returns, and the authors interpret this as evidence of over-investment.

Moreover, the relation between firm performance and managerial ownership has been used in previous work to support the over-investment theory. Morck et al (1988) estimate a piecewise linear relation between managerial ownership and firm performance. They find that firm performance is increasing in managerial ownership-for-ownership levels below 5 percent or over 25 percent but decreasing in ownership-for-ownership levels between 5 and 25 percent of the firm. They interpret their results as evidence that managers make investment decisions that entrench them in their positions for ownership in this range; entrenchment reduces firm performance. Many subsequent papers (McConnell and Servaes (1990), Himmelberg, Hubbard, and Palia (1999) and Palia (2001)) have conducted similar analyses with mixed results. Other
support for the over-investment theory comes from Jensen (1993), who provides illustrative calculations of the destruction of shareholders' value in a number of the world's largest corporations.

However, recent papers by Aggarwal & Samwick (2003) and Hadlock (1998) have concluded with differing results. For instance, Aggarwal and Samwick (2003) estimate a non-linear regression between investment, firm performance and managerial incentives (measured as pay-performance sensitivities and representing incentives provided by direct ownership of stock and stock options for each top management team). They find that firm performance is increasing in incentives for all levels of incentives. Investment is also increasing in incentives. These two results jointly imply that agency problems do not necessarily lead firms to over-invest. More exactly, incentives increase both performance and investment. Accordingly, the argument of Jensen that firms with FCF over-invest will be tested in this chapter. The main hypothesis to be tested is: debt can, to a certain degree, mitigate the over-investment problem, and this should be related to better firm performance.

6.3 Developing the Model

Jensen (1986) explains that firms may engage in projects with negative net present values (NPV), because managers aim for firm growth. Increasing firm growth, as explained before, is not motivated by a desire for maximising shareholders' wealth only, but by opportunities for self-aggrandisement of management. He points out that the reaction to a securities issue depends on the expected purpose of raising the capital and the type of security that is issued. As managers have an incentive to increase the size of the firm, even with negative NPV projects, the firm's performance is determined by the value created through positive NPV investments. Over-investment (i.e. investment in negative NPV projects) cannot be directly observed. Previous studies have used free cash flow together with low growth opportunities measured by Tobin's Q as a proxy for the over-investment problem. Furthermore, previous studies
assumed that a company that has free cash flow together with low growth opportunities suffers from an over-investment problem\textsuperscript{34}.

Although having FCF together with low growth opportunities could lead to over-investment this does not necessarily mean that a company has an over-investment problem. In addition, according to Myers and Majluf (1984) the firm could face asymmetric information in which investors are less informed about the value of the firm than insiders. In this case, equity may be mis-priced. If the firm has a project with a positive NPV available, the under-pricing of the equity may be greater than the value of the project, and it will be passed over. This problem (usually called the under-investment problem) can be overcome by using a less risky form of financing i.e. internal funds. For this reason, firms prefer maintaining financial slack\textsuperscript{35} in order to have internal funds available for valuable projects.

In this study, we introduce a simple way to proxy for the over-investment problem. Logically, firms with low growth opportunities should not invest heavily; if they do, they might over-invest. Tobin’s Q will be used to capture growth opportunities and industry average investment expenditure will be used to capture the expected investment level. Companies with low growth opportunities should not invest above their industry average. In other words, companies may be in a state of over-investment if they have low growth opportunities and are at the same time investing above their industry average. Therefore, the combination of a low Tobin’s Q with capital expenditure (scaled by total assets) in excess of the industry average will be used as a proxy to capture the existence of over-investment. For example, if companies have capital expenditure in excess of the industry average together with low growth opportunities as measured by Tobin’s Q, this indicates that companies belonging to this category are experiencing over-investment.

Accordingly, if firms issue debt, and the effect of it is to discipline the over-investment behaviour, a positive relationship will be expected between the firm’s

\textsuperscript{34} Despite the difficulties in finding good proxies for FCF, this method has been employed by many researchers such as Devereux and Schiantarelli (1990), Lang, et al (1996), Lamont (1997), Chen and Ho (1997), Chakraborty, Kazarosian and Trahan (1999), Jong and Veld (2001), Miguel and Pindado (2001) and Del Brio.

\textsuperscript{35} The difference between financial slack and free cash flow is important. Financial slack is cash, liquid assets and unused borrowing power (Myers and Majluf, 1984). Free cash flow is the cash that remains after all positive NPV projects are undertaken.
performance and level of debt, as the latter is expected to mitigate the over-investment problem. Debt is valuable for firms with low growth opportunities (as measured by Tobin's Q), and also capital expenditure in excess of the industry average, because it commits managers to pay out cash in the future, thereby reducing the free cash flow at their disposal for empire-building investments.

6.4 Definition of the Variables

The dependent variable is a dichotomous variable (dummy variable $D_i$) and is designed to capture the over-investment problem, i.e. $(D_i) = 1$ if Tobin's Q is below 1 and capital expenditure (scaled by total assets) is above industry average and $D_i = 0$ otherwise. For companies with Tobin's Q below 1 the market considers that they have no growth opportunities and therefore one would not expect their investment to exceed the industry average. As a consequence, the over-investment problem can be captured in companies with Tobin's Q below 1 i.e. low growth opportunities, and at the same time a level of capital expenditure is above the industry average.

**Tobin's Q**: This is defined as the ratio of the market value of the firm to the replacement cost of its assets (e.g. Lindenberg and Ross, 1981; Lang and Stulz, 1994). This measure of Q is likely to capture growth opportunities. If Q is greater than one, the inference is that the market values current the firm's assets (including intangibles) more highly than it would in their next best alternative use, their replacement cost. There are a number of methods used in the previous literature to capture Tobin's Q. Perfect and Wiles (1994) argue that the Lindenberg and Ross method for estimating Q values is difficult to use because it relies upon companies' reporting replacement cost estimates which are not available all the time. On the other hand, Chung and Pruitt (1994) and Perfect and Wiles (1994) summarise and compare these measures. Chung and Pruitt (1994) define an approximate Q by assuming that the book value of assets is equal to their replacement cost and find that this simple measure is highly correlated with Lindenberg and Ross' more theoretically correct model. In this study, replacement costs for most of the companies in our sample are not available and
therefore we will assume that the book value of assets is equal to their replacement cost. Accordingly, Tobin’s Q equals:

\[
\frac{BVTA + MVE - BVE}{BVTA}
\]

BVTA (the book value of total assets, Datastream item 392) is the sum of Cash and due from banks, total investments, net loans, customer liability on acceptance, investment in unconsolidated subsidiaries, real estate assets, net property, plant and equipment and other assets.

MVE (market value of equity, Data-stream item MV)) is the share prices multiplied by the number of ordinary shares in issue. The amount in issue is updated whenever new tranches of stock are issued or after a capital change. In order to avoid the effect of financial statement publication on the stock price formation, the average stock price of four weeks prior and 4 weeks after the balance sheet date was used to calculate market value of equity.

BV (book value of equity, Data-stream item 305) is the equity share capital and reserves of the company i.e. ordinary share capital, other equity capital, share premium account and reserves.

**Capital Expenditure:** (Payments-fixed assets, Data-stream item: 1024). The amount specified by the company under capital expenditure. It is cash paid for tangible fixed-assets (expenditure on machinery, equipment, plant, vehicles and buildings) during the year including payments deferred from previous.

**Capital Expenditure, Industry averages:** based on the panel median of US and UK data respectively. We also calculated the industry average based on the means and the results are comparable. Capital expenditure industry average is also calculated based on portfolios. We calculated the mean and median of each industry in our sample and compared each company’s capital expenditure to its industries’ mean and median, and the results are almost the same.
**Capital Gearing**\(^36\): It should be noted that there is contradictory evidence about the use of the market value of equity or book value of equity. Some of the previous studies have used the book value of equity, arguing that although the theory of capital structure suggests that debt ratios should be measured in market value terms management prefers to use the book value. Myers (1977) argues that market values incorporate the present values of future growth opportunities. Debt issued against these values can distort future real investment decisions. Accordingly, in this study capital gearing will be measured using two ratios: the book value of gearing and the market value of gearing defined as follows:

**Book value of Capital Gearing (BG):** is the ratio of total debt plus preference share capital to total assets, where: Total debt is the total of all long and short-term borrowings (Data-stream, item: 1301). Preference share capital is the capital, which has a fixed dividend and does not participate further in the company’s profits (Data-stream, item: 306). Total Assets is the book value of total assets and is defined above (Data-stream, item: 392).

**Market value of Capital Gearing (MG)**\(^37\): this is the ratio of total debt plus preference share capital to total debt plus preference share capital plus the market value of the common stock, where market value of equity capital share is defined as the number of shares outstanding (data-stream, item: NS), multiplied by the stock prices (data-stream, item: P) at the balance sheet date. In order to avoid the effect of financial statement publication on the stock price formation, the average stock price of four weeks prior and 4 weeks after to the balance sheet date was used to calculate market value of equity.

**Firm Performance:** firm performance will be measured by the residual income model component as represented in Equation (3.1) scaled by total assets per share (tested in Chapter 5).

---

\(^36\) We have also used book and market values of debt in calculating the debt-to-equity ratio

\(^37\) One could argue that this ratio is the proportion of debt in the capital structure, rather than the debt-to-value ration. For this reason, we have also used the ratio of total debt plus preference share capital to total assets plus the difference between market value of common stock and book value of common equity. For more information about the results, refer to appendix 6.8.
**Control Variables**: three control variables are used: the natural logarithm of total sales (LnTS), the dividend payout ratio (DPU) and the stock price volatility. Total sales controls for firm size, and a positive and significant coefficient for size is suggestive of managers pursuing size and growth strategies that are not necessarily in the principal's best interest. DPU is quite relevant to over-investment issues; if companies have high dividends it signals one of two things: either they have abundant current cash balances and expect future cash flows or future investment opportunities are not extant. DPU is equal to dividends paid divided by profits after interest and taxes but before dividends. High stock price volatility necessitates careful evaluation of a firm’s investment decisions. The reason for including this control variable is that stock price volatility encourages managers to reduce gearing in order to over-gearing. Stock Price volatility is a measure of a stock’s average annual price movement to a high and low from a mean price for each year.
Table: 6.1 Definition of the Data-stream Variables

<table>
<thead>
<tr>
<th>Item</th>
<th>Not.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>392</td>
<td>BVTA</td>
<td>BVTA. The book value of total assets. This item is used to calculate Tobin’s Q as well as to determine capital gearing.</td>
</tr>
<tr>
<td>305</td>
<td>BVE</td>
<td>Equity Capital and Reserves. This item is used in combination with BVTA to determine Tobin’s Q.</td>
</tr>
<tr>
<td>MV</td>
<td>MV</td>
<td>Market value of equity. Is used together with 392 and 305 to determine Tobin’s Q.</td>
</tr>
<tr>
<td>1024</td>
<td>CE</td>
<td>Payment Fixed Assets. This item is used to calculate capital expenditure.</td>
</tr>
<tr>
<td>1301</td>
<td>TD</td>
<td>Total debt including long and short-term debt</td>
</tr>
<tr>
<td>306</td>
<td>PC</td>
<td>Preference capital. This item is added to item 1301 to determine company’s overall debt.</td>
</tr>
<tr>
<td>NS</td>
<td>NS</td>
<td>The number of shares used in the calculation of market value of equity.</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>Is the market stock price used in combination with NS to calculate the market value of equity.</td>
</tr>
<tr>
<td>Ln(104)</td>
<td>Ln(TS)</td>
<td>Natural logarithm of total sales. This item is used as a control variable for firm’s size.</td>
</tr>
<tr>
<td>DPU</td>
<td>DPU</td>
<td>Dividend Payout Ratio is the ratio of dividends paid to the profit after interests and taxes but before dividends. This item is used as a control variable.</td>
</tr>
<tr>
<td>MG</td>
<td>MG</td>
<td>Market value of gearing calculated as TD (1301) + PC (306) to TD (1301) + PC (306) + MV.</td>
</tr>
<tr>
<td>BG</td>
<td>BG</td>
<td>Book value of gearing calculated as TD (1301) + PC (306) to BVTA (392).</td>
</tr>
<tr>
<td>TQ</td>
<td>TQ</td>
<td>Tobin’s Q is the BVTA (392) + MV + BVE (305) to BVTA (392)</td>
</tr>
<tr>
<td>E002</td>
<td>Vol.</td>
<td>Stock Price volatility is a measure of a stock’s average annual price movement to a high and low from a mean price for each year. This item is used as a control variable.</td>
</tr>
</tbody>
</table>

Source: Data-stream data items.

6.5 The Hypotheses

In this chapter, testing for firm performance, debt and over-investment will be carried out through several steps. Firstly, theoretically, debt can mitigate the over-investment problem to a certain degree, and the following will therefore be hypothesised:
There is no relationship between gearing and over-investment

There is a negative relationship between gearing and over-investment

The above hypothesis will be tested using a discrete choice model since the dependent variable in this case is a dichotomous variable that takes a value of 1 if there is over-investment and 0 otherwise. H₀ will be accepted when the results of the data analysis lead to the conclusion that there is no relation between debt and over-investment. It should be noted that even if the null hypothesis is rejected, and a negative relationship between gearing and over-investment has been found, this should not be considered as enough evidence to support the over-investment argument. Over-investment should be linked to performance at the end to make the final conclusion. Accordingly, testing the second and the third hypotheses (listed below) is very important in order to draw the final conclusion.

If firms over-invest, this should negatively influence performance and it would therefore be hypothesised that:

H₀: There is no relationship between over-investment and firm performance

H₁: There is a negative relationship between over-investment and firm performance.

Theoretically and intuitively, when debt mitigates the over-investment problem this should lead to better performance. Introducing the interaction term between debt and over-investment would capture the link between the first and the second hypothesis and hence the next hypothesis would be as follows:

H₀: the interaction between the presence of over-investment and the level of gearing is not related to firm performance

H₁: the interaction between the presence of over-investment and the level of gearing is related to better firm performance
As mentioned above, the dependent variable in the first hypothesis is a dichotomous variable and takes the value of 1 for companies experiencing an over-investment problem (as defined above) and of 0 otherwise. On the other hand, the explanatory variable is a quantitative or continuous variable. Accordingly, the above stated hypothesis will be tested and analysed by estimating model (1) below using a logit model:

\[ D_i = \alpha_i + \beta G_i + u_i \quad (1) \]

Where \( D \) is the over-investment variable which takes a value of 1 if there is over-investment and 0 otherwise. \( G \) is either the market or book value of gearing. \( \beta \), \( \alpha \) and \( u \) are the slope coefficient, intercept and residual terms respectively. \( \beta \) represents the effect of gearing on the over-investment problem and it is expected to be significantly negative in order to reject the null hypotheses. Before we proceed, it should be noted that in models where \( D \) is qualitative, our objective is to find the probability of something happening. Hence qualitative response regression models are often known as probability models. This is because the conditional expectation of \( D_i \) given \( G_i \), \( E(D_i|G_i) \) can be interpreted as the conditional probability that the event will occur given \( G_i \), i.e. \( Pr(D_i=1|G_i) \). Thus, in our case, \( E(D_i / G_i) \) gives the probability of a company experiences an over-investment problem where the level of gearing is the given amount of \( G_i \).

Table 2.6 shown below provides statistical properties of the sample used for both the UK and US markets:

---

38 We also use Probit and the results are very similar.
### Table 6.2: UK and US Descriptive Statistics

#### UK Descriptive Data

<table>
<thead>
<tr>
<th></th>
<th>Vol.</th>
<th>LnTS</th>
<th>RIC</th>
<th>DPU</th>
<th>CE/TA</th>
<th>BG</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>32.0</td>
<td>13.52</td>
<td>6.17</td>
<td>47.4</td>
<td>0.017</td>
<td>20.92</td>
<td>21.36</td>
</tr>
<tr>
<td>Median</td>
<td>30.33</td>
<td>13.54</td>
<td>4.32</td>
<td>44.8</td>
<td>0.007</td>
<td>19.84</td>
<td>18.42</td>
</tr>
<tr>
<td>Maximum</td>
<td>47.86</td>
<td>9.15</td>
<td>169.76</td>
<td>100</td>
<td>0.97</td>
<td>68.17</td>
<td>92.75</td>
</tr>
<tr>
<td>Minimum</td>
<td>13.42</td>
<td>18.61</td>
<td>-80.15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>8.91</td>
<td>1.54</td>
<td>14.64</td>
<td>20.24</td>
<td>0.05</td>
<td>12.61</td>
<td>16.54</td>
</tr>
</tbody>
</table>

#### US Descriptive Data

<table>
<thead>
<tr>
<th></th>
<th>Vol.</th>
<th>LnTS</th>
<th>RIC</th>
<th>DPU</th>
<th>CE/TA</th>
<th>BG</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>29.56</td>
<td>15.50</td>
<td>7.098</td>
<td>36.08</td>
<td>0.068</td>
<td>26.76</td>
<td>22.60</td>
</tr>
<tr>
<td>Median</td>
<td>27.18</td>
<td>15.55</td>
<td>2.973</td>
<td>32.92</td>
<td>0.057</td>
<td>27.04</td>
<td>17.33</td>
</tr>
<tr>
<td>Maximum</td>
<td>86.63</td>
<td>19.19</td>
<td>166</td>
<td>305.60</td>
<td>0.57</td>
<td>78.58</td>
<td>91.16</td>
</tr>
<tr>
<td>Minimum</td>
<td>12.75</td>
<td>9.87</td>
<td>-90.21</td>
<td>-281.52</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>10.32</td>
<td>1.3</td>
<td>15.83</td>
<td>36.96</td>
<td>0.046</td>
<td>14.42</td>
<td>18.52</td>
</tr>
</tbody>
</table>

CE/TA is the capital expenditure scaled by total assets. BG (the book value of gearing) is the ratio of total debt plus preference capital to total assets. MG (the market value of gearing) is the ratio of total debt plus preference capital plus market value of the common stock. RIC is the performance measured by the Residual Income Components scaled by total assets per share (RIC Chapter 5 Equation 4). LnTS is the natural logarithm of total sales. DPU is the dividend out ratio. Vol. is the market stock price volatility.

### Table 6.3: Correlation Matrix between Over-investment and Gearing Variables.

#### UK Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>BG</th>
<th>BG²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>-.050*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BG²</td>
<td>-.081***</td>
<td>.937***</td>
<td></td>
</tr>
<tr>
<td>BG³</td>
<td>-.089***</td>
<td>.826***</td>
<td>.966***</td>
</tr>
</tbody>
</table>

#### US Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>BG</th>
<th>BG²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>-.126***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BG²</td>
<td>-.169***</td>
<td>.954***</td>
<td></td>
</tr>
<tr>
<td>BG³</td>
<td>-.184***</td>
<td>.871***</td>
<td>.974***</td>
</tr>
</tbody>
</table>

*** Correlation is significant at 1% level (2-tailed). **. Correlation is significant at 5% level (2-tailed). *. Correlation is significant at 10% level (2-tailed). D is the dummy variable takes a value of 1 if there is over-investment 0 otherwise. BG, BG² and BG³: book value of gearing squared book value of gearing and cubed book value of gearing respectively.
Table 6.4: Correlation Matrix between Performance and Interaction Variables of Over-investment and Gearing.

<table>
<thead>
<tr>
<th>UK Correlation Matrix</th>
<th>US Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIC</td>
<td>DBG</td>
</tr>
<tr>
<td>DBG</td>
<td>-.153*</td>
</tr>
<tr>
<td>DBG$^2$</td>
<td>-.176*</td>
</tr>
<tr>
<td>DBG$^3$</td>
<td>-.168*</td>
</tr>
<tr>
<td>V</td>
<td>-.089*</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.01 level (2-tailed). RIC is the firm performance measure based on the residual income components calculated in Chapter 5. DBG is the interaction variables between D (over-investment) and book value of gearing. DBG$^2$ and DBG$^3$ are the squared and cubed of the interaction variables.

The above Table (6.2) reveals that the mean and median for both UK and US samples are analogous, indicating that the distribution of the gearing variable is very close to normality. It also reveals that companies in the US sample are using more debt than companies in the UK market in our sample. For example, the US mean and median are around 27% where the UK mean and median are around 20%. The mean of Market Value of Gearing (MG) in the UK data is slightly higher than the mean of Book Value of Gearing (BG) and the standard deviation of MG is also higher than the standard deviation of the BG. This is unforeseen since one could expect that the BG to be higher than the MG. This indicates that book value of assets in the UK was higher than the market value of assets. The reason might be that firms in the UK are allowed to restate the non-current assets to current values while firms in the US are not. This might suggest that firms in the UK restate their assets at high current values to make their gearing appear smaller. Furthermore, BG and MG are not strictly comparable as MG is the proportion of debt in the capital structure not to debt-to-value ratio.

On the other hand, the mean of BG in the US data is higher than the mean of MG. This is expected since US firms are not allowed to restate their assets to their current values in their balance sheets. The standard deviation of MG for both the UK and the US are higher than for the BG. This is because the market value of equity is subject to the fluctuation of stock prices movement on the market. The above table also
reveals that capital expenditure in the US market is higher than the in UK market. This is could be because the US firms are larger than in the UK. It also reveals that capital expenditure scaled by total assets in the UK market is more skewed than the capital expenditure scaled by total assets in the US market. These differences, although they are not very significant, might affect the results in both UK and US data.
6.6.1 Over-investment and Gearing Assuming Linearity

The following table presents the estimated output of the above model (1) page 140:

Table 6.5: Modelling D by Logit Results from UK and US Data:

<table>
<thead>
<tr>
<th></th>
<th>UK Data</th>
<th>US Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>BG</td>
</tr>
<tr>
<td>Coefficient</td>
<td>-0.0725</td>
<td>0.00347</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.09856</td>
<td>0.00404</td>
</tr>
<tr>
<td>t-value</td>
<td>-0.735</td>
<td>0.858</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.462</td>
<td>0.391</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1072.62296</td>
<td>-1071.09176</td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-1072.992</td>
<td>-1072.992</td>
</tr>
<tr>
<td>Test Chi²2</td>
<td>0.73774</td>
<td>3.8001</td>
</tr>
<tr>
<td>Chi²2 prob.</td>
<td>0.3904</td>
<td>0.0512</td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.38840176</td>
<td>1.38642346</td>
</tr>
</tbody>
</table>

C is the constant. BG is the book value of gearing. MG is the market value of gearing.

Table 6.5 reveals the results from the UK and US markets. Modelling D using BG for UK data, the coefficient of gearing has an unexpected positive sign but is statistically insignificant. This implies that gearing would have no significant effect on the over-investment problem for the UK market. This result is in conflict with our hypothesis, which states that there is a negative relationship between gearing and over-investment. However, modelling D using MG the coefficient is negative and significant at the 5.2% level. Unlike the BG, MG would have a significant effect on the over-investment. Hence, using MG leads us to reject the null hypothesis in this study at the 5.2% level of significance and to accept the alternative, which states that there is a negative relation between gearing and over-investment. MG has different results.

39 We have also used book and market value of debt-equity ratio and the results were significant with the expected negative sign. Furthermore, we have used the ratio of total debt plus preference share capital to total assets plus the difference between market value of common stock and book value of common equity as another way of measuring the market gearing. For more information about these results, refer to appendix 6.A and 6.B respectively.
from BG. This could be explained by the fact that firms in the UK restate their assets (this is why the mean and median for BG and MG are very close) in a way that their asset value may be higher to make their gearing appear smaller. More analysis will be carried out in the next sections to analyse these differences.

Table 6.5 also reveals that for US data the coefficients of both BG and MG are statistically significant at 1% level and both have the expected negative sign. The finding suggests that the higher the gearing, the lower the probability of over-investment. This result, using both BG and MG, leads us to reject the null hypothesis at 1% level of significance in the sample used in this study (over 12 years) and to accept the alternative which states that there is a negative relationship between gearing and over-investment. This finding is in line with Jensen’s (1986) free cash flow theory; he argued that debt provides a means of bonding managers’ promises to pay out future cash flows rather than investing in wealth-destroying ventures. Debt also provides the means for controlling opportunistic behaviour by reducing the cash flow available for discretionary spending. Top managers’ attention is then clearly focused on those activities necessary to ensure that debt payments are made. Companies failing to make interest and principal payments can be declared insolvent. This use of debt as a disciplinary tool makes survival the central issue for all concerned.

6.6.2 Over-investment and Gearing Assuming Linearity (Industry Effect)

The above results do not control for industry effects, which may be an important element in determining both firms’ investment and their capital structures. We follow Lang, Ofek and Stulz (1996) and estimate the above results while adjusting the gearing variables by their industry medians. Capital expenditure average is also calculated based on each industry. Each company’s median capital expenditure was compared with its industry median to see if it is investing above or below its industry median. Industries are as defined before (Chapter 4: methodology). The results reported in the following section are industry-adjusted:
Table 6.6: Modelling D by Logit-Industry-Adjusted Results from UK and US Data:

<table>
<thead>
<tr>
<th></th>
<th>UK Data</th>
<th>US Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>BG</td>
</tr>
<tr>
<td>Coefficient</td>
<td>-0.0792</td>
<td>0.0698</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.09529</td>
<td>0.07601</td>
</tr>
<tr>
<td>t-value</td>
<td>-0.831</td>
<td>0.919</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.406</td>
<td>0.358</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1072.56329</td>
<td>-1068.31514</td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-1072.987</td>
<td>-1072.987</td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>0.84675</td>
<td>9.343</td>
</tr>
<tr>
<td>Chi^2 prob.</td>
<td>0.3575</td>
<td>0.0022</td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.38832466</td>
<td>1.3828361</td>
</tr>
</tbody>
</table>

C is the constant. BG is the book value of gearing industry-adjusted. MG is the market value of gearing industry-adjusted.

Table 6.6 reveals very similar results to the above reported industry-unadjusted results in Table 6.5. BG for the UK sample is still insignificant with unexpected positive sign, while for the US sample BG is still significant with the expected negative sign. On the other hand, MG is negatively significant for both US and UK data.

6.6.3 Over-investment and Gearing Assuming Linearity (Gearing Classification)⁴⁰

The above reported results do not demonstrate whether the effect of gearing on over-investment is a continuous function given that the degree of gearing (i.e. high or low gearing) may have different implications for over-investment. Therefore, the degree of gearing might play an important role in deterring the over-investment problem since in theory the higher the level of debt the greater the pressure on managers to perform. Hence, UK and US data were split into four quartiles. UK and US data

⁴⁰ We tested the relation between over-investment and debt-equity ratio classification in appendix 6.C.
were sorted based on median\textsuperscript{41} gearing from top to bottom. Then each UK and US data set was split into four quartiles, the top and the bottom quartile representing the top and bottom gearing. The two middle quartiles are middle gearing. The following tables present the estimated model (1):

**Table 6.7: Modelling D by Logit Results from UK Data\textsuperscript{42}:**

<table>
<thead>
<tr>
<th></th>
<th>Top Gearing</th>
<th>Middle Gearing</th>
<th>Bottom Gearing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Std. Error</td>
<td>t-value</td>
</tr>
<tr>
<td>Constant</td>
<td>0.794714</td>
<td>0.3460</td>
<td>2.30</td>
</tr>
<tr>
<td>BG</td>
<td>-0.018636</td>
<td>0.009883</td>
<td>-1.89</td>
</tr>
<tr>
<td>Constant</td>
<td>0.782961</td>
<td>0.1895</td>
<td>4.13</td>
</tr>
<tr>
<td>BG</td>
<td>-0.02836</td>
<td>0.008554</td>
<td>-3.32</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.564323</td>
<td>0.1526</td>
<td>-3.70</td>
</tr>
<tr>
<td>BG</td>
<td>-0.003952</td>
<td>0.01189</td>
<td>-0.332</td>
</tr>
</tbody>
</table>

\textsuperscript{41} We also sorted the leverage based on mean gearing and the results are almost the same.

\textsuperscript{42} We also split market value of leverage into three percentiles, high, middle and bottom the results are very similar with MG = -0.0111526 and significant at 10\% level in top percentile, = -0.0131714 and significant at 10\% level in the middle percentile and = -0.00421284 with t-probability of 0.703 which is insignificant in the bottom percentile.
Table 6.8: Modelling D by Logit Results from US Data:

<table>
<thead>
<tr>
<th></th>
<th>Top Gearing</th>
<th></th>
<th>Middle Gearing</th>
<th></th>
<th>Bottom Gearing</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>BG</td>
<td>Constant</td>
<td>BG</td>
<td>Constant</td>
<td>BG</td>
</tr>
<tr>
<td>Coefficient</td>
<td>1.50779</td>
<td>-0.045189</td>
<td>0.879735</td>
<td>-0.01487</td>
<td>-0.060333</td>
<td>0.03218</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.4870</td>
<td>0.01133</td>
<td>0.1939</td>
<td>0.006862</td>
<td>0.1418</td>
<td>0.01139</td>
</tr>
<tr>
<td>t-value</td>
<td>3.10</td>
<td>-3.99</td>
<td>4.54</td>
<td>-2.17</td>
<td>-0.425</td>
<td>2.83</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.030</td>
<td>0.671</td>
<td>0.005</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-370.445456</td>
<td></td>
<td>-731.484751</td>
<td></td>
<td>-357.45498</td>
<td></td>
</tr>
<tr>
<td>Baseline Log-likelihood</td>
<td>-378.9204</td>
<td>-373.8471</td>
<td>-361.5907</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>16.95</td>
<td>4.7247</td>
<td>8.2715</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi^2 prob.</td>
<td>0.0000</td>
<td>0.0297</td>
<td>0.0040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.32072857</td>
<td>1.32877672</td>
<td>1.36157189</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Table 6.8 it is clear that as leverage increases the probability of over-investment in the US data decreases. In the top firms, gearing G is negatively and statistically significant at the 1% level. Moving to the middle firms, gearing G also has the expected negative sign and statistically significant at the 5% level. However, in the bottom firms, gearing G is statistically significant with unexpected positive sign. This suggests that a low level of gearing may increase financial slack. This also suggests that the higher the gearing level the higher the pressure on managers to invest less. These results are also in the same line with Table 6.5 where the overall gearing was significantly negative at the 1% level. This means that US data, using top and bottom gearing, leads us to reject the null hypothesis and accept the alternative, which states that there is a negative relationship between gearing and over-investment. It should be noted that the coefficient of the gearing in tables 6.5 and 6.8 appears to be small. This is because debt is only one of the factors to mitigate the over-investment problem.

43 We also split market value of leverage into four quartiles, 1 top, 2 middle and bottom, the results are very similar with MG = -0.0430527 and significant at 1% level in top percentile, = -0.00620971 and significant only at 5% level in the middle percentile and = 0.00710119 with t-probability of 0.654 which is insignificant in the bottom percentile.
On the other hand, results from Table 6.7 reveal that for UK data, unlike Table 6.5 and US result, gearing is significantly negative at the 6% in the top percentile. It is significantly negative at the 1% level in the middle percentile and non-significant in the bottom percentile. Based on these results, the relationship between gearing and over-investment in the UK market is most significant in the middle percentile, while in the top percentile, this relationship becomes less significant (6%) and it becomes negligible in the bottom percentile. The UK market does not follow the same pattern as the data in the US market where the higher the debt the lower the probability of over-investment. This could be related to sample size and the average of leverage in each market. Moreover, this method of splitting the sample into three sub-samples may be subjective and a regression model with non-linear variables, which will be run in the next paragraph, may provide more robust results.

On the other hand, Table 6.7 reveals that both top and middle BG are significant, while the bottom BG is insignificant, taking into consideration that the total BG for the UK was insignificant. Accordingly, another reason for observing non-significance for the total BG in the UK sample could be because of the bottom percentiles. For this reason, we ran the same logit model using only top and middle BG percentiles for the UK sample. The results are shown in the following table:

**Table 6.9: Modelling D by Logit-Results from UK Top and Middle Percentile Data:**

<table>
<thead>
<tr>
<th></th>
<th>Constant</th>
<th>BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.628161</td>
<td>-0.0175112</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.1453</td>
<td>0.005328</td>
</tr>
<tr>
<td>t-value</td>
<td>4.32</td>
<td>-3.29</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-795.958746</td>
<td></td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-801.4267</td>
<td></td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>10.936</td>
<td></td>
</tr>
<tr>
<td>Chi^2 prob.</td>
<td>0.0009</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.37106314</td>
<td></td>
</tr>
</tbody>
</table>

BG is the book value of top and middle percentiles in the UK sample.
Table 6.9 shows that BG for the UK sample is negatively significant at the 1% level when we excluded the bottom percentile. This means that the bottom percentile could be the reason for driving the total UK BG to be insignificant. As stated above, the method of splitting the sample into three sub-samples may be subjective since the number of observation in each sub-sample is not identical and a regression model with non-linear variables, which will be run in the next paragraph, may provide results that are more robust.

6.6.4 Over-investment and Gearing Assuming Non-Linear Variables

The working hypothesis of most empirical studies in this area is that certain financial policies, in particular, the use of debt financing, should uniformly either boost or hinder firm performance (see for example, Demsetz and Lehn, 1985). However, later analysis has considered non-linear forms (see Morck et al., 1988; McConnell and Servaes, 1990, 1995; Kole, 1995). The non-linear analysis follows from the possibility that influences the relation between a firm’s performance and gearing, that there is an optimal level of gearing, and hence, gearing could increase or decrease a firm’s performance depending on whether a low or high degree of gearing is used. For instance, trade-off theory predicts that the relation between gearing and value is parabolic and takes an “n” shape i.e. gearing increases firm’s value until a certain degree, and then gearing starts to decrease the firm’s value.

Moreover, results from the 3-step linear regression analysis suggest that the relation between gearing and over-investment could also be non-linear. According to our initial results, the gearing coefficient has shown different signs when we split our sample into three sub-samples based on bottom, middle and top gearing. This suggests that the relation between gearing and over-investment may be cubic. Accordingly, below is the estimated non-linear regression using the cubic equation:

\[ D_{it} = \alpha_i + \beta G_{it} + \beta G_{it}^2 + \beta G_{it}^3 + u_{it} \]  \hspace{1cm} (2)

We also tested the non-linear relationship between over-investment and the debt-equity ratio. However; the results revealed that the debt-equity ratio cannot be expressed in a non-linear form with over-investment, as explained in appendix 6.D.

We used a cubic equation because we believe that there may be only two turning points. There could be only one turning point if the results reveal that G^3 is insignificant.
Where $D$ is a dichotomous variable described above and $G$ is the gearing defined above, $G^2$ and $G^3$ are the squared and cubed, respectively, of the gearing variable.

Table 6.10: Modelling $D$ by Logit: the Cubed UK and US Results$^{46}$

<table>
<thead>
<tr>
<th></th>
<th>UK results</th>
<th>US results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cons.</td>
<td>BG</td>
</tr>
<tr>
<td>Coeff.</td>
<td>-0.6034</td>
<td>0.0826</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.1724</td>
<td>0.0256</td>
</tr>
<tr>
<td>t-value</td>
<td>-3.50</td>
<td>3.22</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1063.93067</td>
<td></td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Chi$^2$</td>
<td>18.122</td>
<td></td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.37975539</td>
<td>1.3358042</td>
</tr>
</tbody>
</table>

Cost. is the constant. $BG$ is the book value of gearing. $BG^2$ and $BG^3$ are the squared and cubed book value of gearing.

From Table 6.10, it seems that levels of gearing exert a significant influence on over-investment for both the UK and US data. Moreover, the results provide evidence for the possible functional form of the relationship between gearing and over-investment problem. More specifically, the estimated coefficients of $G$, $G^2$ and $G^3$ suggest that over-investment varies at different stages of the gearing level. The estimated coefficient of $G$ is statistically significant at the 1% level in both US and UK data, but with an unexpected positive sign. The positive coefficient of $G$ in both UK and US suggest that, at a low level of gearing, the higher the debt level the greater the probability of having over-investment. The positive coefficient of $G$ suggests that a low level of gearing increases the over-investment not the opposite i.e. a low level of gearing may increase financial slack.

$^{46}$ We run the same regression using MG and the results are almost the same for more information see appendix 6. E.
On the other hand, $G^2$ is also statistically significant in both UK and US data at less than 5% level with the expected negative sign. This means that the more the gearing the lower the probability of over-investment. The negative coefficient of $G^2$ suggests that a high level of gearing does put pressure on managers to reduce the over-investment. On the other hand, the non-significant results for $G^3$ in both UK and US suggests that more gearing (above a certain level) has no further effect on the over-investment problem. Moreover, because $G^3$ is insignificant, this could suggest that the relationship between gearing and over-investment might be quadratic rather than cubic. Accordingly, we estimated the same logit regression using the following quadratic equation:

$$ D_n = \alpha_i + \beta G_i + \beta G_i^2 + u_i $$

(3)

Table 6.11: Modelling D by Logit the Squared UK and US Results

<table>
<thead>
<tr>
<th></th>
<th>UK results</th>
<th>US results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Const.</td>
<td>BG</td>
</tr>
<tr>
<td>Coefficient</td>
<td>-0.45</td>
<td>0.048</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.140</td>
<td>0.012</td>
</tr>
<tr>
<td>t-value</td>
<td>-3.22</td>
<td>3.88</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>1065.10332</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1072.992</td>
<td></td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>15.777</td>
<td></td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.37997845</td>
<td></td>
</tr>
</tbody>
</table>

Const. is the constant. BG and BG^2 are the book value of gearing and the square book value of gearing respectively.

Results from Table 6.11 reveal no differences from Table 6.10, the level of gearing has an impact on over-investment. At low levels of gearing, the higher the gearing, the greater the probability of over-investment. By contrast, at higher levels of gearing, the higher the gearing the lower the probability of over-investment. Calculations carried out on the coefficients of the variables MG and MG^2 for the UK sample reveal
that gearing below 26.67% has no effect on the over-investment problem, from 26.67% gearing starts to mitigate the over-investment problem. On the other hand, for the US sample, gearing becomes effective after 20.5%\(^{47}\). It should be noted that using a non-linear relation between gearing and over-investment has revealed similar results between book value and market value for both US and UK data.

To interpret the results from a logit model meaningfully, the model itself must first fit the data. In other words, the explanatory variables included in the model must be able to explain the response variable significantly better than the model with the intercept only (Liao, 1994). While in classical regression the F-test is used, in a logit model the most commonly used test is the likelihood ratio statistic, which approximately follows the chi-squared distribution. If chi-squared indicates that the model fits the data significantly better than the model with the intercept only, one can move on to the intercept parameter estimates.

For the UK data (table 6.5), chi-squared is statistically insignificant when using a linear book value of leverage but it is significant when we removed the bottom percentile from the sample, and it is statistically significant at the 6% level when using market value of leverage. On the other hand, the chi-squared probability in Table 6.5 (for US data) is zero and therefore one can argue that the explanatory variable included in the model (BG and MG) is able to explain the response variable significantly better than the model with the intercept only. This is also the case when non-linear regression has been used; all chi-squared probabilities are significant. Both market value and book value for the US and UK data reveal similar results. This also suggests that the explanatory variables including in the model are able to explain the response variable significantly better than the intercept only. This leads us to reject the null hypothesis and to accept the alternative, which states that there is a relation between gearing and over-investment.

The next step involves the assessment of predictive efficiency. In classical regression, the most common measure of goodness-of-fit is R-squared. However, in a logit model it would not be right to think of this as the proportion of variance explained by

\(^{47}\) The maximum point of a quadratic function is calculated as follows: Assuming all other variables are constant and denoting MG by \(x\): \(D = 0.048x - 0.00091x^2\). The maximum point is found by differentiating \(y(D)\) with respect to \(x\), letting \(dy/dx = 0\) and solving for \(x\). This implies that \(0.048 - 0.0018x = 0\) this implies \(x = 26.67\). The same calculation for US gives \(x = 20.5\).
the model. Typically, R-squared reported in logistic regression underestimates the proportion of variation explained by the model (Gujarati, 2002). Accordingly, some other criteria for model selection should be used such as Akaike Information Criterion (AIC). The model, which best fits the data is the one with the lowest AIC. In all the above reported models, AIC is around 1.3, which leads to a conclusion that whether the model is estimated using linear regression or non-linear, BG or MG, industry-adjustment or non-adjusted, AIC is almost the same in all circumstances and gives no preference to one estimated model over the other.

All the above non-linear regression results suggest that the higher the level of gearing the lower the probability of over-investment. This leads us to reject the null hypothesis and to accept the alternative, which states that there is a relationship between gearing and over-investment. However, this should not be interpreted as evidence to support the existence of the over-investment problem, as this should be related to the level of performance. Accordingly, to demonstrate the over-investment problem, one should relate this to performance, expecting to find a negative relation between over-investment and performance since over-investment in negative NPV projects destroys firm value. This test will be carried out in the following section.

6.7 Regression Analysis: Over-investment, Gearing and Firm Performance Assuming Non-Linearity

The previous results have revealed that leverage decreases over-investment. However, to make this sufficient to support the over-investment argument one should relate this to performance; in other words, over-investment destroys firm value and reduces performance. Therefore, to demonstrate the over-investment problem, firstly the expected relationship between performance and over-investment has to be negative, and secondly, where debt mitigates the over-investment, this should be related to better performance. Accordingly, two regressions will be carried out in this section. one to test the relationship between performance and over-investment, and the
second to test performance from one side and the interaction variable (leverage and over-investment) from the other side.

6.7.1 Over-investment and Firm Performance

We now proceed to test the relationship between firm performance and over-investment. It is expected that this relationship would be negative since over-investment reduces a firm's performance. The static relation between firm performance and over-investment implicitly assumes that an investment decision can instantaneously have an impact on a firm's performance. However, an investment decision at year \( t \) would probably affect firm's performance in later years. Therefore, in this project, we adopt a more realistic approach recognising that a firm's performance would be affected for years following an investment project.

In this study we use a method which is very similar to the one adopted by Rajan and Zingales (1995). In their method, they estimate a cross-sectional model where their dependent variable is in year 1991 and all the regressors are four-year averages (1987-1990). Averaging the explanatory variable in our case would help us capturing the affect of investment decision on firm's performance through accounting for adjustments and reducing the noise. On the other hand, lagging the explanatory variable one period could help reducing the problem of endogeneity.

In our model, we used the same technique as described above, but we used a 4 year moving average for the explanatory variable and we pooled the data. This is because the total number of firms in Rajan and Zingales (1995) were 2583, which is extremely large compared to our sample (317 firms). This would suggest that a cross-sectional model with many observations could give significant results, while in this case, we used 8 cross-sectional models and the results were mainly insignificant. Accordingly, the following regression will be tested:

\[
Performance_{t+5} = \alpha + \beta D_{\text{average}(t+1:t+4)} + u_{it} \quad (4)
\]

Performance is the performance measured by RIC scaled by total assets per share at time \( t \). \( D \) is the over-investment dummy 4 years moving average, which takes a value
of 1 for over-investment and 0 otherwise. Results from the above regression for both the UK and the US data are presented in Table 6.12 below:

Table 6.12: Modelling Performance and Over-investment the UK and US Results

<table>
<thead>
<tr>
<th></th>
<th>UK Results</th>
<th>US Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>-0.008549</td>
<td>-0.287850</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.02474</td>
<td>0.1753</td>
</tr>
<tr>
<td>t-value</td>
<td>-0.345</td>
<td>-1.64</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td>Wald test</td>
<td>48.78</td>
<td>2.697</td>
</tr>
</tbody>
</table>

*: Significant at 1%, **: Significant at 5%, ***: Significant at 10%

As can be seen from the above table, both results are in the expected direction but insignificant (US data was significant at 11% very close to the 10% level of significance). Next, we use a 5-year performance moving average instead of 4 years. The results from the 5 years moving average are insignificant (worse than the 4 years moving average) but in the expected direction as well.

The lack of significance may be due to losing many observations, since by using 4-years moving average we are losing 3 years, which could explain why they are worse than the 5-years moving average. To increase the number of observations, we combined both UK (132 firms) and US (185 firms) data in the same sample, giving 317 firms. We then use a 4-year moving average over a period of 12 years and at the same time we assigned another Dummy (D2) which takes a value of 1 for UK firms and 0 otherwise. The results are as follows:
Table 6.13: Modelling Performance and Over-investment the Total UK and US Sample:

<table>
<thead>
<tr>
<th></th>
<th>D1</th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>-0.151651</td>
<td>0.338687</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.09155</td>
<td>0.09155</td>
</tr>
<tr>
<td>t-value</td>
<td>-1.66***</td>
<td>3.70*</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>0.68</td>
</tr>
<tr>
<td>Wald test</td>
<td></td>
<td>7.172**</td>
</tr>
</tbody>
</table>

*: Significant at 1%, **: Significant at 5%, ***: Significant at 10%. D1 is the over-investment dummy, D2 is the country dummy.

The combined sample in Table 6.13 reveals that the coefficient of over-investment is significantly negative at the level of 9%. This means that, in line with our hypothesis, over-investment has a negative impact on firms’ performance. This suggests that the null hypothesis, which states that there is no relation between firm’s performance and over-investment, is rejected and we accept the alternative, which states that there is a negative relation between firm’s performance and over-investment. This is expected; according to the over-investment argument, over-investment reduces performance.

6.7.2 Firm Performance, Over-investment and Gearing Assuming Non-Linearity:

Having shown that gearing affects over-investment in the first hypothesis and over-investment decreases firm’s performance in the second hypothesis, it is expected to find the following relationship:

1- At the lower level of gearing (below 26.67% for UK and below 20.5% for US), we have found that the higher the gearing the greater the probability of over-investment. On the other hand, in the second hypothesis, over-investment has a negative impact on firm performance. Therefore, it is expected that at a low level
of gearing the interaction between gearing and over-investment will have a negative impact on firm's performance.

2- At the higher level of gearing (above 26.67% for UK and above 20.5% for US), we have found that the higher the gearing the smaller the probability of having over-investment; high level of gearing mitigates the over-investment problem. Therefore, it is expected that at a high level of gearing, the interaction between gearing and over-investment may be significantly positive.

The above two points lead us to test the non-linearity between the interaction of gearing, over-investment and performance. We use three control variables: the natural logarithm of total sales (LnTS), the dividend payout ratio (DPU) and market stock price volatility. Total sales controls for firm size; a positive and significant coefficient for size is suggestive of managers pursuing size and growth strategies that are not necessarily in the principal's best interest. DPU is relevant to over-investment issues; if companies pay out dividends it signals two things: either they have abundant current cash balances and expect future cash flows, or they have future no investment opportunities. Firms with high stock return volatility should be careful in evaluating their investment decisions. The reason for including stock price volatility as a control variable is that if the stock price is more volatile, careful managers will reduce gearing in order to avoid being over-geared. Accordingly, the relationship between the interaction of gearing and over-investment from one side and performance from the other side will be tested by using the following model:

\[
\Delta P_{t+5} = \alpha_t + \beta_1(D_1G) + \beta_2(D_1G)^2 + \beta_3(D_1G)^3 + \beta_4DPU + \beta_5TS + \beta_6VOI + \beta_8D_2 \] 

\textit{average(t+1+4)} + \epsilon_{it} \tag{6}

The above regression is similar to the Rajan and Zingales (1995) method described above, where \( \Delta P \) is firm's performance, measured by the RIC in chapter (5); (\( D_1G \)), (\( D_1G \)^2) and (\( D_1G \)^3) are the interactions between gearing and over-investment, the squared interaction and the cubed interaction respectively; DPU is the dividend payout ratio; TS is the natural logarithm of total sales; Vol. is the stock price volatility and \( D_2 \) is the country dummy which takes a value of 1 for the UK and 0 for the US. The results from the above model are as follows:
Table 6.14: Modelling Performance and the Integrated Over-investment and Gearing: the Pooled UK and US Results\(^{48}\)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D(_1)G</td>
<td>-0.2616</td>
<td>0.125</td>
<td>-2.09**</td>
</tr>
<tr>
<td>(D(_1)G)(^2)</td>
<td>0.0126</td>
<td>0.0072</td>
<td>1.74***</td>
</tr>
<tr>
<td>(D(_1)G)(^3)</td>
<td>-0.00017</td>
<td>0.00011</td>
<td>-1.53</td>
</tr>
<tr>
<td>DPU</td>
<td>-0.00014</td>
<td>0.00014</td>
<td>-1.05</td>
</tr>
<tr>
<td>Ln(TS)</td>
<td>-0.52</td>
<td>0.693</td>
<td>-0.752</td>
</tr>
<tr>
<td>Vol.</td>
<td>0.0902</td>
<td>0.078</td>
<td>1.16</td>
</tr>
<tr>
<td>D(_2)</td>
<td>1.645</td>
<td>1.224</td>
<td>1.34</td>
</tr>
<tr>
<td>R(^2)</td>
<td></td>
<td>0.649</td>
<td></td>
</tr>
<tr>
<td>Wald Test</td>
<td></td>
<td>19.32*</td>
<td></td>
</tr>
</tbody>
</table>

* : Significant at 1%, ** : Significant at 5%, *** : Significant at 10%

D\(_1\)G is the interaction variable between leverage and over-investment, (D\(_1\)G)\(^2\) is the squared of D\(_1\)G and (D\(_1\)G)\(^3\) is the cubed of D\(_1\)G; DPU is the dividend payout ratio; Ln(TS) is the natural logarithm of total sales; Vol. is the stock price volatility and D\(_2\) is the country dummy variable, which takes a value of 1 for the UK firms and 0 for the US firms.

Table 6.14 reveals that, at low levels of gearing, the coefficient of the interaction variables between gearing and over-investment is statistically significant at the level of 5% with a negative sign. This means that a low level of gearing together with over-investment have a negative impact on performance. This is expected because we have found that at a low level of gearing, over-investment increases i.e. debt does not mitigate the problem; rather debt makes it worse and therefore the problem of over-investment should have a negative impact on firm’s performance. On the other hand, the positive and significant coefficient (significant at the level of 8%) of the squared of the interaction variable between gearing and over-investment suggests that a middle level of gearing does mitigate the over-investment problem and therefore, it is associated with a positive effect on firm performance. Moreover, and in line with the previous results, the non-significant coefficient of (DG)\(^3\) suggests that a very high

\(^{48}\) The turning points of a cubic function can be calculated by differentiating \(P_x\) with respect to the \(x\) variables i.e. (G,G\(^2\),G\(^3\), DG, DG\(^2\) and DG\(^3\), letting \(dP/dX = 0\) and solve for \(X\). to determine whether \(X\) is a maximum or minimum turning point, calculate the value of \(d^2P/d^2X\). If \(d^2P/d^2X > 0\), the turning point is a maximum, if \(d^2P/d^2X < 0\), the turning point is a minimum.
level of gearing has no further effect on firm performance; high gearing does not monitor the over-investment problem. This suggest, and in line with the previous results, that the relationship between gearing, over-investment and performance is quadratic rather than cubic. Moreover, the negative sign of \((DG)^3\) (although it is not significant) could be because gearing reduces the over-investment problem and a very high level of gearing (extreme values) might negate the performance. In addition, the positive and non-significant coefficient of the country dummy suggests that there are no differences across our sample, namely, the UK and the US sample.

The implications of this finding are: 1) the relation between firm's performance from one side and the interaction of gearing and over-investment from the other side can be explained by a non-linear function. 2) The non-linear function has revealed that gearing has a different impact on over-investment at different stages; neither low nor high levels of gearing mitigate the over-investment problem — rather it gets worse. This could be explained by the fact that neither low nor high level of gearing does play a monitoring role in forcing managers not to waste the cash on non-profitable projects, because managers might not feel the pressure of high debt serving payments at low level of gearing. On the other hand, a middle level of gearing plays a monitoring role in reducing the over-investment problem, and that has a positive impact on firm performance. However, a high level of gearing is again associated with a non-significant but negative effect on performance. This could be explained by the fact that a high level of gearing could be associated with financial distress. This finding leads us to reject the null hypothesis and to accept the alternative, which states that a certain level of gearing does reduce the over-investment problem and therefore improves firm performance. This finding is in line with Jensen (1986, 1993) in that it confirms the existence of the free cash flow problem and introduces one factor that could mitigate this problem.

### 6.8 Conclusions

In this chapter, Jensen’s argument (1986, 1993) has been tested; he argues that managers undertake wasteful negative net present value projects because they derive
private benefits from controlling more assets. This is what is called over-investment or empire building. He provides illustrative calculations of the destruction of shareholder value in a number of the world's largest corporations. Several important features of our analysis we believe extend the literature on the empirical work concerning the agency problem, more precisely, Jensen's free cash flow argument. Our analysis, distinct from previous empirical studies, introduces a simple way of identifying the over-investment problem. Second, we use a performance measure that is based on theoretical and empirical work and accounts for the cost of capital. Our analysis incorporates the dynamic nature of the response of a firm's performance to investment decisions and controls for endogeneity through lagging the explanatory variables one period.

Using both FTSE 500 and S&P 500 firms, the results support the principal-agent models based on private benefits of investment (empire building, managerial entrenchment, perquisite consumption). We find that at low levels of gearing (<26% for UK and <20.5% for US), the higher the gearing the greater the over-investment problem, whilst at a higher level of gearing (>26% for UK and >20.5% for US), the higher the gearing the lower the probability of over-investment, except that very high levels of gearing have no impact on the over-investment problem (G3 was insignificant). This could suggest that the relationship between gearing and over-investment is a parabolic function. Moreover, we find that over-investment and firm performance move in opposite directions. At a low level of gearing, where we find that gearing increases the over-investment problem, the interaction between gearing and over-investment has a significant negative impact on firm performance. At a middle level of gearing, where gearing mitigates the over-investment problem, the interaction between gearing and over-investment has a significant positive impact on firm performance. Again, at a high level of gearing, the interaction variable between gearing and over-investment has again a non-significant negative impact on firm performance.

The above findings lead us to reject the null hypothesis and accept the alternative, which states that gearing, to a certain level, can mitigate the over-investment problem. Moreover, our results suggest that only a middle level of gearing mitigates the over-investment problem, since a low level of gearing increases the probability of over-
investment and hence affects a firm's performance negatively. On the other hand in line with the first finding, which suggests that a high level of gearing has no impact on the over-investment problem, the interaction of gearing and over-investment has no impact on firm performance but the sign was negative. This could be because when over-investment is financed by very high gearing, the interaction effect on performance is negative. An important implication of our model is that certain levels of gearing, i.e. a middle level, play a monitoring role in reducing the over-investment problem and hence leads to better firm performance. We have found that debt at a low level increases the over-investment rather than decreases it. This could be because a low level of gearing does not put enough pressure on managers in terms of interest rate and in fact increases a firm's financial slack.
Chapter 7

Summary and Conclusions

7.1 Introduction

In this thesis the relationship between capital structure and firm performance was examined first theoretically and then empirically using panel data from both FTSE 500 and S&P 500 over 12 year period. In these firms, individual shareholders are partial owners, whereas the firms are controlled by professional management teams. In this context, the inherent separation between ownership and control plays an important role. For this reason, we adapt Jensen's (1986) argument in which he explains that firms may engage in projects with negative NPV, because managers pursue growth (which is not necessarily in the interest of shareholders) and derive benefits from investment. Managers are hypothesised to be empire-builders and continue to choose investment projects even after all positive net present value investments have been taken. According to agency theory, this problem arises from incentive misalignment between shareholders and managers. In this thesis, we establish a link between the previous mentioned theory and firms' performance. To illustrate further, thesis adapts Jensen (1986) argument, and empirically tests the proposition that debt disciplines managers and mitigates the over-investment problems and therefore this should lead to better firm performance. In this concluding chapter, we will summarize the preceding chapters and discuss the results of the empirical study. In addition, we will describe the contribution and the limitation of this thesis and topics for further research.

7.2 Summary of Chapters

The capital structure literature starts with the seminal paper by Nobel Laureates M&M (1958). In this paper, under stringent assumptions, capital structure choice is
found to be irrelevant for the value of the firm. Subsequent theoretical studies have investigated the relevance of capital structure by relaxing the assumptions in the MM paper and studying the consequences. Empirical studies provide tests of these theories. This literature is reviewed in Chapter 2. Agency problems play a crucial role in the recent capital structure literature. Information differences and divergences of interests between groups of stakeholders lead to agency problems. With these problems agency costs are associated. Agency problems are closely related to capital structure decisions. On the one hand, capital structure decisions may cause agency problems. On the other hand, capital structure may serve to mitigate agency problems.

Studying listed firms highlights agency problems between managers and shareholders; the shareholders have ownership right in the firm, but the management controls the firm – the managers may have incentives to maximise their own wealth at the expense of the shareholders. This leads to the over-investment problem and is one of several agency theory-based hypotheses that explain capital structure decisions. Other agency-based capital structure hypotheses are derived from alternative agency problems including moral hazard, i.e. between shareholders and bondholders or between insiders and outsiders stakeholders, or agency problems including adverse selection. Aside from the agency-based explanations, tax and bankruptcy-based hypotheses may explain capital structure choice. In the empirical literature, the above-mentioned theories have been reviewed extensively. The majority of the empirical studies are based on data sets of US firms. According to the empirical researchers, most of the theories have been proven relevant in explaining the firms’ capital structure choices.

From Chapter 2, we realised that in Jensen’s (1986) model there are incentives and disciplining roles in using debt. These incentives and disciplining roles in using debt should logically lead to a better performance. Therefore, a theoretical relationship between capital structure and firm performance is established. This theoretical relation was tested empirically in Chapter 5 to conclude whether a certain level of gearing can lead to a better firm performance.

The performance measure is presented in Chapter 3. It can be seen from this chapter that there is substantial criticism continues of traditional performance measures based on ratio analysis (i.e. profitability ratios) that ignore the cost of capital. For this
reason, it was decided to exclude such performance measures and to look at models that take into consideration the spread between the return on capital and the cost of capital. It is clear from the literature that the RI models and the DCF are the most competing models in many of the recent academic books and research papers. Although, these models are preferable to ROA, they are measuring value rather than performance. However, the first difference can be used as a performance measure but with correction for dividend. It was difficult to infer from the prior literature whether one of these approaches is superior to the others. Because of their theoretical equivalence, these models should provide the same valuations when the cash flows are projected consistently to infinity and equivalent discount rates are applied. In practice, however, horizons over which the cash flows can be reasonably projected are limited, and discount rates are estimated subject to errors. These practical considerations cause some academics and practitioners to prefer one valuation model to another.

DCF, as discussed in the literature, looks at the valuation of firms more explicitly than performance, while RI can be used for performance more explicitly than DCF. For example, RI models have an advantage over the DCF model in that they are useful measure for understanding a company’s performance in any single year, while discounted cash flow is not. Since the purpose of this project was to look at the performance measurements rather than valuation, it seemed logical to use RI models, in accordance with their function to evaluate periodic performance, rather than DCF. Therefore, DCF was not used in this project as a performance measure.

On the other hand, as can be seen from the literature, an obvious limitation concerns the absence of a leverage concept in the Ohlson (1995) model. His model was based on risk neutrality, and this does not permit the required return to reflect any compensation for the inherent risk in equity securities or firm specific risk. In addition, it should be noted that certain difficulties would be encountered in implementing the linear information dynamics in Ohlson (1995) model, which frames the stochastic time series behaviour of abnormal earnings. The difficulty is in setting parameter values such as $v_t$ which represent other information in the model. On the other hand, Frankel and Lee (1998) adapted the residual income model in which the cost of capital is risk-adjusted by using an overall assessment of the perceived risk of
the investment as a whole. This leads to a more easily implemented model and therefore leads to preference for the Frankel and Lee's (1998) modified RI model over Ohlson's (1995) model. Chapter 3 concludes that the RI in Frankel and Lee (1998) model will be tested empirically in order to see the possibility of using it as a performance measure.

The first empirical study is presented in Chapter 4. In this study, we tested whether the defined RI (the Frankel and Lee (1998) model) model can, to a significant degree, capture cross-sectional variation in market stock prices of the non-financial FTSE 500 and S&P 500 firms using panel data over a 13-year period of time. Parametric and non-parametric tests were carried out. In parametric tests, we used a fixed effect within group to estimate the regression. The results for both UK and US firms support the hypothesis that RI can explain more than 75% of the cross-sectional variation in market stock prices; this is in line with Frankel and Lee (1998) who find that the model explains more than 70% of the cross-sectional variation in stock prices. We further tested this model cross-sectionally and we controlled for heteroscedasticity. The results from cross-sectional regression also support the hypothesis and support the findings from the panel data regression. We also tested this model for each portfolio in our sample and the results also support the hypothesis.

We used non-parametric tests to compare our results with Frankel and Lee (1998). The results of Spearman's correlation between the RI model and the stock prices were very high (more than 80% for UK and more than 72% for US) and significant at the 1% level. This is in line with Frankel and Lee (1998) who found that the model explains more than 70% of the cross-sectional variation in stock prices.

The second part of Chapter 4 was to determine firm performance based on the RI model, since this model estimates the firm's equity value and not its performance. The latter can be done in one of two ways:

- From an accounting point of view, the first difference in Equation (2), namely the first difference between \( V_{t+1} \) and \( V_t \) divided by \( V_t \) represents firm performance. Therefore, the determination of firm performance using Equation (1) will lead to the following Equation (2, page 122):
Performance, = \(V_t - V_{t-1} + \text{Dividend},\) \hspace{1cm} (2)

The reason we add back the dividend is that it is a component of performance that is not included in \(V_t.\)

- The value of equity at a particular point in time \((V_t)\) is largely dependent on the value of equity in the previous period \((V_{t-1}).\) The actual value created in a certain period of time consists of earnings (including dividend) over and above the cost of capital employed, which is the Residual Income Component (RIC) of the residual income model as defined in Equation (1) and is represented in Equation (3.1), page 122:

\[
\text{Performance}_{t,t+1} = \frac{\text{FROE}_t - r_e}{(1 + r_e)} B_t + \frac{\text{FROE}_{t+1} - r_e}{(1 + r_e)^2} B_{t+1} \quad (3)
\]

The RIC from the accounting point of view represents a performance measure over 2 years. Furthermore, Equation 3.1 was scaled by total assets per share to take into consideration firms' size and to insure that the calculated performance would not be driven by firms' size. As the concern of this study was performance rather than valuation, it can be argued that the performance measure can manifest itself in the residual income component rather than in book value and continuing value. The last two parts (book value and continuing value) are important to value companies and not to calculate their performance.

Accordingly, two regression analyses were carried out to test for Equations (2.1) and (3.1). The results from both regressions lead us to conclude that both Equation (2.1) and Equation (3.1) can partially explain stock price performance (first difference). The two results support the argument that the RIC model or the first difference from equation (1) can explain to a certain level the stock price performance and both can therefore be used as a dependent variable to test the relationship between firm performance, gearing and over-investment problems.

In Chapter 6 we present our second empirical model. Issues of gearing, over-investment and firm performance are studied in order to test the relationship between them. Over-investment plays a key role in the agency theory, and finding a good proxy is important in order to provide a good empirical investigation of the agency
theory. In this thesis, we introduced a very simple way to proxy for over-investment problem. We used low growth opportunities (as measured by Tobin’s Q) together with company’s capital expenditures above its industry median. We used this proxy because first, the difficulty in measuring FCF and second, even if we found a good proxy to measure FCF, FCF together with low growth opportunities does not necessarily lead to an over-investment problem as suggested in the literature. We use publicly listed data of non-financial FTSE 500 and S&P 500 companies. Initially, we use a logit model in order investigate whether gearing mitigates the over-investment. The results from both FTSE 500 and S&P 500 indicate that gearing mitigates the over-investment. As expected we find that the higher the gearing the lower the over-investment. We use both book value of gearing and market value of gearing. Furthermore, we control for industry differences by dividing each company’s book value of gearing and market value of gearing by its industry median; the results reveal no differences from the initial ones.

We next address the issue of linearity i.e. is the effect of gearing on over-investment a continuous function. Different levels of gearing may have different impacts on over-investment. For this reason, we split our sample into four quartiles based on median gearing. We took the top to represent the top level of gearing and the bottom to represent the bottom level of gearing. The two quartiles in the middle are middle gearing. We run the same logit model for each group. As we expected, the results showed that different levels of gearing have a different impacts on over-investment.

This led us to test for a non-linear relationship between gearing and over-investment. Interestingly, we found out that a quadratic relationship between gearing and over-investment was significant. More precisely, we found that at low levels of gearing, (i.e. < 26% for UK firms and < 20.5% for US firms) as gearing increases the probability of over-investment increases as well. On the other hand, at a higher level of gearing (i.e. > 26% for UK firms and > 20.5% for US firms) the higher the gearing the lower the probability of over-investment. The explanation of the above results is that managers at a low level of gearing do not feel the pressure of debt serving payments, and therefore debt may not play a monitoring role. However, results from a higher level of gearing revealed that debt has a negative impact on over-investment,
managers are motivated not to over-invest and therefore, a higher level of debt does monitor managers’ behaviour.

We next tested the relationship between firm performance and over-investment. This is important because the results will reveal whether the over-investment is a problem or not (i.e. value destroying). In order to test the hypothesis we expected a negative relationship between firm performance and over-investment. If the opposite is found, then the whole concept of over-investment as a problem would be questionable. We used a fixed effect regression to test the relationship between firm performance and over-investment. The static relation between firm performance and over-investment implicitly assumes that an investment decision can instantaneously have an impact on firm’s performance. However, an investment decision at year t would probably affect firm’s performance in later years. Therefore, in this thesis, we adopt a more realistic approach recognising that firm’s performance would be affected for years following an investment project. The results provide some support for the hypothesis that over-investment destroy firm value since the coefficient of over-investment is significant and negative only at the level of 9%. This implies, in line with our hypothesis, that over-investment has a negative impact on firms’ performance.

The final step in Chapter 6 was to empirically test the relationship between the interaction variable of gearing and over-investment from one side and firm performance from the other side. We expected to find the following relationship:

1. At a low level of gearing (< 26.67% for UK and < 20.5% for US), we found that the higher the gearing the greater the probability of over-investment. On the other hand, in the second hypothesis, over-investment has a negative impact on firm performance. Therefore, it is expected that at a low level of gearing the interaction between gearing and over-investment will have a negative impact on firm’s performance.

2. At the higher level of gearing (> 26.67% for UK and > 20.5% for US), we found that the higher the gearing the smaller the probability of having over-investment; a higher level of gearing mitigates the over-investment problem. Therefore, we expected that at a high level of gearing, the interaction between gearing and over-investment might be significant and positive.
The two points above led us to test the non-linearity between the interaction of gearing, over-investment on one side and the performance on the other side. We also used three control variables: the natural logarithm of total sales (LnTS), the dividend payout ratio (DPU) and volatility. The results revealed that, at a low level of gearing, the coefficient of the interaction variables between gearing and over-investment was statistically significant at 5% with a negative sign. This implies that a low level of gearing together with over-investment have a negative impact on firm’s performance. This was expected because we found that at a low level of gearing, over-investment increases with gearing; debt does not mitigate the problem of over-investment but makes it worse and the over-investment should therefore have a negative impact on firm’s performance. On the other hand, the positive and significant coefficient (significant at the level of 8%) of the quadratic of the interaction variable between gearing and over-investment suggests that a middle level of gearing does mitigate the over-investment problem and, it is therefore associated with a positive effect on firm performance. Moreover, the non-significant coefficient of \((DG)^3\) suggests that a very high level of gearing has no impact on over-investment and therefore, has no further impact on firm performance and therefore, very high gearing does not monitor the over-investment problem. However, the associated negative sign of the of \((DG)^3\) coefficient on firm performance, although it is not significant, could be because financing over-investment, with very high gearing, negatively affects performance. Summery of the main hypotheses and the results is presented in the next page:
Table 7.1: Summery of the Tested Hypotheses

<table>
<thead>
<tr>
<th>The Hypotheses tested in this thesis</th>
<th>Approved</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: The defined Residual Income model can, to a significant degree, capture cross-sectional variation in market stock prices of the sample over a significant period of time.</td>
<td>Yes</td>
</tr>
<tr>
<td>H2: The first difference of the defined Residual Income model can, to a significant degree, capture the first difference of the cross sectional variation in market stock prices of the sample over a significant period of time.</td>
<td>Yes</td>
</tr>
<tr>
<td>H3: The defined Residual Income Components can, to a significant degree, capture cross-sectional differences in the performance of the market stock prices of the sample over a significant period of time.</td>
<td>Yes</td>
</tr>
<tr>
<td>H4: There is a negative relationship between gearing and over-investment.</td>
<td>Yes</td>
</tr>
<tr>
<td>H5: There is a negative relationship between over-investment and firm performance.</td>
<td>Yes</td>
</tr>
<tr>
<td>H6: The interaction between the presence of over-investment and level of gearing is related to better firm performance.</td>
<td>Yes</td>
</tr>
</tbody>
</table>

7.3 Contributions and Managerial Implications of the Study

After summarising the results of this thesis, we can ask whether we have advanced the knowledge in the area of capital structure and firm performance. There are several important features of our analysis that we believe, extend the literature on the empirical work of agency problem; more precisely, Jensen’s free cash flow argument.
Firstly, whereas most studies investigate US firms our sample comprises both UK and US firms. Secondly, we used a performance measure (RIC) with a solid theoretical underpinning that is also empirically tested. Thirdly, in Chapter 5, we highlighted the possibilities of using the RIC as a performance measure, since it outperforms the simple residual income model (higher $R^2$). The RIC represents the residual income component of the value in the RI model, and accounts for cost of capital. Fourthly, our analysis, distinct from previous empirical studies, introduced a simple way of measuring the over-investment problem i.e. a combination of capital expenditure and growth opportunities (Chapter 4). Fifthly, our analyses are based on panel data. The panel character of the data permits the use of statistical techniques that reduce the model specification bias or omitted variable bias. Sixthly, in Chapter 6 we have shown that the relationship between gearing, performance and over-investment is non-linear. Finally, our analysis in Chapter 6 incorporates the dynamic nature of the response of firm performance to investment decisions and controls for the endogeneity by lagging the explanatory variables one period.

The main finding of this study suggests that only a middle level of gearing mitigates the over-investment problem, since a low level of gearing increases the probability of over-investment and hence affects a firm's performance negatively. On the other hand in line with the first finding, which suggests that a high level of gearing has no impact on the over-investment problem, the interaction of gearing and over-investment has no impact on firm performance but the sign was negative. This could be explained by the fact that gearing reduces over-investment and extreme value might negate the performance. An important implication of our model is that certain levels of gearing, i.e. a middle level, play a monitoring role in reducing the over-investment problem and hence leads to better firm performance. We have found that debt at a low level increases the over-investment rather than decreases it. This could be because a low level of gearing does not put enough pressure on managers in terms of interest rate and in fact increases a firm’s financial slack. Therefore, middle level of gearing plays an important role in disciplining managers, reduces the over-investment problem and hence increases firm’s performance.
7.4 Limitations of this Study

The primary limitation of this study is the lack of complete data availability, primarily due to the incomplete data sets provided by financial databases. The available databases did not provide information for certain variables and had missing data for some of the variables that were present. Managerial ownership is a very important variable in the content of agency problems and firm performance, and incorporating such a variable could help us in explaining more about the motivation of over-investment. However, the managerial ownership variable was unavailable in the financial Datastream, and we had to omit this variable from our analysis. Moreover, the levels of significance for certain results were not very high (9% rather than 5%).

7.5 Suggestions for Further Research

The investigation in this thesis is not complete and further research remains to be done in the field of capital structure and firm performance; more precisely, in the relationship between investment, gearing and firm performance. For example, one could extend this study by incorporating the under-investment together with the over-investment hypotheses as well, to test the possibilities of obtaining the optimal investment. A second topic that arises from this research is to investigate the relationship between gearing, firm performance and diversification (such as acquisitions). Empire-building is arguably more likely to be expressed in acquisitions rather than purchases of fixed assets (we have seen from Chapter 6 that the levels of significance were low). A third topic for further investigation is the composition of both debt and equity to identify the gearing level. In this thesis, debt and equity have each been treated as homogeneous. However, several forms of debt and equity exist and each form has specific characteristics that are relevant to capital structure theories. In this area, Barclay and Smith (1995) study the determinants of debt maturity, while Houston and James (1996) and Johnson (1997) investigate the choice between public debt, bank debt and private debt. However, these studies are single equation models in which the choice of the structure of debt and its determinants are
assumed to be independent of other capital structure decisions. A study that includes both the choice for the maturity and sources of debt, and the debt-equity choice may yield interesting results with respect to capital structure choice.
Bibliography


Appendices

Appendix 3.A: Derivation of the Continuing Value Formula

This appendix proves the equivalence of the two recommended continuing formulas by Copeland et al (1994): the free cash flow perpetuity formula and the value-driver formula. The Two formulas are as follows:

\[
CV = \frac{FCF}{WACC - g} = \frac{NOPLAT \left(1 - \frac{g}{r}\right)}{WACC - g}
\]

Since the denominators are identical, we only need to prove that free cash flow can be expressed by the following equation:

\[
FCF = NOPLAT \left(1 - \frac{g}{r}\right)
\]

Where:
FCF = free cash flow.
NOPLAT = net operating profits less adjusted taxes.
g = growth rate in NOPLAT.
r = rate of return on net new capital invested.

First, they define free cash flow as the company’s operating profits less the net new capital invested.

\[
FCF = NOPLAT - In,
\]

Where In is the net increase in invested capital over and above replacement capital.

As long as the return on existing capital employed remains constant, they state that, a company’s NOPLAT in any period equals last period’s NOPLAT, plus the return in earns on last period’s net investment in new capital.
\[ \text{NOPLAT}_T = \text{NOPLAT}_{T-1} + (r \times \ln_{T-1}) \]

This Equation can be rearranged to show that the change in NOPLAT equals the rate of return on new investment times the amount of new investment.

\[ \text{NOPLAT}_T - \text{NOPLAT}_{T-1} = (r \times \ln_{T-1}) \]

Dividing both sides by last year's NOPLAT calculates the growth rate in NOPLAT:

\[ g = \frac{\text{NOPLAT}_T - \text{NOPLAT}_{T-1}}{\text{NOPLAT}_{T-1}} = r \times \frac{\ln_{T-1}}{\text{NOPLAT}_{T-1}} \]

\[ g = r \times \frac{\ln_{T-1}}{\text{NOPLAT}_{T-1}} \]

Solving for the amount of investment required increasing NOPLAT at the rate g, and substituting for the first definition of free cash flow gives the free cash flow calculation in terms of the key value drivers.

\[ \ln = \text{NOPLAT} \times \frac{g}{r} \]

\[ \text{FCF} = \text{NOPLAT} - \left( \text{NOPLAT} \times \frac{g}{r} \right) \]

\[ \text{FCF} = \text{NOPLAT} \left( 1 - \frac{g}{r} \right) \]

The ratio \( g/r \) can be called the net investment rate, as it represents the ratio of new investment to NOPLAT, which is consistent with a growth rate of \( g \) and a rate of return \( r \).
### Appendix 3.B: Balance Sheet before and after Adjustments

Table 3.1: Accounting-Based Balance Sheets

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities and Net Worth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current assets</td>
<td>Non-interest-bearing current liabilities (NIBCLs)</td>
</tr>
<tr>
<td>Net goodwill</td>
<td>Interest-bearing current liabilities (IBCLs)</td>
</tr>
<tr>
<td>Fixed assets</td>
<td>Long-term debt</td>
</tr>
<tr>
<td></td>
<td>Equity (net of write-off)</td>
</tr>
<tr>
<td>Total assets</td>
<td>Total Liabilities and net worth</td>
</tr>
</tbody>
</table>

Adapted from Bacidore *et al* (1997)

Table 3.2: Economic Book Value Balance Sheet

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities and Net Worth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current assets (with inventory at FIFO) - NIBCLs</td>
<td>IBCLs</td>
</tr>
<tr>
<td>Gross goodwill</td>
<td>Debt (+capitalised leases)</td>
</tr>
<tr>
<td>Fixed Assets</td>
<td>Equity (+adjustment for</td>
</tr>
<tr>
<td></td>
<td>Deferred taxes, goodwill</td>
</tr>
<tr>
<td></td>
<td>Amortisation, write-offs,</td>
</tr>
<tr>
<td></td>
<td>LIFO reserves, etc.)</td>
</tr>
<tr>
<td>Economic book value of assets in place</td>
<td>Total liabilities and net worth</td>
</tr>
</tbody>
</table>

Adapted from Bacidore *et al* (1997)
Appendix 3.C: Development of REVA

Bacidore et al (1997) state that the financiers must earn at least their opportunity cost of capital on the invested capital. This condition implies that this cost of capital must be subtracted from operating profits to gauge the firm’s financial performance. EVA, for that reason, defines net operating profit after tax (NOPAT) and subtracts a capital charge for the economic book value of assets in place. The economic book value of assets in place is the measure of the capital provided to the firm by its financiers, but does this amount truly represent the capital used to generate this operating profit? They believe the answer to that question is negative. For the firm to create a true “operating” surplus for its financiers in a given period, its operating profit at the end of the period must exceed a capital charge that is based on the total market value of the capital used at the beginning of the period, not simply the economic book value of its assets in place. The capital commitment of the firm’s financiers is represented by the total market value of the firm, not simply the economic book value of the assets in place.

For that reason, they modified the EVA measurement and defined a new performance measure, a Refinement of EVA (REVA). The motivation for the REVA refinement to EVA stems partly from EVA’s use of the economic book value of assets while the capital charge for the firm is derived from a market-based WACC. The REVA for a given time $t$ is defined as:

$$ REVA_t = NOPAT_t - k_e (MVA_{t-1}) $$

Bacidore et al (1997)

Where $NOPAT_t$ is the firm’s NOPAT at the end of period $t$ and $MVA_{t-1}$ is the total market value of the firm’s assets at the end of period $t-1$(beginning of period $t$). $k_e$ is the weighted-average cost of capital and $MVA_{t-1}$ is given by the market value of the firm’s equity plus the book value of the firm’s total debt less non-interest-bearing current liabilities, all at the end of period $t-1$. 
Bacidore et al (1997) pointed out that the difference between the market value of the firm and the economic book value of its assets in place represents the market’s assessment of the value of the firm’s current and future investment opportunities. They added that for the firm to create a true “operating” surplus for its financiers in a given period, its operating profit at the end of the period must exceed a capital charge. This capital charge is based on the total market value of the capital used at the beginning of the period, not simply the economic book value of its assets in place. The capital commitment of the firm’s financiers is represented by the total market value of the firm, not simply the economic book value of the assets in place. The key distinction between EVA and REVA is that REVA assesses its capital charge for period \( t \) on the market value of the firm at the end of period \( t-1 \) (or the beginning of period \( t \)) rather than on the economic book value of the assets in place.

Their empirical results indicate that the proportion of positive REVA that corresponds to positive abnormal returns is significantly higher than the same proportion for EVA. Thus, although EVA on its own predicts abnormal returns fairly well, REVA performs significantly better.

However, Ferguson and Leistikow (1998) show that REVA is inconsistent with finance theory and with wealth maximization. Consequently, it is inappropriate for measuring operating performance and rewarding management. They also show that EVA is consistent with finance theory and wealth maximization and is appropriate for measuring operating performance and rewarding management. Based on the argument that Bacidore et al (1997) made which stated that an appropriate measure of operating performance must correlate highly with abnormal stock returns, Ferguson and Leistikow (1998) argue that this view seems plausible because management decisions that change shareholder wealth also cause corresponding abnormal returns. Financial theory suggests, however, that no appropriate single-period measure of operating performance should be highly correlated with abnormal stock return, for the following reasons:

As stated by Ferguson and Leistikow (1998), an unanticipated management operating decision produces a corresponding abnormal stock return at the time it becomes known. Multiplying this return by the beginning stock price gives the decision’s
impact on shareholder wealth. The value of operating decisions, however, typically comes from their impact on operating performance in subsequent periods. Thus, an operating decision can add substantial value in the period in which it is made even if it reduces that period’s operating performance and there is no reason to expect an appropriate measure of one period’s operating performance to be significantly correlated with the same period’s abnormal stock returns. Therefore, they concluded that Bacidore et al’s belief that a good financial performance measure should correlate highly with abnormal stock returns is wrong. They also added that Bacidore et al provided no confidence that their REVA-based independent variable is useful for explaining abnormal stock returns.

Ferguson and Leistikow (1998) thus, disagreed with the Bacidore et al opinion that a true operating surplus for investors in a given period requires an operating profit at the end of the period that exceeds a capital charge based on the firm’s market value. In fact, Ferguson and Leistikow (1998) argue that the operating link to investors’ abnormal stock returns reflects a capital charge based on net asset value, not stock price. They state that the capital used to generate the firm’s operating profit is not the market value of its stock, as Bacodore et al believe. Certainly, they stated, it is the investors’ capital, and a firm’s capital is not the same thing as the investors’ capital. Thus, they argue that the invested capital which generated operating profit is net asset value, not the market value of the firm’s stock.

3.1.D: Derivation of the formula: \( d_t = x_t^a - y_t + R_f y_{t-1} \)

\[
x_t^a = x_t - (R_f - 1)y_{t-1}
\]

This means that \( x_t^a = x_t - R_f y_{t-1} + y_{t-1} \) meaning that: \( R_f y_{t-1} = x_t - x_t^a + y_{t-1} \) ...(1)

Combined with formula (A2a): \( y_t = y_{t-1} - d_t + x_t \) ...(2)

(1) - (2) implies: \( d_t = x_t^a - y_t + R_f y_{t-1} \)

3.2.D: Derivation of the formula: \( P_t = y_t + \sum_{r=1}^{\infty} R_f^{-r} E_t \left[ \bar{x}_{t+r} \right] \)

\[
d_{t+1} = x_{t+1}^a - y_{t+1} + R_f y_t
\]

(From the previous derivation)

Substitute the above equation with the PVED formula implies that:

\[
P_t = \sum_{r=1}^{\infty} R_f^{-r} E_t \left[ x_{t+1}^a - y_{t+1} + R_f y_t \right] = \sum_{r=1}^{\infty} E_t \left[ \frac{y_{t+1}}{R_f} \right] - \sum_{r=1}^{\infty} E_t \left[ R_f^{-r} R_f y_t \right] \]

Since \( \sum_{r=1}^{\infty} E_t \left[ \frac{y_{t+1}}{R_f} \right] \rightarrow 0 \) as \( r \rightarrow \infty \) and \( \sum_{r=1}^{\infty} E_t \left[ R_f^{-r} R_f y_t \right] = 0 \), this implies that:

\[
P_t = y_t + \sum_{r=1}^{\infty} R_f^{-r} E_t \left[ \bar{x}_{t+r} \right]
\]
Appendix 5.A: Non-Parametric Test- Spearman’s Correlation

Table A.1: Speaman’s Correlation Matrix between V and P.

<table>
<thead>
<tr>
<th></th>
<th>UK Correlation Matrix</th>
<th>US Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>.809*</td>
<td>V</td>
</tr>
</tbody>
</table>

*Correlation is significant at 1% level (2-tailed).

The above table reveals that the correlation between the value represented in Equation (2) and the stock prices is almost 81% for UK sample and 72% for US sample. This is in line with Frankel and Lee (1998) who find that the model explains more than 70% of the cross-sectional variation in stock prices.
Appendix 5.B: The Transformed Variables

Both V and P in both UK and US firms do not fall in a nice normally distributed curve. Unfortunately, nonparametric techniques tend to be less “powerful”. An alternative is to transform the variables logarithmically. For this reason, the original variables in both US and UK were transformed into logarithms and the new results are presented in the following table:

**Table 5.6: UK and US Results, the Transformed Variables**

<table>
<thead>
<tr>
<th></th>
<th>UK Results</th>
<th>US Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.502643</td>
<td>0.489471</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.04632</td>
<td>0.017357</td>
</tr>
<tr>
<td>t-Statistic</td>
<td>10.9*</td>
<td>28.2*</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.809399</td>
<td>0.879204</td>
</tr>
<tr>
<td>Chi^2-Statistic</td>
<td>117.8</td>
<td>123.2</td>
</tr>
<tr>
<td>Prob (Chi^2-Statistic)</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

* Means significant at 1% level

Looking at the results in the above table, R-squared (goodness of fit) is 80.9% and 87.92% for UK and US respectively, so it is clear that the new regression based on variables logarithmically transformed provides higher $R^2$ but the level of significant did not change for both UK and US data sets. This is expected since natural logarithm usually smooth the relationship between variables and this usually leads to a higher $R^2$. 

The sample is analysed for both Price/Book value and Price/Residual Income valuation model. The results of these analyses can be found in the following table:

<table>
<thead>
<tr>
<th>Method</th>
<th>R-Squared</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK</td>
<td>US</td>
</tr>
<tr>
<td>Price/Book value (Bt)</td>
<td>56%</td>
<td>68%</td>
</tr>
<tr>
<td>Price/Residual Income Valuation (Vt)</td>
<td>77%</td>
<td>76%</td>
</tr>
</tbody>
</table>

The above results are in line with the findings of Frankel and Lee (1998) who finds that the Price/Book value has the ability to predict the value of a firm, but the Price/Residual Income model performs better.

<table>
<thead>
<tr>
<th>Method</th>
<th>R-Squared</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK</td>
<td>US</td>
</tr>
<tr>
<td>RIC from F&amp;L (1998)</td>
<td>22%</td>
<td>29%</td>
</tr>
<tr>
<td>F&amp;L (1998) first difference</td>
<td>22%</td>
<td>26%</td>
</tr>
<tr>
<td>The Simple Residual Income</td>
<td>11%</td>
<td>20%</td>
</tr>
</tbody>
</table>

From the above table it is clear that the RIC outperforms both F&L first difference and the simple residual income model. R^2 from the simple residual income model, on the other hand, is lower than R^2 in both RIC and first difference.

We estimated model (1) using logit model and we used debt-equity ratio instead of gearing the results were as follows:

Modelling D (over-investment) by Logit results from UK and US data:

<table>
<thead>
<tr>
<th>C</th>
<th>BD/E</th>
<th>C</th>
<th>MD/E</th>
<th>C</th>
<th>BD/E</th>
<th>C</th>
<th>MD/E</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14789</td>
<td>-0.278</td>
<td>0.1358</td>
<td>-0.3751</td>
<td>0.523</td>
<td>-0.3442</td>
<td>0.4817</td>
<td>-0.7132</td>
</tr>
<tr>
<td>0.07156</td>
<td>0.08449</td>
<td>0.06528</td>
<td>0.1159</td>
<td>0.06267</td>
<td>0.04965</td>
<td>0.05708</td>
<td>0.09742</td>
</tr>
<tr>
<td>2.07</td>
<td>-3.29</td>
<td>2.08</td>
<td>-3.24</td>
<td>8.35</td>
<td>-6.93</td>
<td>8.44</td>
<td>-7.32</td>
</tr>
<tr>
<td>0.039</td>
<td>0.001</td>
<td>0.038</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-1024.4862</td>
<td>-1066.75225</td>
<td>-1470.4028</td>
<td>-1466.64219</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1031.337</td>
<td>-1072.992</td>
<td>-1503.528</td>
<td>-1503.528</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.702</td>
<td>12.479</td>
<td>66.25</td>
<td>73.771</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.37968576</td>
<td>1.38081686</td>
<td>1.34835421</td>
<td>1.34491043</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C is the constant. BD/E is the book value of debt-equity ratio. MD/E is the market value of debt-equity ratio.

The above table reveals that for US and UK data the coefficients of both book value and market value of debt-equity ratio are statistically significant and both have the expected negative sign. The finding suggests that the higher the gearing, the lower the probability of over-investment. This result, leads us to reject the null hypothesis in the sample used in this study (over 12 years) and to accept the alternative which states that there is a negative relationship between gearing and over-investment.

We estimated model (1) using logit model and we used the ratio of total debt plus preference share capital to total assets plus the difference between market value of common stock and book value of common equity and the results were as follows:

Modelling D (over-investment) by Logit results from UK and US data:

<table>
<thead>
<tr>
<th></th>
<th>UK Data</th>
<th></th>
<th>US Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>MG</td>
<td>C</td>
<td>MG</td>
</tr>
<tr>
<td>Coefficient</td>
<td>0.01778</td>
<td>-0.1154</td>
<td>0.65216</td>
<td>-0.2637</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.08530</td>
<td>0.4446</td>
<td>0.07284</td>
<td>0.3357</td>
</tr>
<tr>
<td>t-value</td>
<td>0.208</td>
<td>-0.260</td>
<td>8.95</td>
<td>-7.86</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.835</td>
<td>0.795</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1072.95814</td>
<td></td>
<td>-1471.62859</td>
<td></td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-1072.992</td>
<td></td>
<td>-1503.528</td>
<td></td>
</tr>
<tr>
<td>Test Chi-Squared</td>
<td>0.06738</td>
<td></td>
<td>63.799</td>
<td></td>
</tr>
<tr>
<td>Chi-Squared prob.</td>
<td>0.7952</td>
<td></td>
<td>0.00000</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.3888348</td>
<td></td>
<td>1.34947673</td>
<td></td>
</tr>
</tbody>
</table>

C is the constant. G is the book value industry-adjusted gearing.

US data reveal similar results to Table 6.5 where the coefficient is negatively significant at the 1% level of significant and hence, lead to reject the null hypothesis and to accept the alternative, which states that there is a negative relation between gearing and over-investment. However, despite the fact that the coefficient from US data is negative however, it is not significant. Based on the above results, this measurement of market value of gearing will be ignored.
Appendix 6.C: Over-investment and Debt-Equity Ratio Classification
Assuming Linearity.

Modelling D by Logit results from US D/E ratio classification:

<table>
<thead>
<tr>
<th></th>
<th>Top Gearing</th>
<th>Middle Gearing</th>
<th>Bottom Gearing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Constant</td>
<td>Constant</td>
</tr>
<tr>
<td>Coefficient</td>
<td>0.2689</td>
<td>0.3562</td>
<td>0.2408</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.167</td>
<td>0.1374</td>
<td>0.1335</td>
</tr>
<tr>
<td>t-value</td>
<td>1.61</td>
<td>2.59</td>
<td>1.80</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.108</td>
<td>0.010</td>
<td>0.072</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-362.052506</td>
<td>-733.51569</td>
<td>-365.660653</td>
</tr>
<tr>
<td>Baseline Log-likelihood</td>
<td>-374.5745</td>
<td>-733.5287</td>
<td>-368.1093</td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>25.044</td>
<td>0.0260</td>
<td>4.8972</td>
</tr>
<tr>
<td>Chi^2 prob.</td>
<td>0.0000</td>
<td>0.8717</td>
<td>0.0269</td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.31903082</td>
<td>1.36206609</td>
<td>1.33210382</td>
</tr>
</tbody>
</table>

Modelling D by Logit results from UK D/E ratio classification:

<table>
<thead>
<tr>
<th></th>
<th>Top Gearing</th>
<th>Middle Gearing</th>
<th>Bottom Gearing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Constant</td>
<td>Constant</td>
</tr>
<tr>
<td>Coefficient</td>
<td>-0.05779</td>
<td>0.544237</td>
<td>-0.421946</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.1386</td>
<td>0.1381</td>
<td>0.1496</td>
</tr>
<tr>
<td>t-value</td>
<td>-0.417</td>
<td>3.94</td>
<td>-2.82</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.677</td>
<td>0.0000</td>
<td>0.005</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-261.798138</td>
<td>-507.09587</td>
<td>-260.96829</td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-263.867</td>
<td>-511.1735</td>
<td>-261.4618</td>
</tr>
<tr>
<td>Test Chi^2</td>
<td>4.1378</td>
<td>8.1552</td>
<td>0.98693</td>
</tr>
<tr>
<td>Chi^2 prob.</td>
<td>0.0419</td>
<td>0.0043</td>
<td>0.3205</td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.37394863</td>
<td>1.36853729</td>
<td>1.36962651</td>
</tr>
</tbody>
</table>
The above three percentile in both US and UK data were sorted in the same way that gearing was sorted. It should be noted that both book value and market value of debt to equity ratio revealed similar results. The US data for debt-equity ratio revealed different results from gearing ratio where the middle percentile was negatively significant when we used gearing however it is in-significant when debt-equity ratio was used. Top percentile in the US data revealed no differences between gearing and debt-equity ratio. The bottom percentile for debt-equity ratio is positive and significant as in gearing ratio. Moreover, The UK data for debt-equity ratio revealed very similar results to gearing ratio.

Modelling D by Logit: the UK and US results using debt-equity ratio

<table>
<thead>
<tr>
<th></th>
<th>UK results</th>
<th></th>
<th>US results</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cons.</td>
<td>BD/E</td>
<td>BD/E$^2$</td>
<td>BD/E$^3$</td>
</tr>
<tr>
<td>Coeff.</td>
<td>0.01216</td>
<td>-0.0211</td>
<td>0.00026</td>
<td>2.2e$^{-05}$</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.0567</td>
<td>0.0408</td>
<td>0.0012</td>
<td>3.67e$^{-05}$</td>
</tr>
<tr>
<td>t-value</td>
<td>0.215</td>
<td>-0.518</td>
<td>0.216</td>
<td>0.0617</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.830</td>
<td>0.605</td>
<td>0.829</td>
<td>0.951</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1072.58508</td>
<td></td>
<td>-1461.51231</td>
<td></td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-1072.992</td>
<td></td>
<td>-1503.528</td>
<td></td>
</tr>
<tr>
<td>Test Chi$^2$</td>
<td>0.81351</td>
<td></td>
<td>84.031</td>
<td></td>
</tr>
<tr>
<td>Prob.</td>
<td>0.8462</td>
<td></td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.39093679</td>
<td></td>
<td>1.34204424</td>
<td></td>
</tr>
</tbody>
</table>

Market value of debt-equity ratio revealed very similar results to book value of debt-equity ratio. The above table reveals that for UK data, the results are insignificant. It is significant only when we drop both D/E$^2$ and D/E$^3$. This could indicate that the relationship between over-investment and debt-equity ratio, unlike gearing, cannot be expressed by non-linearity. On the other hand, for US data, the D/E is only significant while both D/E$^2$ and D/E$^3$ are insignificant. This means as well, for US data, the relationship between over-investment and debt-equity ratio, unlike gearing, cannot be expressed by non-linearity.
Appendix 6.E Over-investment and Gearing Assuming Non-Linear Variables

Modelling D by Logit: the UK and US results

<table>
<thead>
<tr>
<th></th>
<th>UK results</th>
<th>US results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cons.</td>
<td>MG</td>
</tr>
<tr>
<td>Coeff.</td>
<td>-0.2968</td>
<td>0.05329</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.1370</td>
<td>0.0183</td>
</tr>
<tr>
<td>t-value</td>
<td>-2.17</td>
<td>2.19</td>
</tr>
<tr>
<td>t-prob.</td>
<td>0.030</td>
<td>0.004</td>
</tr>
<tr>
<td>Loglikelihood</td>
<td>-1061.7925</td>
<td></td>
</tr>
<tr>
<td>Baseline loglikelihood</td>
<td>-1072.992</td>
<td></td>
</tr>
<tr>
<td>Test Chi²</td>
<td>22.399</td>
<td></td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>AIC/n</td>
<td>1.37699289</td>
<td></td>
</tr>
</tbody>
</table>

The above table reveals very similar results to Table 6.10.
# Table of Contents

Abstract ........................................................................................................................... I

List of Tables .................................................................................................................... II

List of Figures ................................................................................................................... IV

List of Abbreviations ....................................................................................................... V

Table of Contents ........................................................................................................... vii

Acknowledgments .......................................................................................................... X

Chapter 1 ........................................................................................................................ I

Introduction ..................................................................................................................... 1

1.1 Introduction ............................................................................................................ 1

1.2 The Aim of this Study .......................................................................................... 3

1.3 Outline of the thesis .............................................................................................. 5

Chapter 2 ........................................................................................................................ 7

Capital Structure ............................................................................................................. 7

2.1 Introduction ............................................................................................................ 7

2.2 Modigliani and Miller Theory ............................................................................. 9

2.2.1 Introduction .................................................................................................... 9

2.2.2 MM Proposition I ....................................................................................... 9

2.2.3 MM Proposition II ................................................................................. 12

2.2.4 MM with Corporate Taxes .................................................................... 14

2.2.5 Implications of the MM Theory ............................................................ 16

2.3 The Miller Model ............................................................................................. 17

2.4 Cost of Financial Distress ................................................................................ 18

2.5 The Trade-Off Theory ..................................................................................... 20

2.6 Agency Costs ................................................................................................... 23

2.6.1 Introduction ............................................................................................. 23

2.6.2 Agency Cost of External Equity ............................................................ 24

2.6.3 Agency Cost of Debt ............................................................................... 25

2.7 Signalling and Asymmetric Information Theory ............................................. 29

2.8 The Pecking Order Theory ............................................................................ 31

2.9 Free Cash Flow Theory: Jensen (1986) ......................................................... 34

2.10 Empirical Evidence on Capital Structure ................................................. 35

2.11 Conclusion ....................................................................................................... 41

Chapter 3 ...................................................................................................................... 46

Financial Performance Measures ................................................................................. 46

3.1 Introduction ........................................................................................................ 46

3.2 Financial Ratios ............................................................................................... 47

3.3 Discounted Cash Flow Model (DCF): The Free Cash Flow Model .......... 48

3.4 Discounted Cash Flow: The Economic Profit Model .................................. 53