The Use of Balanced Scorecard Measures in Executive Incentives and Organizational Performance

By

Kangwu Kenneth Xi, CFA, MBA

A thesis submitted to

The Faculty of Graduate and Postdoctoral Affairs

In particular fulfillment of the requirements for the degree

Doctor of Philosophy

Eric Sprott School of Business

Carleton University

Ottawa, Canada

September 20, 2010

©Copyright 2010, Kangwu Kenneth Xi

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-70565-0
Our file Notre référence
ISBN: 978-0-494-70565-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The Balanced Scorecard (BSC) has been argued as a superior combination of non-financial and financial measures of performance. The power of this concept can be used only if BSC measures are linked to reward systems. Despite an increasing usage of BSC, however, virtually no large-scale empirical studies have documented the prevalence of BSC in executive compensation plans. This thesis attempted to address this deficiency. Three primary questions are central to the study: Do contextual factors affect the usage of BSC in compensation plans? Does BSC have an impact on organizational performance? Is this performance associated with the fit between BSC and firm characteristics?

The hypotheses were tested using a three-stage regression model including logistic and OLS regressions, based on a sample of 165 firms that adopted BSC measurement schemes and 165 control firms over the period of 1992–2009.

The results underscore the importance of considering firm characteristics prior to adopting BSC. This study finds that 1) the usage of BSC in executive compensation plans is related to strategy, certain industries, size, quality, structure, Anglo culture and ownership; 2) there is a strong association between firm performance and the adoption of BSC; and 3) firm performance is an increasing function of the level of fit of BSC with its contingent factors.

In addition, this study confirms that firm characteristics explain the relative weights placed on non-financial components of BSC. Yet, the evidence regarding the relationship between firm performance and the relative weights of non-financial measures and the fit is mixed.

This research enriches the existing knowledge of performance measurement by making contributions to a limited literature. As the first empirical study that systematically examines BSC compensation plan in the context of contingency theory, this thesis provides concrete evidence on the validity of this theory. It also reminds shareholders and managers that firms should seek to improve the alignment between BSC incentive package and firm characteristics to ensure better performance. Over-relying or under-relying the relative weights placed on financial or non-financial measures may not help firm performance.

Acknowledgement

I owe my most sincere gratitude to my thesis advisor, Dr. Raili Pollanen for her continuous support in the Ph.D. Program. Raili has shown me how to approach a research problem and write an academic paper. She was always there to meet with me and give detailed and constructive comments.

I am deeply grateful to Dr. Claude Laurin and Dr. Gregory Richards for their precious time and expertise dedicated to my thesis defence.

I would like to give heartfelt, special thanks to Dr. David Cray, Dr. Howard Nemiroff and Dr. Hilary Becker who guided me on some of the theoretical and econometric issues from the proposal to the final thesis.

Last but not least, I am indebted to my wife Yanling Wang who has provided valuable suggestions on econometrics.

Thank you all!

Dedication

This dissertation is dedicated to my family, wife Yanling Wang, son Richard Dingyuan Xi, and daughter Suri Zifei Xi, without whose unconditional love and longstanding support I would have never accomplished this goal.

Table of Contents

Abstract.....	ii
Acknowledgement.....	iv
Dedication.....	v
Table of Contents.....	vi
List of Tables.....	xi
List of Figures.....	xiii
List of Acronyms.....	xiv
List of Appendices.....	xv
I. Introduction.....	1
1. Importance of the Study.....	1
2. Research Objectives.....	4
3. Major Findings.....	5
II. Literature Review	8
1. Balanced Scorecard.....	8
1.1 BSC Overview.....	10
1.2 BSC Criticisms.....	12
1.3 What Constitute a Successful BSC?	15
2. Executive Incentives.....	24
2.1 Managerial Compensation.....	24
2.2 Do Incentives Work?.....	28
2.3 Performance Measurement.....	29

3. Theoretical Foundations.....	31
3.1 Contingency Theory.....	32
3.2 Cons and Pros.....	35
3.3 The Concept of “Fit”	36
4. Summary.....	38
III. Conceptual Model and Hypotheses Development.....	40
1. Framework and Model.....	40
2. Hypotheses–BSC Use in Exec. Compensation and Contingent Factors....	43
2.1 Strategy.....	44
2.2 Industry.....	47
2.3 Size.....	49
2.4 Quality	51
2.5 Structure	52
2.6 Culture.....	53
2.7 Ownership.....	58
3. Hypotheses–BSC Use in Executive Comp. and Corporate Performance...	59
4. Hypotheses–The Fit of BSC and Corporate Performance.....	65
5. Summary.....	67
IV. Research Design and Methods.....	68
1. Research Design.....	68
1.1 Qualitative vs. Quantitative.....	68
1.2 Deduction vs. Induction.....	70
1.3 Reliability and Validity.....	71

1.4 Levels of Measurement.....	74
2. Data Collection.....	76
2.1 Sampling Design.....	77
2.2 Sample Characteristics	80
2.3 Sub-sample of Firms with Non-financial Weights.....	82
3. Measurement of Variables.....	84
3.1 BSC Usage in Executive Compensation Plans.....	84
3.2 Contingent Variables.....	86
3.3 Performance Variables.....	91
3.4 Control Variables.....	95
3.5 Measurement of Fit.....	99
4. Data Preparation, Verification and Testing.....	99
4.1 Database Construction.....	99
4.2 Missing Data.....	101
4.3 Normality Assessment.....	103
4.4 Parametric vs. Nonparametric Statistics.....	106
5. Summary.....	107
V. Results: Full Sample of Firms Adopting BSC	108
1. Model 1—BSC Usage and Firm Characteristics	109
1.1 Descriptive Statistics and Univariate Tests.....	109
1.2 Composite Measures of Contingent Variables.....	117
1.3 Logistics Regression for BSC Usage.....	119
2. Model 2—Firm Performance and BSC Usage	124

2.1 Descriptive Statistics and Univariate Tests.....	125
2.2 Composite Measures of Firm Performance.....	131
2.3 OLS Regressions for Performance.....	133
2.4 Additional Tests on Individual Measures.....	136
3. Model 3—Contingent Fit and Firm Performance.....	139
3.1 Descriptive Statistics.....	139
3.2 OLS Regressions for Fit.....	139
3.3 Additional Tests on Individual Measures.....	142
4. Summary.....	143
VI. Results: Sub-sample of Firms with Non-financial Weights.....	145
1. Model 1—Non-financial Weights and Firm Characteristics.....	145
1.1 Descriptive Statistics and Univariate Tests.....	146
1.2 Composite Measures of Contingent Variables.....	149
1.3 OLS Regressions for Non-financial Weights.....	150
2. Model 2—Firm Performance and Non-financial Weights.....	153
2.1 Descriptive Statistics and Univariate Tests.....	154
2.2 Composite Measures of Firm Performance.....	157
2.3 OLS Regressions for Performance.....	158
2.4 Additional Tests on Individual Measures.....	160
3. Model 3—Contingent Fit and Firm Performance	162
3.1 OLS Regressions for Fit.....	162
3.2 Additional Tests on Individual Measures.....	164
4 Summary.....	165

VII. Discussions.....	166
1. Analysis of Findings.....	166
2. Major Contributions.....	170
3. Ex-post Model Development.....	173
4. Methodological Consideration.....	178
5. Summary.....	181
VIII. Conclusions and Future Directions.....	182
1. Summary.....	182
2. Limitations.....	183
3. Future Research.....	186
4. Concluding Remarks.....	189
IX. References.....	190

List of Tables

Table 2.1: Summary statistics of articles.....	9
Table 4.1: The year of BSC adoption	80
Table 4.2: Two-sample t-test (total assets).....	81
Table 4.3: Measurement of contingent variables.....	86
Table 4.4: Industry classification by SIC	87
Table 4.5: Global society/culture clusters	89
Table 4.6: Measurement of performance variables	92
Table 4.7: Measurement of control variables	96
Table 4.8: Product development and life cycles	98
Table 5.1: Descriptive statistics (contingent variables).....	110
Table 5.2: Two-sample difference tests (contingent variables)	112
Table 5.3: Correlations between dependent and independent variables.....	113
Table 5.4: Correlation matrix (contingent variables)	115
Table 5.5: Eigenvalues for composite contingent variables.....	118
Table 5.6: Test of H1 - BSC and firm characteristics.....	120
Table 5.7: Test of H1 - BSC and firm characteristics (individual measures).....	122
Table 5.8: Descriptive analysis (performance and control variables).....	125
Table 5.9: Two-sample difference test (performance & control variables)	126
Table 5.10: Correlations between dependent and independent variables.....	130
Table 5.11: Correlation matrix (independent and control variables)	131
Table 5.12 Eigenvalues for composite firm performance.....	132

Table 5.13: Test of H2 - BSC and firm performance.....	134
Table 5.14: Test of H2 - BSC and firm performance (with control variables).....	135
Table 5.15: Two-sample t-test (fit)	139
Table 5.16: Test of H3-BSC fit and firm performance.....	140
Table 5.17: Test of H3 – BSC fit and firm performance (with control variables).....	141
Table 6.1: Descriptive statistics (sub-sample contingent variables).....	146
Table 6.2: Correlations between dependent and independent variables.....	147
Table 6.3: Correlation matrix (sub-sample contingent variables)	148
Table 6.4: Eigenvalues for composite contingent variables (sub-sample).....	149
Table 6.5: BSC and firm characteristics (sub-sample).....	150
Table 6.6: Non-financial weight and firm characteristics (individual measures).....	151
Table 6.7: Descriptive analysis (sub-sample performance and control variables).....	154
Table 6.8: Correlations between dependent and independent variables.....	155
Table 6.9: Correlation matrix (sub-sample).....	156
Table 6.10: Eigenvalues for composite firm performance (sub-sample).....	157
Table 6.11: Non-financial weights and firm performance.....	158
Table 6.12: Non-financial weights and firm performance (with control variables)...	159
Table 6.13: Non-financial fit and firm performance.....	162
Table 6.14: Non-financial fit and firm performance (with control variables).....	163
Table 7.1: Measurement levels of firm performance.....	175

List of Figures

Figure 3.1: Framework for the study	41
Figure 4.1: Levels of measurement	76
Figure 4.2: The distribution of BSC adoption by year	81
Figure 4.3: The distribution of BSC adoption by year (sub-sample)	83
Figure 7.1: Ex-post model.....	174

List of Acronyms

AR	=	Accounts receivables
BSC	=	Balanced scorecard
CEO	=	Chief executive officer
CF	=	Cash flow
COGS	=	Cost of goods sold
D/E	=	Debt to equity
EBIT	=	Earnings before interest and tax
EPS	=	Earnings per share
EVA	=	Economic value added
GAAP	=	Generally accepted accounting principles
M&A	=	Merger and acquisitions
MV/BV	=	Market to book value
NI	=	Net income
OLS	=	Ordinary least squares
P/E	=	Price to earnings
PCA	=	Principle component analysis
R&D	=	Research and development
ROA	=	Return on assets
ROE	=	Return on equity
ROI	=	Return on investment
SEC	=	Securities and Exchange Commission
SEM	=	Structural equation modelling
SIC	=	Standard industry code
VIF	=	Variance inflation factor

List of Appendices

Appendix 1: Excerpts from proxy statements	199
Appendix 2: The list of sample firms.....	200
Appendix 3: K-density estimations: continuous data.....	206
Appendix 4: Histograms: categorical data.....	212
Appendix 5: Tests of significant control variables (full-sample).....	214
Appendix 6: Test of H2 – BSC and firm performance (no control variables).....	215
Appendix 7: Test of H2 – BSC and firm perform. (with control variables).....	216
Appendix 8: Tests of H3 – BSC fit and firm perform. (no control variables).....	218
Appendix 9: Tests of H3–BSC fit and firm performance (with control variables)....	219
Appendix 10: Tests of significant control variables (sub-sample).....	221
Appendix 11: Non-financial weights and firm performance (no control variables)...	222
Appendix 12: Non-financial weight and firm performance (with control variables).	223
Appendix 13: Non-financial fit and firm perform. (no control variables).....	225
Appendix 14: Non-financial fit and firm perform. (with control variables).....	226

I. Introduction

Recent and widespread use of the Balanced Scorecard (BSC) motivates this study. BSC, popularized by Kaplan and Norton (1992, 1993 and 1996) and adopted widely around the world, has been argued to be a superior combination of non-financial and financial measures of performance. Current non-financial measures have been advocated to be better measures of future corporate performance, as they help refocus managers on the long-term aspects of their actions. Interest in BSC continues to grow. A survey estimates that 60% of Fortune 1000 firms have experimented with BSC (Silk, 1998). A worldwide study of management tools shows that BSC achieves a utilization rate of 44% (Rigby, 2001). Atkinson et al. (1997, p. 94) noted that “the balanced scorecard is among the most significant developments in management accounting and thus, deserves intense research attention.” They further suggested that BSC has the potential to provide planners with a way of expressing and testing a sophisticated model of cause-and-effect relationships in the organization—a model that provides managers with a basis on which to manage the drivers of desired outcomes.

1. Importance of the Study

Despite the increasing usage of BSC measures in managerial compensation, little empirical evidence is available on the relationships between BSC and various potential contingent factors, and even less is known about performance impacts of incorporating BSC in executives' annual incentives contracts. This thesis aims to make unique

contributions by addressing the above deficiencies. This study is significant for the following five reasons:

First, even though the literature seems to support the impression that BSCs are widely used in many companies, only limited systematic research has been done on BSC applications. Speckbacher, Bischof and Pfeiffer (2003) pointed out that one obvious reason is that most previous studies suffer from methodological shortcomings, such as a bias with respect to the selected companies, low response rates or unreliable estimates. Accordingly, the findings are quite different and often incoherent.

Second, few empirical studies have put BSC in a theoretical context. This study tested the practice of BSC from the angle of contingency theory that argues that the design and use of performance measurement depends upon organizational and environmental contexts (Hansen, Mowen, Senkow and Pollanen, 2004). A better match between BSC and contextual variables, such as external environment, technology, structure, size, culture and strategy, is hypothesized to result in increased corporate performance (Fisher, 1998; Sila, 2007), while poor fit implies diminished performance.

Third, Banker, Potter and Srinivasan (2000) found that although there had been calls for greater emphasis on non-financial measures in internal performance measurement systems (Kaplan and Norton, 1996) and firms had implemented such systems, empirical evidence on BSC in performance evaluation was inadequate (Ittner and Larcker, 1998). In addition, incentive plans that include BSC can affect financial performance (Wagner,

Rubin and Callahan, 1988; Banker, Lee and Potter, 1996), yet empirical support for the hypothesized performance impacts of including BSC measures in compensation plans is at best weak (Ittner and Larcker, 1998). Despite survey evidence that a growing number of firms are using BSC for compensation purposes, relatively little is known about the implementation issues associated with scorecard-based reward systems (Ittner, Larcker and Meyer, 2003). Ittner, Larcker and Rajan (1997), for example, analyzed the determinants of the use of non-financial measures in CEO compensation, but did not examine the performance impacts of such compensation plans.

Fourth, there has been little work relating the adoption of BSC to corporate performance, in particular, market-based measures. There have been increasing calls for the investigation of stock market reaction to BSC adoption. McConnell and Muscarella (1985) reported positive associations between announcements of increases in capital investment plans, management control system and share price movement. Larcker (1983) found that firms adopting incentive performance plans experienced an increase in capital investment and a positive security market reaction to disclosure of the plan to the market. Yet, in contrast to Larcker's findings, Gaver and Gaver's (1993) replication provided little evidence that capital spending increases following adoption.

Finally, previous studies have postulated that the fit between performance measures and firm characteristics is positively associated with corporate performance (Chenhall and Langfield-Smith, 1998), but there are few empirical tests on how to measure fit and how this fit affects a firm's performance. HassabElnaby, Said and Wier (2005) recently related

the match of non-financial measures and firm characteristics to subsequent enhanced performance. This thesis extends their study by investigating the interfaces between the fit of BSC and its performance consequences.

This study will shed additional light on these key issues in these five areas.

2. Research Objectives

There is a great demand for research on the proper construction of BSC bonus plans and their effects. As virtually no large-scale studies have documented the prevalence of BSC in executive compensation contracts, this study aims to undertake the task by examining the following research questions: a) how contingent contextual factors (strategy, industry, size, structure, quality, culture and ownership) affect BSC usage in executive compensation plans, b) how BSC usage in executive compensation plans affects organizational performance (effectiveness and efficiency), and c) how the fit between BSC and contingent factors affects organizational performance. This empirical study has important implications for the design of executive incentives. It is widely held that the real power of the BSC concept can be used only if BSC measures are linked to reward systems.

The BSC consists of financial and non-financial components. Not all firms apply BSC in the same way since the weights placed on financial and non-financial measures vary across firms. Using non-financial weight as a proxy to the extent of BSC usage, therefore,

this study also investigates the factors influencing the relative weights placed on non-financial measures, the association of corporate performance with non-financial weights, and whether this performance is contingent on the fit between non-financial weights and firm characteristics.

This study, which is rooted in a variety of innovative data collection and analysis methods, will provide valuable evidence for both academic research and management.

3. Major Findings

The hypotheses were tested using a three-stage regression model including logistic and OLS regressions, based on a sample of 165 firms that have adopted BSC measurement schemes and 165 matched firms that have not over the period of 1992–2009. First, the results underscore the importance of considering firm characteristics prior to adopting BSC. This study reports that firms pursuing prospector strategy and quality management are more likely to use BSC. Firms in financial and utility industries and in the Anglo cultural cluster tend to adopt BSC. The study finds positive associations of BSC with the size and structure of firms. It also supports the hypothesis that more concentrated institutional ownership provides motivation for managers to adopt BSC. Second, although one cannot conclude from this research that a BSC-based measurement scheme will “cause” managers to make decisions that are valued by shareholders, a strong association is found between the adoption of BSC and corporate performance. Relative to the control firms, treatment firms report higher accounting returns, higher productivity,

more intense use of assets, and lower leverage. Also, strong evidence suggests that market participants may respond favourably to the adoption of BSC-based compensation plans, as evidenced in higher stock returns, P/E ratio, dividend yields and price/cash flow ratio. Ideally, the performance measures selected for management compensation contracts should be those measures that best motivate management to maximize the value of firms. Finally, consistent with expectations, the results indicate that firm performance is an increasing function of the level of fit between BSC and its contingent factors.

At the same time, this study confirms that firm characteristics explain the relative weights placed on non-financial measures in compensation contracts. The prospector firms, larger firms, quality-oriented firms, firms with more diversified structure, and firms with more concentrated institutional ownership rely more heavily on non-financial components of BSC. Firms engaged in financial and utility industries and in the Anglo cultural cluster put more emphasis on non-financial measures. However, the results regarding the relative weights of non-financial measures and firm performance is mixed.

As one of the first empirical studies that examines BSC compensation plan in the context of contingency theory, the results provide concrete empirical evidence on the validity of this theory. Also as the first known empirical study that examines the association of BSC and its relative non-financial weights with firm performance, this study enriches existing knowledge of performance measurement by making contributions to a limited literature. This study reminds shareholders and managers that firms should improve the alignment between their interests by implementing a better BSC incentive package. Yet, the

cost/benefit aspects and financial/non-financial weights in designing each component of BSC system should be considered at an early stage. Unlike prior research, this study uses multiple indicators to represent each of the contingent, performance and control variables, which gives greater explanatory power. Some variables, such as structure and culture, are introduced for the first time. Both accounting- and market-based metrics are used to measure organizational performance.

This thesis is structured as follows. The subsequent chapter provides a critical review of extant literature regarding BSC, executive compensation, theoretical foundations and related issues. Chapter III illustrates the research model and hypothesis development. Chapter IV introduces research design, data collection, measurement of variables, and data preparation and verification. Chapter V discusses empirical results for the full sample of BSC firms and their matched control firms. Chapter VI reports empirical results for the sub-sample of firms that provide non-financial weights. Chapter VII discusses findings, contributions, ex-post model development and methodological issues. The last chapter concludes with a review of limitations and future research.

II. Literature Review

This chapter presents a critical analysis of extant literature on BSC, executive incentives, and theoretical foundations of the study. Given the breadth of research topics and methods, it is unrealistic for a single thesis to adequately cover the entire field. 87 papers, 29 books and 3 working papers are referred in this study. This study chooses references based on their relevance (performance measurement and BSC), journal ranking (over 70% of papers are from top accounting/finance/management journals—see Brown, Jones and Steele, 2007), publishing date (86% were published in/after 1990), and a balance between empirical and theoretical issues. Extensive tables, charts and figures are employed for the organization and analysis of these materials. Table 2.1 provides a snapshot of the papers and their sources.

1. Balanced Scorecard

This section offers a detailed review of BSC, which represents a notable development in the field of performance measurement. BSC is defined as a set of causally linked non-financial and financial objectives and performance measures designed to align managers' actions with an organization's strategy (Kaplan and Norton, 1996). Academic and practitioner interest in BSC continues to grow. There is no perfect BSC system in all contexts, but effective performance management systems share some features in common.

Table 2.1: Summary statistics of articles

Source of Studies	Number of Items
<i>Journals (total papers)</i>	87
Accounting, Organizations and Society	13
Journal of Accounting and Economics	9
Harvard Business Review	6
Journal of Financial Economics	3
Journal of Management Accounting Research	6
The Academy of Management Review	4
Contemporary Accounting Research	1
Management Accounting Research	9
The Journal of Finance	1
The Accounting Review	7
Journal of Political Economy	1
Management Accounting	1
International Journal of Operations and Production Management	3
Journal of Operations Management	1
Journal of Accounting Research	3
The International Journal of Accounting	1
Quarterly Journal of Economics	2
British Journal of Management	1
Behavioural Research in Accounting	2
International Journal of Public Sector Management	1
Economic Journal	1
Journal of Business and Economic Statistics	1
Accounting, Auditing and Accountability Journal	1
California Management Review	1
American Sociology Review	1
Journal of International Business Studies	1
Organizational Research Methods	1
Others	5
<i>Books</i>	29
<i>Working papers</i>	3

1.1 BSC Overview

The BSC system attempts to address the needs of the different stakeholders of an organization by creating a blend of strategic measures: outcome and driver measures, financial and non-financial measures, and internal and external measures (Anthony and Govindarajan, 2007). Outcome measurements relate to the results of strategy. They typically are “lagging indicators” that tell management what has happened. By contrast, driver measures are “leading indicators” that relate to processes of key areas in implementing a strategy. Otley (2005) suggested that the financial and customer perspectives can be viewed as providing results measures, whereas the internal business and innovation and learning perspectives serve as a means by which results are obtained. The most important aspect of this method is its ability to measure outcomes and drivers in a way that should cause the organization to act in accordance with its strategies. Kaplan and Norton (1996) stressed the importance of the link between an organization’s strategy and its performance measurement system. They argued that BSC helps management in four steps: 1) translating the vision, 2) communicating the strategy and linking it to departmental and individual objectives, 3) business planning, and 4) feedback and learning.

Companies should adopt a range of measures covering both financial and non-financial aspects of performance. Both internal and external benchmarking can be used to set standards against which performance is assessed. Effectiveness may be more fundamental than efficiency in that cost reductions should not be given greater

importance than the achievement of output objectives. Efficiency is “the ratio of outputs to inputs, or the amount of output per unit of input” and effectiveness “the relationship between a responsibility center’s outputs and its objectives” (Anthony and Govindarajan, 1998, pp. 130-131). As Hansen et al. (2004) pointed out, managing costs may increase an activity’s efficiency, but if the activity itself is unnecessary, it does not matter whether it is performed efficiently or not. An unnecessary activity is wasteful and should be eliminated. An appropriate package of incentive and reward arrangements should be central to the design of a performance management system. There should be a balance and tradeoff between rewarding results and behaviour (Otley, 2005).

Overall, BSC has several potential benefits. Firstly, compared with traditional measurement systems that include only financial measures, such as Economic Value Added (EVA) and stock return, BSC is designed to improve managers’ decision making by guiding their attention to a broader vision of the company’s operations (Kaplan and Norton, 1992). Secondly, as a holistic performance measurement system, BSC provides causal links connecting the multiple classes of non-financial measures (“drivers of the performance”) and the financial measures (“final outcome”) (Ittner, Larcker and Randall, 2003). As such, it clearly shows the links by which specific improvements in the drivers are expected to lead to desired outcomes according to the strategy. Thirdly, BSC can be used as a strategic management system (Kaplan and Norton, 1996, 2001). In particular, the pattern of results for the causally linked driver and outcome measures may provide clues about the effectiveness of the strategy. This then can provide companies with the capacity for strategic learning, thereby enabling them to modify strategies when

necessary. The unique features of BSC noted above suggest that its use can be expected to reduce the likelihood of disagreement and conflict between raters and ratees in performance evaluation. Lastly, BSC can be incorporated into executives' compensation plan.

The cause-and-effect chain is central to BSC. Recent studies, however, suggest that the causal links between driver and outcome measures are often overlooked. For example, Ittner et al.'s (2003b) study of financial services firms found that of those claiming to use BSC, 76.9% placed little or no reliance on their strategically linked causal business models. In another study of manufacturing and service companies, Ittner and Larcker (2003) noted more specifically that among those that created causal models, only 21% went on to validate the causal links between drivers and outcome measures. They also observed that "businesses often fail to establish such links partly out of laziness or thoughtlessness" (p. 89). Thus, the evidence from these studies indicates that, although information about strategy effectiveness can be available in BSC, it is not used as much as would have been expected, presumably because of cognitive limitations.

1.2 BSC Criticisms

BSC is not free from challenges. A key question is how to retain "balance" in managerial actions and performance evaluations in the presence of tradeoffs. Neely, Gregory and Platts (1995) criticized that BSC contains a flaw because, if a manager were to introduce a set of measures based solely on it, he would not be able to answer one of the most

fundamental questions of all – what are their competitors doing? Also, little research has been conducted on the implementation or performance consequences of BSC, despite widespread practitioner interest in the subject. Some advocates of EVA, which provides a single financial measure, claim that BSC of financial and non-financial measures hinders performance because there is no single overall measure of performance on which managers can concentrate their efforts to improve. Finally, a diverse set of performance measures may result in information overload, causing managers to spread their valuable efforts over too many objectives and reducing the effectiveness of BSC (Ittner and Larcker, 1998).

BSC might also make invalid assumptions about causal relationships, leading to the anticipation of performance indicators which are faulty, thus resulting in dysfunctional organizational and individual behaviours and sub-optimized performance. Nørreklit (2000) criticized that, if a cause-and-effect relationship requires a time lag between the cause and its effect, then it is problematic that the time dimension is not part of the scorecard. Because it measures causes and effects at the same time without considering any time lag, it has no time dimension. Nevertheless, the effects of the measures will occur at different points of time, because the effects of the different areas involve different time scales. While the introduction of more efficient processes may yield more satisfied customers within a period of three months, for example, innovation may not affect the financial results until a few years have passed. Because numerous factors may influence the results, it may be difficult to determine when the financial effect of an action will occur or what the long-term impacts of that effect will be. In the case of

excess demand, a minor cut in quality performance may similarly permit the handling of an additional workload. Likewise, any profitability derived from customer satisfaction or loyalty is neither a necessary outcome nor a highly probable one. It has been noted that customers who are not satisfied do not lead to financial success. However, this does not allow to conclude that satisfied customers lead to financial success. Atkinson et al. (1997, p. 27) viewed the modern corporation as “a complex web of contracts, both explicit and implicit, that specify relationships between the company and its stakeholders”. They believe that a company’s strategy is reflected by the relationships that the company negotiates with each stakeholder group. These contracts specify what the company expects from each stakeholder group to achieve its main objectives (financial performance) and what each stakeholder expects from the company in return for his/her co-operation. As a consequence they propose an approach to performance measurement that views the performance measurement system as the tool the company uses to monitor those contractual relationships. Since the approach of Atkinson et al. proposes simultaneously managing stakeholder relations, it is clearly different from Kaplan and Norton’s approach that proposes a sequential consideration of stakeholders through cause-and-effect relationships. Similarly, in response to the view of a linear causal link, Norreklit (2000) argued that the four perspectives of BSC are in fact interdependent.

Despite widespread practitioner interest in the subject, little research has been conducted on the implementation consequences of BSC. Ittner and Larcker (1998) found that studies investigating the link between customer satisfaction and future financial performance, such as stock return, have produced mixed results while many studies have found positive

associations between advanced manufacturing practices and the provision of non-financial measures, such as defect rates, on-time delivery, and machine utilization. The use of BSC does not mean just using more measures: it means putting a handful of strategically critical measures together in a single report, in a way that makes cause-and-effect relationships transparent and keeps managers from sub-optimizing by improving one measure at the expense of others. To achieve a balance among the four dimensions of BSC, a company should pay attention to all of them.

There is also the question as to whether the adoption of BSC improves organizational performance. Chenhall (2003) found that many studies do not employ contingency-based approaches as they explore only the main effects between share price movement and the adoption of elements of performance measurement. Progress in this area may be limited due to the difficulties in extracting the effects of adopting different performance measurement systems on share prices from other events that may be associated with share price movements. With the numerous possible events affecting share prices, control problems can become acute. Also, data collection is complicated by the need to collect data on the adoption and implementation of BSC by survey methods and to then match these with share price changes.

1.3 What Constitute a Successful BSC?

How to design a winning BSC? This question is essential to managers and researchers. BSC should consist of various mechanisms, both financial and non-financial, giving

sufficient consideration to the tradeoffs between efficiency and effectiveness. In general, a measurement package, which is properly designed and implemented, can serve as a convenient planning, monitoring, feedback, evaluation and motivational device in business administration. BSC can bring together, in a single management report, many of the seemingly disparate elements of a company's competitive agenda (Kaplan and Norton, 1992). BSC complements financial measures with operational measures on customer satisfaction, internal processes, and the organization's innovation and improvement activities, which are the drivers of future financial performance. Kaplan and Norton (1993) showed that BSC is most successful when it is used to drive the process of change. Successful BSCs have four characteristics: they are balanced, holistic, dynamic, and aligned with incentives.

First, the components of BSC should be balanced. The emphasis on traditional performance measures distracts from non-financial factors such as market share, customer satisfaction, productivity, product quality and employee satisfaction. One interesting property of BSC is its attribute “balanced.” This was described by Kaplan and Norton (1992, p. 73) as follows: “By forcing senior managers to consider all the important operational measures together, the BSC lets them see whether improvement in one area may have been achieved at the expense of another.” The logic behind this statement is that the firm’s strategy links all the different perspectives and their measures, thereby eventually characterizing the firm’s long-term objective. By this means, actions that are not value-adding, but improve the measured performance, can be ruled out.

A critical implementation issue that arises when incorporating multiple performance measures in reward systems is determining the relative weights placed on the four elements of BSC when determining bonuses. One option is to use a formula that explicitly weights each measure. Ittner et al. (2003a), however, warned that potential difficulties with this option include determining the appropriate weights to place on each measure, the “game-playing” associated with any formula-based plan, the possibility that bonuses will be paid even when performance is “unbalanced” (i.e., overachievement on some objectives and underachievement on others), and the likelihood that all relevant dimensions of managerial performance are not captured by the selected performance measures.

A second option is to introduce subjectivity into the bonus award process. This subjectivity can take the form of flexibility in weighting quantitative performance measures when computing a manager's bonus, the use of qualitative performance evaluations, and/or the discretion to adjust bonus awards based on factors other than the measures specified in the bonus contract. Baker, Gibbons and Murphy's (1994) theoretical work indicated that greater subjectivity can improve incentive contracting, because it allows the firm to exploit non-contractible information that might otherwise be ignored in formula-based contracts, and to mitigate distortions in managerial effort by “backing out” dysfunctional behaviour induced by incomplete objective performance measures. On the other hand, Ittner et al. (2003a) showed that the subjectivity in the scorecard plan allowed superiors to reduce the “balance” in bonus awards by placing most of the weight on financial measures, to incorporate factors other than the scorecard

measures in performance evaluations, to change evaluation criteria from quarter to quarter, to ignore measures that were predictive of future financial performance, and to weight measures that were not predictive of desired results. The high level of subjectivity in BSC plans led many branch managers in a major financial services firm to complain about favouritism in bonus awards and uncertainty in the criteria being used to determine rewards. The system ultimately was abandoned in favour of a formulaic bonus plan based solely on revenues. Although the ability to generalize their results is limited by the analysis of only a single firm, they provide one of the first detailed studies of scorecard-based compensation plans.

Second, BSC should be holistic. A holistic approach is needed to consider all the important measures at the same time, letting us see whether improvement in one area is achieved at the expense of another. Advocates of BSC suggest that each unit in the organization should develop and use its own scorecard, choosing measures that capture the unit's business strategy. Lipe and Salterio (2000) examined judgmental effects of BSC – specifically, how BSC that includes some measures common to multiple units and other measures that are unique to a particular unit affect superior's evaluations of that unit's performance. They demonstrated an important limitation to using BSC. Without providing a way to ease the cognitive burden on users, decision makers, when making comparative evaluations, will tend to focus attention only on measures common among managers and ignore measures unique to division manager. Under-use of unique measures reduces the potential benefits of BSC because the unique measures are important in capturing the unit's business strategy.

Lipe and Salterio (2000) attributed the common-measures bias to the superior's need to employ simplifying cognitive strategies. BSC typically requires four to seven performance measures in each of the four categories. As a result, evaluators using BSC could potentially have 16 to 28 indicators to process, holistically, in assessing the performance of a firm manager. Measures that are common across units tend to be lagging and financial indicators of performance, such as return on sales. In contrast, unique measures are more often leading and non-financial measures, such as market shares. Thus, ignoring the unique measures in BSC is tantamount to ignoring many leading indicators and focusing managerial attention more on lagging indicators. If the unique measures of BSC do not affect managers' decision, firms will not reap the expected benefits of BSC adoption.

Roberts, Albright and Hibbets (2004) extended the above study by using a "disaggregated/mechanically aggregated" BSC. They argued that superiors could utilize unique as well as common measures, thus overcoming the common-measures bias. Disaggregating the process, therefore, is one approach for improving effectiveness of BSC. Thus, common-measures bias could possibly be mitigated by either 1) requiring BSC users to evaluate performance on each BSC measure at a time, and then make a holistic judgment, and/or 2) suggesting weights for each measure. This would substantially lessen the cognitive burden.

The performance measures literature also suggests that BSC should be tailored for each business unit. A measurement system can be examined at three different levels (Neely et al., 1995). At the individual level, the generic terms quality, time, cost and flexibility encompass a variety of different dimensions. It is frequently argued that performance measures should be derived from strategy. At the entity level, there are two basic types of performance measures in any organization – those that relate to results (competitiveness, financial performance), and those that focus on the determinants of the results (quality, flexibility, resource utilization and innovation). At the environmental level, the performance measurement system has to be consistent with the organization's culture and consider customers and competitors. Benchmarking can be used to evaluate the external environment. Eccles (1991) predicted that a focus on quality would evolve naturally into strategies based on customer service as competition continues to stiffen. He identifies five areas of activity that need to be addressed: developing an information architecture; putting the technology in place to support this architecture; aligning incentives with the new system; drawing on outside resources; and designing a process to ensure that the other four activities occur.

In the original BSC concept, emphasis is placed on BSC as a shareholder rather than a stakeholder model. Kaplan and Norton (1996) argued against the development of the employees' own perspective, and included instead "an employee focus" in the learning and growth perspective. They claimed that a separate employee perspective may distract attention from the business strategy. In the Swedish BSC package, however, the employee perspective is very visible (Ax and Bjørnenak, 2005). The great majority of the

companies including ABB, Electrolux and SKF, for example, have developed their own employee perspective. The Swedish BSC package is more than a tool for cultivating a pure shareholder perspective; it also has definitive stakeholder characteristics. Some empirical findings show that the original BSC presented by Kaplan and Norton has been supplemented with other administrative innovations and adapted to the existing business culture to form a potentially more attractive set of elements. This bundling of BSC with other ideas can be seen as a fashion-setting process that affects the diffusion of performance measurement.

Third, BSC should be dynamic. The continued success of BSC is contingent on its environment. Most organizations have only a static performance measurement system (Bititci et al., 2000). Consideration is being given to what should be measured today, but little attention is being paid to the question of what should be measured tomorrow. Performance measures do not seem to change as products move through their life cycle (Neely et al. 1995). However, the appropriateness of established BSC systems must be re-evaluated in view of the current competitive environment, because measures tend to lose their relevance over time. BSC should reflect the context to which they are applied. Researchers, for example, have considered the influences of environmental uncertainty or technology in moderating the relationships between accounting systems and outcomes. Bititci, Turner and Begemann (2000) explored the use of IT-based management tools as a self-auditing “dynamic performance measurement system”, which would ensure that an organization’s BSC remains integrated, efficient and effective at all times.

Kaplan and Norton (1996) emphasized that BSC helps management in four steps: translating the vision, communicating the strategy and linking it to departmental and individual objectives, business planning, and feedback and learning. This is certainly a dynamic perspective. Kennerley and Neely (2002) grouped the change factors into drivers and barriers in terms of process, people, systems and culture, using a multiple case study approach. There are four stages of evolution – use, reflect, modify and deploy. A well-designed BSC should be accompanied by the existence of a process for reviewing, modifying and deploying measures; the availability of the required skills to use, reflect on, modify and deploy measures; the availability of flexible systems that enable the collection, analysis and reporting of appropriate data; and the existence of a measurement culture ensuring that the value of measurement and importance of maintaining relevant and appropriate measures are appreciated.

Therefore, it is important that BSC be dynamic, so that performance measures remain relevant and continue to reflect the issues of importance to the business. Yet, few organizations appear to have systematic processes in place for managing the evolution of their BSC.

Finally, BSC should be aligned with incentives. Although Kaplan and Norton (1996) argued that BSC should affect compensation, they provided no guidelines for this linkage. Depending on the evaluation style, reliance on BSC can result in dysfunctional consequences including disagreement and conflict between supervisors and subordinates. Roberts et al. (2004) found that performance and bonus allocations are highly correlated.

Wong-On-Wing, Guo, Li and Yang (2007) hypothesized that individuals in top-manager roles do not take into account strategy effectiveness unless they are explicitly required to do so. In contrast, individuals in the store-manager role automatically consider the quality of strategy without being prompted to do so. Yet, the studies of Lipe and Salterio (2000), Roberts et al. (2004) and Wong-On-Wing et al. (2007) suffered similar flaws, as they used only MBA students for experiments in place of managers. Because these studies were conducted in a controlled environment, they omitted many contextual factors that would be present in the natural organizational settings.

Although the general importance of such incentive provisions is hardly disputed, their implementation is still controversial. Kaplan and Norton (1996, 2001) considered explicit, formula-based incentive plans with some scepticism. They found that companies using BSC for incentive purposes experienced particular difficulties in quantifying measures based on the customer and learning and growth perspectives. Such measures also varied significantly across the companies. BSC can be used to influence behaviour. Rewards or sanctions reinforce or modify behaviour depending on the employee's performance and on the appropriateness of the use of performance information (Neely et al., 2005). BSC can also be used as a means of surveillance, motivation, monitoring performance, stimulating learning, sending signals or introducing constraints.

Performance and bonus allocations have been observed to be highly correlated (Roberts et al. 2004). Organizations that align their incentive plans' performance measures with contingency factors achieve higher performance (Ittner and Larcker, 2001). Significant

conflict and tension between superiors and evaluatees was observed when evaluation was perceived as subjective (Ittner et al. 2003a). Perceptions of subjectivity led to the rejection of BSC and a return to financial performance measures at a large firm. To be effective as a management control device, BSC should result in evaluations that are accurate, objective and verifiable.

2. Executive Incentives

What gets measured gets attention, particularly when rewards are tied to performance measures (Eccles, 1991). It is problematic to ask managers to focus on BSC measures if managers' compensation and evaluation are based on traditional financial measures. BSC compensation plans provide one method to mitigate agency conflicts by providing incentives for managers to make decisions that are in the best interest of the stakeholders. Unlike contingency theory operating at the organizational level, agency theory has focused primarily on the individual level. However, Kaplan and Norton (1996) did not provide specific recommendations regarding how to link BSC to compensation. This section offers a review of the debates on executive incentives.

2.1 Managerial Compensation

The existing performance evaluation and compensation literature is largely grounded in agency theory. The incentive problems resulting from separation of ownership and control have long been recognized in management accounting and corporate finance

literature (Jensen and Meckling, 1976). Managers who control the firm are supposed to act in the shareholders' best interests, but they may be tempted to take advantage of their control power to satisfy their own needs and desires, often to the detriment of the shareholders.

To inhibit poor managerial behaviour, shareholders try to align the executives' interests with their own through incentive programs involving stocks and stock options in addition to offering competitive salaries, bonuses and perquisites. Equity-linked compensation is a natural way to achieve this end. Jensen and Meckling (1976) analyzed conflicts between managers and shareholders and argued that to reduce agency costs a manager's compensation should be linked to shareholder wealth. When executives work hard to increase the firm's stock price over the long term, both the shareholders and the executives reap benefits. Murphy (1985) found that executive compensation is positively related to corporate performance as measured by shareholder return and growth in sales. Yet, he had no mention of efficiency and effectiveness of performance measurement. Since the various theories of managerial remuneration are based on diverse assumptions regarding states of nature, capital markets, and managerial characteristics, it is not surprising that competing models offer different sets of theoretical and empirical implications. Common to all theories, however, is the important result that, to some extent, compensation should be tied to observed productivity. Econometric studies have typically indicated that firm size (e.g., sales) is the only important determinant of executive compensation, and performance (e.g. return on equity, profits) plays at best a minor role. Yet, economic theories of efficient compensation suggest that, in addition to

current performance, contracts will depend on other factors such as entrepreneurial ability, managerial responsibility, and past performance.

Jensen and Murphy (1990) noted that companies paid their CEOs like bureaucrats rather than the value-maximizing entrepreneurs. Citing data on thousands of CEOs spanning five decades, they argued that the so-called “excessive” pay is not the biggest issue, because the real problem is how CEOs are paid. With respect to pay for performance, CEO compensation is actually getting worse rather than better. CEOs receive only \$3.25 for every \$1000 increase in shareholder wealth (p. 225). Therefore, they argued that more aggressive pay-for-performance systems and a higher probability of dismissal for poor performance should be introduced. Increases in compensation, which are driven by improved business performance, would not represent a transfer of wealth from shareholders to executives. Hall and Liebman (1998) documented that on average a 1% increase in a firm’s stock price led to an increase of \$124,000 in the CEO’s wealth in 1994, while in 1998 this value exceeded \$500,000 (p. 678).

Conyon and Murphy (2000) reported that both the level and the pay–performance sensitivity of CEO compensation increased sharply over the 1990s. One can interpret these studies in the 1990s as evidence of a period of high changeability and a commitment to high compensation. Murphy (1999) found widespread use of stock options in most industry groupings, but not in utilities. One can interpret this, again, as evidence of higher strategic discretion of CEOs and more changeable environments in growth industries. An alternative explanation is that higher compensation in high-growth

industries simply reflects higher agency costs of effort. Plausibly, however, while managerial jobs in growth industries may be more challenging and more onerous, this difference in effort seems unlikely to be large enough to explain very large divergences in compensation packages.

The level of pay has very little to do with whether or not CEOs have incentives to run companies in the shareholders' interests—incentives are a function of how pay, whatever the level, changes in response to corporate performance (Jensen and Murphy, 1990). The most powerful link between shareholder wealth and executive wealth is direct ownership of shares by the CEO. What really matters is the percentage of the company's outstanding shares the CEO owns. This will reduce the negative effects of agency problem. Second, cash compensation has certain advantages. Creating better incentives for CEOs means increasing financial risk that CEOs face. The weak link between cash compensation and corporate performance would be less troubling if CEOs owned a large percentage of corporate equity. Third, the prospect of being fired as a result of poor performance can provide powerful monetary and non-monetary incentives for CEOs to maximize company value. Compensation systems should be weighted toward the contributions made by individuals. Money is not everything, but unless non-monetary rewards vary positively with company value, they are not effective in motivating CEOs to act in the shareholders' interests. Prendergast (2002) considered a setting in which the agent has discretion in deciding how to solve the problem. If a problem is uncertain, the principal will not know how to solve it and thus will allow the agent to determine the solution, motivating him or her with strong incentive pay. On the other hand, if a problem

is predictable, the principal knows how to solve it, and will instruct the agent with regard to what to do and then verify compliance by directly monitoring his or her actions.

2.2 Do Incentives Work?

Whether or not the incentives work has resulted in many debates. Kim and Nofsinger (2006) found mixed evidence that executives manage earnings and other performance benchmarks to maximize their performance-based compensation. Also, these incentives or “rewards” potentially create yet another managerial temptation—the temptation to adopt inappropriate strategy or commit fraud to reap these rewards. Performance-related compensation creates an incentive to look for overly ambitious, hard to implement strategies. At a cost, shareholders can curb this tendency by pre-committing to a regime of CEO overcompensation in highly changeable environments. Alternatively shareholders can commit to low CEO pay, although this requires a commitment mechanism (either by the board of the individual company, or by the society as a whole) to counter the incentive to renegotiate upwards. Kim and Nofsinger (2006) proposed two ways to examine whether or not incentive-based compensation works. First, one could try to see if there is a positive relationship between firm performance and management compensation. This would be defined as *ex post* evidence. Another way to assess the efficacy of incentive-based compensation is to see if the firms that enacted these compensation mechanisms subsequently experienced superior performance. This can be defined as *ex ante* evidence.

The market uses accounting information to infer firm prospects and value. Given the recent increased linkage between manager compensation and firm value, it is important to examine the relation between compensation and accounting choices. Implicit in the argument that aggressive accounting can be used to affect stock prices is the assumption that manipulative accounting affects the financial markets. The significant average negative return around the announcement of a restatement of earnings suggests the market is not able to detect the “purposeful” use of accounting practices that led to the restatement, *ex ante*. Murphy (1985, 1999) found significant evidence that CEO compensation packages affect the adoption of aggressive accounting practices. This empirical study showed a strong causal relationship between the structure of managerial compensation and value-critical managerial decisions, specifically those derived from investment policy and debt policy. Ideally, the performance measures selected for compensation contracts are measures that best motivate management to maximize the value of the firm. While these actions are generally considered consistent with mitigating agency costs and therefore increasing shareholder value, the same actions can also be associated with sub-optimal decisions (e.g., reducing investment in otherwise worthwhile projects in order to avoid an explicit capital charge). Similarly, Core and Guay (2002) found a strong positive relation between firm risk and CEO pay–performance sensitivity.

2.3 Performance Measurement

Firms have traditionally measured and rewarded executive performance using financial metrics such as net earnings or return on investment (ROI). Annual bonus awards (also

referred as short-term incentives) that are based solely on financial results have been criticized for promoting an over-emphasis on short-term accounting returns and discouraging long-term investments. The usage of BSC has the potential to address the incentive problems caused by agency conflicts by directing managers' attention to long-term organizational performance.

Shareholder value has been recently accepted as an appropriate standard for performance (Ittner and Larcker, 2001). Advocates of EVA argue that this measure reflects risk and the time value of money and provides a framework for balancing profits and losses and balance sheet tradeoffs. Earnings-per-share is not the most important measure. Nor is growth for growth's sake. What matters is long-term cash generation. That is what drives long-term stock performance, and how executives should manage. Shareholder value analysis has been argued to be a powerful tool for deciding how to act. It can help managers get a handle on the value of all three types of corporate restructuring: portfolio restructuring, asset redeployment, and financial restructuring. Business-unit managers can use it to create value by expanding their focus from accounting profits and income statement effects alone to all the factors that affect shareholder value.

However, shareholder value analysis still only focuses on financial metrics. Effective measurement should combine both financial and non-financial measures. Ittner et al. (1997) examined the factors influencing the relative weights placed on financial and non-financial performance measures in CEO bonus contracts. They argued that such non-financial measures include market share, efficiency/productivity, product quality,

customer satisfaction and employee satisfaction. Two critical concerns of corporate executives are delivering consistent and predictable financial results, and ensuring that customers are satisfied (Lambert, 2001). When properly executed, both of these activities will in turn deliver increased value to shareholders. To continually succeed, it is imperative that executives stay abreast of market conditions and adjust strategic business objectives when necessary. Executives must operate their businesses nimbly enough to shift from one major initiative to the next within a very short timeframe. Whether the key initiative is revenue growth, driving a competitor out of the market, reducing discounts, expanding internationally, or introducing a new product line, the executive incentive program can play a crucial role in aligning the actions and the pace of strategic changes adopted by management with the creation of shareholder value. BSC costs and benefits must be balanced in the first place. An optimal BSC incentive contract involves a pay-for-performance scheme which ties the managers' pay-off to production indicators that correlate with his effort level.

3. Theoretical Foundations

Research on performance measurement has been linked to several theories, such as contingency theory, agency theory, institutional theory and anthropological theory (Berry, Broadbent and Otley, 2005). In particular, executive compensation studies are often based on agency theory, whereas performance measurement research has relied more on other theories. The fundamental appeal of contingency theory has resulted in a constant and continuing stream of studies seeking to explore the contingent nature of

accounting (Chapman, 1997). Perhaps the lack of contingency research in executive incentives is due to the difference in the levels of study. Agency theory is suitable for individual level studies, whereas contingent theory is appropriate to organizational level research.

3.1 Contingency Theory

The roots of contingency theory lie in the organization theory literature back to the 1960s (Otley 1980). As a milestone in organizational design, this theory has developed quickly since the early 1980s. By contending that the appropriate control mechanisms depend upon the circumstances surrounding an organization, this theory has significantly expanded Anthony's (1965) attempt to create the best and universal formal systems for organizations to implement strategies and objectives. Early researchers investigated the importance of environment, technology, uncertainty and structure for organization designs. Today, contingency theory is supported by an impressive body of literatures, and is capable of providing guidance to this thesis.

Scientific management implied that there exists one best way to design performance measures in order to maximize efficiency and effectiveness. Contingency theory arose in response to the universalistic approach, arguing that the design and use of performance measurement depends upon organizational and environmental context (Hansen et al., 2004). A better match between BSC and contextual variables, such as external environment, technology, structure, size, culture and strategy, is hypothesized to result in

increased organizational performance (Fisher, 1998; Sila, 2007)), while poor fit implies diminished performance. Thus, according to contingency theory arguments, a universalistic control approach does not exist.

Performance measurement has been conventionally perceived as passive tools providing information to assist managerial decision making. Yet, approaches following a sociological orientation see it as more active, furnishing individuals with power to achieve their own ends (Chenhall, 2003). Organizations purposefully adapt their structures in order to maintain fit with changing contingency factors. Donaldson (1995) argued that the implied role of management is vague but positive. Managers are able to rationally orchestrate this adaptation to environment and thereby seek to maintain a high performance for their organization. The move from misfit to fit is adaptive change that is the essence of contingency theory. Thus the contingency theory views management in a positive way, as the controllers coordinate the adaptation of the organization to its environment through implementing better-fitting control systems. This approach helps understand how and why different performance systems and BSC become established and widely diffused in different contexts.

Much of the empirical research based on contingent theory involves a search for systematic relationships between performance measurement system and particular environmental variables. Waterhouse and Tiessen (1978) reported that the issue of authority and power distribution is crucial to understanding of the control processes within an organization. The distribution of organizational authority and the extent to

which procedures can be specified are contingent on technology and environment. Fisher (1998) identified five contingent factors: 1) task and external environment uncertainty, 2) firm technology and interdependence, 3) industry, firm and unit variables, such as size, diversification and structure, 4) competitive strategy and mission, and 5) behaviour and outcome observability. Chenhall (2003) provided a critical review of findings from contingency-based studies, deriving a series of propositions relating control to organizational contexts, in terms of external environment, generic concepts of technology, contemporary technologies, organizational structure, size, strategy and culture. Hansen et al. (2004) spotted eight organizational factors (size, strategy, structure, systems, resources, staff skills, management philosophy, and organizational norms) and seven environmental factors (technology, competition, economy, culture, social values, regulation and legislation, and politics. Pollanen (2001) found that more emphasis has been placed on internal factors over which firms can exercise reasonable control rather than external factors that are largely uncontrollable forces.

Taking a contingency theoretic perspective, Hoque and James (2000) considered the size of organizations, product life-cycle stage, and market position as potential contextual factors of BSC usage and explored how organizational performance is affected by different uses of BSC in different settings. Yet, this study is confined only to manufacturing firms instead of using a cross-sectional design. Ittner et al. (1997) found that the relative weights placed on non-financial measures is greater in firms following an innovation-oriented “prospector” strategy than in firms following a cost leader or “defender” strategy. Firms that have adopted quality-oriented strategies also place more

weight on non-financial measures, as do utilities and telecommunication firms, many of which face regulatory and competitive pressures to improve non-financial dimensions, such as safety and customer satisfaction.

3.2 Cons and Pros

While contingency theory can provide significant insight into the antecedents and consequences of measurement systems, it also has faced serious criticisms. Chenhall (2003) argued that the definitions and measures of variables have often lacked conceptual clarity. The term contingency means that something is true only under specified conditions. As such, there is no “contingency theory” but rather that a variety of theories may be used to explain and predict the conditions under which particular measurement systems will be found, or where they will be associated with enhanced performance. In addition, the full set of possible contingent factors has not been, or is not possible to be, identified. Other variables, such as ownership, merger and acquisition, uncertainty, and product-life-cycle, may also play a role in the adoption of BSC. The relationships among these variables also need to be explored. Chenhall further contended that contingency-based research has focused on specific elements of management controls, with a limited number of studies examining broader elements of control. The absence of a process element in contingency theory is another shortcoming. Chapman (1997) criticized that, despite a promising start and many subsequent studies, the literature has not yet developed a credible and comprehensive contingency theory of accounting. A refined contingency theory should be able to identify specific aspects of an accounting system,

which are associated with certain defined circumstances and demonstrate an appropriate matching (Otley, 1980).

Yet, Donaldson (1995) contended that contingency theory was supported by an impressive body of literatures, and is capable of providing guidance for the design of BSC systems. It is plausible to assume, from the pragmatic managerial point of view, managers can play a constructive role in selecting and implementing fitting measurement systems to achieve the objectives of their organizations. Theories that are nurtured in many interrelated subjects play a major role in the development of performance measurement and BSC. He further acclaims, "Contingency theory remains as a cogent body of theory which is intellectually sound and empirically valid, certainly more so than any of the would-be alternative theories" (p. 215). Theory construction and empirical research interact. Just as theories stimulate empirical work, rich empirical settings stimulate theory. Likewise, Otley (1980) maintained that the justification for adopting a contingency perspective of performance measurement was that it emerges as a necessary means of interpreting the results of empirical studies at the organizational level.

3.3 The Concept of "Fit"

The concept of "fit," which originated from contingency theory and population ecology, occupies a central role in theory construction of many management disciplines (Venkatraman, 1989). It involves the matching or the art of reconciling the various organizational elements, such as strategy, structure, systems, style, staff, shared values,

and skills, and how such elements affect the relationships of interest among key variables in a specific study (Venkatraman and Camillus, 1984, p. 514). This congruence is a prerequisite for organizational success. Milgrom and Roberts (1995) proposed the idea of complementarity, which gives substance to previously elusive notion of “fit”. Their econometric model provides some basis for interpreting claims such as the need for strategy and structure to fit each other, give an approach to model such issues formally, enrich the directions of causation, and suggest reasons why “fit” may be hard to achieve and change may be slow, painful and uncertain. Management literature often uses this notion to investigate the contingent relationship between management systems, their contextual variables, and organizational performance. The “fit” in related research has been measured as moderating or mediating effects, analyzed through multiple regression techniques (Venkatraman and Camillus, 1984). Tests of fit as mediation can also be carried out using path-analytic frameworks. The relative proportion of the two effects – indirect versus direct – provides an index of the relative effect of fit (namely, indirect effects) versus the direct effects (Venkatraman, 1989). Recent studies have proposed structural equation modeling as a technique for testing “fit” (McDonald and Moon-Ho, 2002).

Many attempts have been made recently to construct the fit between performance measurement and its contextual variables. Each attempt is motivated by a conviction that the latter is inadequate to explain the former. Chenhall (2005) showed that the influence of strategic performance measurement systems on firm performance is indirect through the mediating roles of alignment of manufacturing with strategy and organizational

learning. Banker et al. (2000) investigated the relation between non-financial measures and financial performance. By using moderated regression analysis, Hartmann and Moers (1999) tested contingency hypotheses that predict interaction effects between budgetary and contextual variables. Building on fitness landscape theory, Jermias and Gani (2005) defined contingent-fit as the weighted sum of the independent contribution of each contextual variable to the chosen strategy. HassabElnaby et al. (2005) viewed the fitted values as the precise extent to which the firms have invested in BSC performance measures.

4. Summary

BSC can provide performance measures that cut across traditional functional areas of an organization. The integration of both financial and non-financial measures into a unified framework has become popular. There is a need to satisfy potentially competing requirements of multiple stakeholders. Performance management should identify these requirements to attract and maintain shareholders, employees and customers and to do so in ways that are socially acceptable (Kaplan and Norton, 1992; Chenhall, 2003). BSC stresses activities and processes that drive cash flows. It therefore enables companies to track financial results, while simultaneously monitoring processes in building the capabilities and acquiring the intangible assets they need for future growth. Much more than a measurement exercise, BSC is also praised by its authors as an incentive system that can motivate executives to maximize stakeholders' value and achieve other performance objectives. BSC contracts can be structured in many ways for executive

incentives. One critical implementation issue that arises when incorporating multiple performance measures in reward systems is determining the relative weights placed on various measures when determining bonuses. BSC, with its specification of causal relationships between performance drivers and objectives, allows executives to use their periodic review sessions to evaluate the validity of the unit's strategy and the quality of its execution.

Organizational and external environments contain many powerful contextual variables that are at the foundation of contingency-based research (Sila, 2007). Contingency theory suggests that organizations that can establish a fit between BSC and environmental context will achieve higher organizational performance, while a misfit would have negative effects on organizational performance. The primary feature of agency theory that has made it attractive to performance measurement researchers is that it allows them to explicitly incorporate conflicts of interest, measurement problems, uncertainty, and controlling mechanisms into measurement systems. This theory provides a framework for rigorously examining the link between managerial behaviours and conflicts with BSC.

III. Conceptual Model and Hypotheses Development

This chapter explains research models and develops hypotheses. Ryan, Scapens and Theobald (1992) argued that empirical research was generally concerned with establishing relationships among variables. That is, for example, how does one variable change as another variable changes? This thesis will probe relationships between contingent factors and BSC usage and explores how corporate performance is affected by the usage of BSC in different settings. The “fit” hypotheses in this study relate to the motivation that greater reliance on BSC results in increased performance of organizations. The model and hypotheses are based on extensive reviews on extant research.

1. Framework and Model

This section lays a conceptual foundation for the study. The model is based on previous research, yet, it is still a priori model as it is untested. The results provided in later chapters will test how well the model works, and accordingly, a revised model, or ex-post model, will be proposed.

Figure 3.1 highlights the relationships among the variables included in this study. Stage 1 focuses on how contingent variables affect the usage of BSC in executive compensation plans. Stage 2 examines whether BSC usage in executive compensation plans has an impact on organizational performance. Stage 3 tests performance consequence of the “fit”

between BSC and firm characteristics. The variables considered in the model are dichotomized as dependent or independent variables. Ryan et al. (1992) observed that an independent variable is the variable that is manipulated by the researcher, and it is the effect of this variable that is being studied by the researcher. The dependent variable measures the response to the manipulation of the independent variable. Thus, the researcher is interested in determining the impact of changes in the independent variable upon the dependent variable.

Figure 3.1: Framework for the study

According to contingency theory, the design and use of performance management depends upon organizational and environmental contexts. This study involves an analysis of systematic relationships 1) between BSC usage and particular contextual variables, including strategy, industry, size, quality, structure, culture and ownership, 2) between

BSC usage and organizational performance, and 3) between contingent fit and organizational performance.

Despite the critique that contingency-based studies should include organizational performance as the dependent variable, many studies still tend to focus on contingent variables and use system design characteristics such as BSC as dependent variables (Chenhall, 2003). By contrast, this thesis tests whether the adoption of BSC enhances corporate performance in terms of effectiveness and efficiency for treatment firms, compared to matched control firms. This will extend several recent attempts. Wallace (1997) included corporate performance as a dependent variable, while using the adoption of residual income as an independent variable. Ittner et al. (2003b) used customer growth and financial performance as dependent variables and non-financial performance measures as independent variables. Hoque and James (2000) treated organizational performance as dependent variable and organization characteristics and BSC usage as the independent variables. These studies can provide important insights into the extent of adoption, use and usefulness of BSC. However, it should not be assumed that the constructs modelled in this study the models necessarily lead to enhanced organizational performance. The reverse relationship is also possible, that is, well-performing firms adopt more sophisticated incentive systems. Similarly, if performance is the dependent variable, then a compelling theory is required to show how the combination of BSC and organizational characteristics enables managers to take more effective decisions that enhance corporate performance.

The model in Figure 3.1 depicts two sets of direct effects: from A (contingent variables) to B (BSC usage), and from B (BSC usage) to C (organizational performance). Indirect effects between A and C are also possible. A key assumption is that the “fit” between BSC usage and its context can improve organizational performance. The higher the level of fit, the better the performance. Jermias and Gani (2005) considered “contingent fit” as one of the most appropriate approaches for examining the relationship between evolving system and performance. Organizational performance in this thesis is studied along two major dimensions, effectiveness and efficiency, as discussed in Section 3 of this chapter.

2. Hypotheses - BSC Use in Executive Compensation and Contingency Factors

Contingency theory advocates that the design and use of BSC depends upon organizational and environmental context. A better match between BSC and contextual variables is hypothesized to result in increased organizational performance, while poor fit implies diminished performance. There are a myriad of possible contextual variables that may affect the design and the effectiveness of BSC. Important contextual factors are strategy, industry, size, structure, technology, culture, quality, ownership and so on. Hussain and Hoque (2002), for example, examined the performance measurement practices of four Japanese banks. The results indicate that several institutional forces influenced the banks to implement a particular performance measurement system. Of these, economic constraints appeared to be the most forceful factor, followed by the central bank’s regulatory control, accounting standards and financial legislation,

management's strategic focus, bank size, competition, and organizational tendencies to copy best practices from others.

2.1 Strategy

Prior research argues that performance measures used in compensation contracts should be closely linked to the organization's strategy to ensure that managers' incentives are aligned with the company's goals (Govindarajan and Gupta, 1985; Simons, 1990). Corporate strategy is concerned with the way resources are focused to convert distinct competences into competitive advantage. Some research follows a positivist approach, assuming that strategy is an outcome of rational choice. Alternatively, strategy may be considered a craft (Chenhall, 2003). BSC, as part of resources, should be tailored explicitly to support the strategy of the business to lead to superior performance. Contingency-based research predicts that certain types of performance measures will be more suited to particular strategies. The studies indicate that there are systematic differences in performance measurement systems among firms that compete in different ways. Hoque and James (2000) found that BSC usage is positively associated with businesses with products at the early/growth stage. Simons (1990) and Langfield-Smith (1997) showed that strategies characterized by a conservative orientation, defenders, harvest and cost leadership are best served by centralized control systems, specialized and formalized work, simple coordination mechanisms, diagnostic controls, and attention directing to problem areas. Strategies characterized by an entrepreneurial orientation, prospectors, build and product differentiation are linked to lack of standardized

procedures, decentralized and results-oriented evaluation, interactive control, flexible structures and processes, complex coordination of overlapping project teams, and attention directed to curbing excess innovation. Non-financial measures, subjective judgment and behaviour control are considered appropriate for this type of strategy.

Of the many frequently studied aspects of strategy, Jermias and Gani (2005) focused on low cost and differentiation. Since costs are typically easily determined and outputs are more observable, low-cost business units tend to use output measures for maximum effectiveness. The type of formal performance measures appropriate for companies emphasizing a low-cost strategy focus mainly on controlling costs. Thus, traditional financial mechanisms, such as budgetary performance measures, variance analyses, and activity-based costing, are suitable for these companies. Product differentiation business units tend to have difficulty implementing comprehensive control systems due to changing environment. Behaviour measures are considered appropriate for this type of strategy. These types of business units will benefit more by using systems that provide measures of customer satisfaction, timely and reliable delivery, and measures of key production activities, quality, benchmarking, employee-based measures, and strategic planning. Rather than providing comprehensive coverage, Govindarajan and Gupta (1985) emphasized the relationship between corporate strategy and one particular aspect of performance management - bonus remuneration. Long-term evaluation criteria and subjective, non-formula bonus calculations are effective for businesses following a “build” strategy, but detrimental to business units pursuing a “harvest” strategy. Interactive reward systems based on subjective evaluation of effort are appropriate for

firms that need to motivate organizational learning in rapidly changing environments and where rewarding team effort is important - typical conditions of firms in a growth phase. This approach is, however, generally uneconomic for businesses in slow decline.

The corporate strategy literature suggests that competitive strategy can be broadly conceptualized as a continuum between two different strategic orientations (Porter, 1980). At one extreme, organizations are characterized as prospectors or firms that exhibit differentiation strategy. These firms attempt to identify new product/service market opportunities, quickly adapt to changes in the external environment, and follow a “first-to-market” strategy. At the other extreme, organizations are characterized as defenders or firms that exhibit a cost leader strategy. These firms attempt to provide a stable set of products and services to a well-defined portion of the total market while emphasizing improvements in current operating efficiencies in order to lower costs. The latter tends to focus on financial measures, whereas the former favours to non-financial measures. Ittner et al. (1997) showed that the relative weight placed on non-financial measures was greater in firms following an innovation-oriented “prospector” strategy than in firms following a cost leader or “defender” strategy.

The above review leads to Hypothesis 1A that predicts differential use of BSC for firms at the different ends of the strategy spectrum from cost leader (defender) to differentiator (prospector).

H1A: The usage of BSC in executive compensation plans is positively associated with product differentiation strategy (prospectors), but negatively associated with cost leadership strategy (defender).

2.2 Industry

Research has related some specific industries to the adoption of BSC. Ely (1991) found that the choice between alternative performance measures varies by industry, indicating that these measures must be tailored to reflect industry-specific values and surrounding environments. Speckbacher et al. (2003) compared the companies which have adopted a BSC and the companies which have not. They found that firms belonging to the “consumer and retail” industry are associated with a significantly lower usage of BSC; but they found no significant association for other industries. Also, regulators may have a particular interest in the stakeholder perspective of BSC. Ittner et al. (1997) argued that a situation where non-financial performance measures are expected to be relatively more informative about the managerial actions desired by the owner arises when public utility regulators link profits or rate increases to the achievement of non-financial goals, such as customer satisfaction, reliability and employee safety. They found that utilities and telecommunication firms place more weight on non-financial measures. Indeed, organizations create control mechanisms to look legitimate to important stakeholders (Sila, 2007). This is due in part to regulations, procedures and structures imposed on them by environmental institutions and conditions. A company, for instance, may be forced to change their performance measurement system due to governmental pressures,

imitates the performance metrics of other organizations as a result of competitive pressures, or conforms to normative standards developed by industrial associations as accreditation (DiMaggio and Powell, 1983). A good example is a change in internal controls in response to the U.S. government-mandated Sarbanes-Oxley Act. Thus, this study views organizations as embedded within larger inter-organizational networks.

Breakthrough technologies have been among the main driving forces for change in today's markets in some industries. Waterhouse and Tissen (1978) argued that organizational authority is contingent on technology. A just-in-time environment tends to be best suited to open, informal, organic forms of performance measures. Flexible manufacturing is associated with performance measures focused on time, quality, operating efficiency and flexibility. Chenhall (2003) stressed that organizations producing highly specialized, non-standard, differentiated products were likely to employ complex unit/batch technologies. A need for flexible responses to specific customers increases interdependencies across the value chain, involving reciprocal interactions with customers, suppliers and functional units. These types of technologies would require controls to encourage flexible responses, high levels of open communication within the work force and systems to manage interdependencies. Traditional, mechanistic measures based on financial controls would not seem to suit these circumstances. Organizations that produce standard, undifferentiated products employing capital-intensive, automated processes are likely to employ mass production and process technologies, which will involve highly analyzable processes and few exceptions. Knowledge of processes and measures of output will be more readily available. Thus technology-intensive companies

require standardized, administrative controls such as traditional, formal financial controls. Yet, more research is needed to investigate how performance measurement is best suited to different stages in the growth of firms.

Thereby, this study hypothesizes that some regulated industries (such as utility and telecommunication), will place more emphasis on BSC measures in the annual bonus contract, whereas other industries, such as consumer and retail companies are less associated to the usage of BSC.

***H1B:** The usage of BSC in executive compensation plans is affected by the type of industry.*

2.3 Size

The contingency theory literature suggests that the size of organizations is related to the design of organizational control and performance measurement. Murphy (1985) showed that firm size yields information relevant for determining levels of managerial effort. Indeed, several theories of managerial production suggest that compensation should be partially determined by firm size or growth, reflecting the quantity of resources controlled by the individual executive and the scope of managerial responsibilities. Chenhall (2003) argued that performance measurement systems tend to be more specialized and sophisticated in larger firms. Two forms of control have been identified to associate with size: administrative with large firms and personal with small firms.

Based on a survey of 66 Australian manufacturing companies, Hoque and James (2000) discovered that BSC usage is positively associated with organization size. Larger firms are associated with more decentralized organizational structures and more specialized functions and processes and, therefore, co-ordination and communication problems increase with size. As a broader set of information and measurement issues arises in larger firms, more advanced and more sophisticated management accounting systems are required. Based on this reasoning, it seems plausible that larger organizations are more likely to use BSC in executive compensation plans than smaller ones. The “fit” hypothesis proposes that because large organizations depend more on sophisticated information and control systems using diverse measures, they derive more benefit from reliance on BSC than small organizations.

Thus, contingency theory suggests that size may affect the way organizations design and use BSC. As firm size increases, accounting and control processes tend to become more specialized and sophisticated. Yet, compared to large manufacturing firms, smaller-sized entities have received little attention in the contingency-based performance measurement literature. This study will also address the above flaw by including companies with different sizes. Based on this a priori reasoning, the following hypothesis is proposed:

H1C: Larger organizations place a greater reliance on BSC measures in executive compensation plans than smaller organizations.

2.4 Quality

According to the quality management literature, successful total quality management requires greater reliance on non-financial quality measures in order to foster management commitment to the quality program, to communicate the significance of quality management to employees, and to ensure that quality improvement results are elevated to the same level of importance as financial performance. Quality management is linked to timely problem-solving information and flexible revisions to reward systems. Since the benefits from current quality improvement activities may not be fully reflected in short-term financial measures, Ittner et al. (1997) argued that non-financial quality measures provide additional information regarding current managerial performance. Firms that have adopted quality-oriented strategies also place more weight on non-financial measures. Many of these firms also face regulatory and competitive pressures to improve non-financial dimensions, such as safety and customer satisfaction. These companies tend to receive quality awards that show their commitment to improving competitiveness and the pursuit of performance excellence. The relationship between quality and the usage of BSC is presented in the following hypothesis:

***H1D:** The usage of BSC in executive compensation plans is positively related to a firm's use of quality management programs.*

2.5 Structure

Structural arrangements remain an important factor in understanding the design of BSC.

Growth enables firms to improve efficiency and results in specialization and increased division of labour. Large firms that are decentralized have been associated with a strong emphasis on formal controls; and large, diverse, decentralized firms use more administrative controls (Chenhall, 2003). Performance measurement in decentralized organizations is more complex. Anthony and Govindarajan (2007), according to the nature of monetary inputs and/or outputs that are measured for control purposes, classified firms as revenue, expense, profit and investment centers. Waterhouse and Tiessen (1978) conjectured that decentralization, which is broad decision-making discretion at lower levels, existed because operators have access to information that is not available to their managers. Consequently procedures are difficult to specify and document. The implication is that the superior must rely on subordinates for information on efficient means for coordinating and operating. Thus, decentralization or divisionalization has advantages. Decisions can be made at the place of their execution, where there is presumably better knowledge of environmental conditions. This structure provides a means of filtering information flows to the top managers who could focus on strategic issues. Decentralization enables the quicker allocation of resources. But one problem is that communication difficulties up and down the hierarchy could lead to loss of central control. Decisions that were suboptimal for the organization as a whole might be adopted because they were optimal for a particular division.

Organizational structures are not only shaped by environment contingencies, but also by the need to fit with preconceived notions in the wider community about forms of organization which are legitimate, effective and rational (Donaldson, 1995). The central tenet is that organizations are pressured to become isomorphic with an environment of institutionalized rules, practices, symbols, beliefs and normative requirements in order to receive support and legitimacy (Abernethy and Chua, 1996).

These arguments provide the basis for H1E. Decentralized companies, which typically have greater number of and more independent segments, are more likely to use BSC in executive compensation plans. This thesis is the first known study to investigate such relationships.

***H1E:** The complexity of corporate structure in a firm is positively associated with the usage of BSC in executive compensation plans.*

2.6 Culture

Compared to studies of other contextual variables, research into culture is somewhat exploratory. As multinationals expand globally, they are forced to cope with multiple national cultures, languages and differing environments. National culture impacts on a firm's ability to transfer competencies and standardize processes and work practices. If the depth and width of globalization is limited by the reach of cultural boundaries, does it not logically follow that national cultures impose constraints on the usage of BSC? Can

BSC become universal while culture remains local? Over the past two decades, national culture has become important in the design of performance measures, as many companies have developed multinational operations. These companies face the issue of whether to transfer their domestic measurement systems overseas, or redesign their systems to fit the cultural characteristics of the offshore entities. The literature supports the use of corporate culture as a way of uniting a dispersed organization through commitment to a common purpose and behaviour built on shared values. Empirical findings provide evidence of the ability of the international firms to transplant their organizational practices, which are consistent with and supported by their home culture, into their overseas affiliates. Chow, Harrison, McKinnon and Wu (2002) showed differences in the organizational cultures of Taiwanese local firms and Taiwanese affiliates of the “Big Five” accounting firms. The local firms demonstrated organizational culture characteristics consistent with Taiwanese national cultural attributes, but the affiliates demonstrated organizational culture characteristics more consistent with US national culture. In a study of strategic investment decisions, Carr and Tomkins (1998) reported that U.S., German and Japanese firms pay greatest attention to competitive advantage analysis as compared to financial analysis techniques such as discounted cash flow. While the U.S. firms pay almost as much attention to financial numbers as competitive advantage analysis, the U.K. firms give the greatest weight by far to financial measures. They conclude that national culture influences management styles and financial analysis approaches.

The relationship between the design of BSC and culture represents an extension of contingency-based research from its organizational foundations into more sociological concerns (Chenhall, 2003). There is no universally agreed-upon definition of culture among social scientists. The focus is on the “sharedness” of cultural indicators among members of the collectivity: “shared motives, values, beliefs, identities, and interpretations or meanings of significant events that results from common experiences of members of collectives that are transmitted across age generations” (Chhokar, Brodbeck and House, 2008, p.3). National culture represents the shared values of people within a certain national environment. Measurement systems and practices are man-made artefacts that take place in an environment that is culturally predefined. Values are learned from the socialization process in which the individual participates as a member of society. Different organizations and countries possess particular cultural characteristics. BSC, therefore, is never culturally neutral, but is influenced by and influences the cultural context in which it is executed. This predisposes individuals from within these cultures to respond in distinctive ways to performance management. Practitioner-oriented books, for example, revealed the secrets of the success of Japanese companies: a strong, clan-like and commitment-oriented organizational culture, shaped by the basic rules of Japanese society (Scheytt and Soin, 2005). In a study of the relationships between cultural dimensions and human relationship management practices across twenty-four countries, Schuler and Rogovsky (1998) found that national culture was an important factor in multinationals’ use of compensation practices. They found that pay-for-performance individual incentive compensation practices, e.g. share options and stock-ownership, have a better fit in countries with higher levels of individualism and lower levels of

uncertainty avoidance and power distance. In such environments, the managers of strategic business units become more reliant on leading indicators.

The situation that BSC faces in France provides some insight into the relationship between BSC and culture. BSC has attracted a great deal of attention, especially in the United States, but also in many other countries such as Germany, U.K. and Italy. It has not, however, received a particularly warm welcome in France, where the *tableau de bord* has been used for 50 years. One of the explanations for the French reluctance to adopt BSC may be that, French firms have used the *tableau de bord*, which 100% of the French companies in the survey reported that they used. The *tableau de bord* is in many ways similar to BSC. Bourguignon, Malleret and Nørreklit (2004) stated that BSC and the *tableau de bord* were coherent with the local ideologies in the countries of origin. BSC creates a hierarchy through performance measurements and rewards, which is not useful in France, where hierarchies and a local sense of obedience are constructed otherwise: the demand for fairness and the subsequent need for performance measurements are not very important in a society based on honour, rank and privilege. The “mechanical” top-down deployment of BSC disregards the “incremental and collective construction” of strategy in France. The focus in BSC on accountability and external control may fit badly with the French concept of honour because it assumes self-control, which is more closely related to self-provided rewards and self-realization and which creates a defensive attitude towards external control and performance measurements. As it does not link rewards to performance measurements, the *tableau de bord* matches the French ideology better than does BSC.

Ax and Bjørnenak (2005) looked at the communication, diffusion and transformation of BSC in Sweden. Similarly, they stressed the importance of making innovations compatible with the societies to which they are transmitted. In the original BSC concept, emphasis is placed on BSC as a shareholder rather than a stakeholder model. Kaplan and Norton (1996) argued against the development of the employees' own perspective, and included instead "an employee focus" in the learning and growth perspective. They claimed that a separate employee perspective may distract attention from the business strategy. In the Swedish BSC package, however, the employee perspective is very obvious. The great majority of the companies including ABB, Electrolux and SKF, for example, have developed their own employee perspective. This might mean that the perspectives have been developed not just as causal factor of improved financial performance, but rather as perspectives with their own objectives and measures. In this sense, the Swedish BSC package is more than a tool for cultivating a pure shareholder perspective; it also has definitive stakeholder characteristics.

For that reason, hypothesis 1F presents the relationship between culture and BSC. While rigorous research has been conducted in relation to performance measurement systems, little research has been done regarding the implications of BSC for multinational corporations operating in a global environment. This study aims to address this deficiency, and explore the role of BSC within the processes of managing multinational companies and the way in which financial and non-financial resources are integrated. The

cultural cluster of parent firms is used as a proxy for national culture, because local firms are shaped by national culture of their parent companies.

***H1F:** The usage of BSC in executive compensation plans is positively associated with firms whose culture is consistent with the North American culture.*

2.7 Ownership

Prior research suggested that institutional owners that hold large blocks of shares have both the incentives to monitor management and the voting power to bring pressure to bear on management to effect change when control problems surface (Ashbaugh-Skaife, Collins and Kinney, 2007). However, concentrated ownership, especially when the controlling shareholder is either an individual or a family member, will have a negative impact on a company's performance (Jermias and Gani, 2005). This thesis also posits that managers of firms with more concentrated ownership face greater incentives to adopt BSC due to increased monitoring and greater litigation threats from concentrated owners. Hence, this study predicts a positive relationship—the higher the concentration of institutional ownership, the more likely a firm is to adopt BSC-based performance measures in executives' compensation plan.

***H1G:** Firms with a higher concentration of institutional ownership are more likely to adopt BSC-based compensation plans.*

This set of hypotheses makes unique contributions to the literature. The hypotheses incorporate several contingent variables and test their associations with BSC use in executive compensation plans, some for the first time, such as structure and culture.

3. Hypotheses - BSC use in Executive Compensation and Corporate Performance

Explicit consideration of organizational performance is a vital part of contingency theory research. It has been a much neglected topic from a theoretical stance, and its development is urgently needed since many surveyed studies do not even attempt to measure performance (Otley, 1980). In response to this call, this thesis will contribute to the literature by investigating whether and how the usage of BSC in executive compensation affects corporate performance.

It has been broadly accepted that employee compensation should be tied to BSC measures. Both Kaplan and Norton (2001) and Banker et al. (2000) provided empirical evidence of this assertion, through examples of alteration in employee behaviour or the organization's financial performance, after remuneration had been related to non-financial measures. According to economic theories of compensation, firm performance should affect an executive's remuneration only to the extent that it serves as a proxy for unobservable managerial effort or productivity. Although these theories uniformly suggest a relationship between compensation and observed performance, most analyses are silent about the measurement of performance. Since firm shareholders are generally considered to be the principals in agency theory research, it seems appropriate to define

performance in terms of shareholder returns rather than in terms of accounting profits (Murphy, 1985).

Budde (2007) studied the incentive effects of a BSC within a multi-task agency framework under both formal and relational contracts. First, the main characteristics of BSC were analyzed with respect to performance measure congruity. It is shown that under complete verifiability, a properly designed BSC is capable of perfectly aligning the interests of owners and employees by means of an explicit contract. He then investigated whether subjective performance evaluation is beneficial when not all scorecard measures are contractible. It emerged that congruity of the contractible scorecard measures constrains a purely implicit incentive contract, but the first-best solution may still be obtained through a combination of formal and relational contracts. Furthermore, a purely explicit contract in most cases can be improved by incorporating subjective rewards.

Kaplan and Norton (1996) assumed the following causal relationship: from measures of organizational learning and growth, to measures of internal business processes, to measures of the customer perspective, and to financial measures. The measures of organizational learning and growth are therefore the drivers of the measures of the internal business processes. The measures of these processes are in turn the drivers of the measures of the customer perspective, which are the subsequent drivers of the financial measures. A good BSC should have a mix of outcome measures (lag indicators) and performance drivers (lead indicators). The causal paths from the measure indicators on the scorecard should be linked to financial objectives. This procedure implies that

strategy is translated into a set of hypotheses about cause and effect. The assumption that there is a cause-and-effect relationship is essential because it allows the measurement in non-financial areas to be used to predict future financial performance. The validity of the model relies, however, on the assumption that the cause-and-effect relationship exists and that it can be adequately measured.

Non-financial measures may increase the efficiency in contracting, given managerial knowledge regarding the measures' informativeness about dimensions of managerial effort desired by the owner. Consistent with theoretical work on performance evaluation using multiple signals in agency settings (Feltham and Xie, 1994), financial measures alone may not provide the most efficient means to motivate managers to act in the manner desired by the firm's owners. In theory, the bonus contract should place a non-zero weight on any performance measure that provides incremental information content about the dimensions of managerial action that the owner wished to motivate. Ittner et al. (1997) proposed that as long as measures other than stock price convey information about desired managerial actions, they should be included in the bonus contract to efficiently motivate managers. However, they do not mention the impact of BSC on organizational performance.

Banker et al. (2000) revealed that non-financial measures of customer satisfaction were significantly associated with future financial performance and contain additional information not reflected in the past financial measures. Furthermore, both non-financial and financial performance improved following the implementation of an incentive plan

that includes non-financial measures. Hoque and James (2000) found that overall usage of BSC was significantly correlated with organizational performance, but this relationship does not depend on organization size, product life cycle, or market position. Yet, Ittner and Larcker (2003) found that managers made little attempt to link non-financial measures to advance their chosen strategies. Moreover, only 23% of these managers were able to show that they built causal models and most did not validate the causal links.

In a study of banks, Davis and Albright (2004) found that a group of branches that used BSC outperformed a group that did not use BSC on common composite financial measures. Other studies have shown that the association between strategic performance management systems and performance depends on the type of organizational performance being considered, with some evidence suggesting that strategic performance management systems are associated with medium-to-long-term performance. Ittner et al. (2003b) found that financial service firms using a broad set of financial and non-financial measures, relative to firms with similar strategies or value drivers, earned higher stock returns. Chenhall (2005) supported the view that the three dimensions of integrative strategic performance management system can enhance strategic competitiveness for firms emphasising both product differentiation and low cost-price strategies. These results are consistent with arguments that performance measurement can be a strategic management tool.

Market reaction to BSC use in executive compensation plans in the form of change in stock price may be considered as a key performance measure. Given the stock market is forward looking and anticipates changes in future management actions, these expectations should be impounded in the period surrounding the adoption of a BSC-based compensation plan. As Murphy (1985) pointed out, a corporate board of directors, unable to observe managerial effort directly, would base an executive's compensation on stock performance in addition to other variables that provide incremental information about the true productivity of the executive. Thereby, the market will react favourably to the adoption of BSC-based compensation plans, as evidenced by increased stock price, relative to firms without adopting BSC.

The notions of efficiency and effectiveness are at the heart of performance measurement, as they allow the “delineation of managerial responsibilities and accountability for the achievement of interim and ultimate organizational objectives at different stages of operations” (Pollanen, 2005, p. 6). A responsibility center is efficient if it does things right, and it is effective if it does the right things. The more this output contributes to the objectives, the more effective the unit. Therefore, efficiency is concerned with achieving given results with minimum resources, and effectiveness with attaining organizational objectives (Hansen et al., 2004). Because the corporate objectives aim to maximize shareholders’ value, for example, return on equity, net income, dividend and stock price are good indicators of effectiveness measures. In contrast, revenue per employee, accounts receivable turnover, net assets turnover and current ratio are examples of efficiency measures. However, a trade-off exists between efficiency and effectiveness,

making it necessary to balance the two types of measures carefully. Yet, efficiency and effectiveness are not necessarily mutually exclusive; every responsibility center ought to be both efficient and effective – in which case, the organization ought to be meeting its goals in an optimum manner.

These arguments lead to the second primary hypothesis that the usage of BSC in executive compensation will enhance organizational performance in terms of efficiency and effectiveness. Each element of this hypothesis is directed at management actions predicted to be valued by shareholders.

***H2A:** Organizational effectiveness is positively associated with the usage of BSC in executive compensation plans.*

***H2B:** Organizational efficiency is positively associated with the usage of BSC in executive compensation plans.*

This second set of hypotheses reveals whether there are more positive relationships between the usage of BSC in executive compensation plans and corporate performance, relative to firms that do not adopt BSC. It is also the first time that a dichotomy of effectiveness and efficiency has been used to gauge organizational performance in BSC and executive compensation studies. The two sets of variables also allow the exploration of possible “fit” of contextual factors and BSC usage in executive compensation and their relationships with corporate performance. In the next section, specific fit hypotheses will be proposed.

4. Hypotheses—Contingent Fit and Corporate Performance

This study posits that corporate performance is a joint function of the usage of BSC, firm-specific economic attributes and external events. The previous hypotheses assume that performance measures are efficacious for all firms. The match or fit between performance measures as a management practice and the firm's organizational environment is ignored when examining their performance consequences. Contextual factors, environmental factors, and strategic plans vary across firms. Contingency theories suggest that the choice of appropriate techniques of performance measurement depends on circumstances surrounding the firm. Therefore, Said, HassabElnaby and Wier (2003) argued that the adoption and use of performance measures is an endogenous choice, with the potential net benefit depending on contextual factors. Since the performance consequences of measures may be contingent on exogenous variables, the ability to draw inferences about the performance consequences of using those measures might be affected by specification errors. Similarly, Ittner and Larcker (1998) proposed that the optimal choice of performance measures was a function of a variety of factors such as strategic plans, organizational environment, investment opportunity, and executive compensation.

Although many studies contend that the adoption of BSC in compensation contracts is contingent on a variety of factors, little evidence exists on the relation between the match of BSC bonus plan measures with firm characteristics and the performance of firms using

BSC. The endogenous nature of the adoption of BSC measures raises a potential problem in assessing the performance consequences of BSC measures. Said et al. (2003) criticized that ignoring the possible endogeneity between the use of BSC measures and firm performance may result in inappropriate inferences regarding the usefulness of BSC measures. Firms that adopt BSC measures and fail to improve performance may incorrectly estimate the importance of BSC measures. These firms may not enhance their performance because they fail to match their characteristics and the use of BSC measures. Conversely, firms that adopt BSC measures and achieve higher performance are believed to have achieved the proper fit between their characteristics and the use of BSC measures. Previous studies have found a positive association between fit and performance (Chenhall and Langfield-Smith, 1998; Jermias and Gani, 2005).

Therefore, this study examines the performance differences associated with the degree of fit between the use of BSC measures and the contextual factors that reflect the firm's characteristics. This study expects positive performance consequences to be more pronounced in firms with a higher level of fit between BSC measures and their firm characteristics.

***H3A:** Firms that match BSC-based performance measures with their firm characteristics tend to achieve higher effectiveness.*

***H3B:** Firms that match BSC-based performance measures with their firm characteristics tend to achieve higher efficiency.*

5. Summary

BSC incentive reward schemes are likely to be one of the main mechanisms through which an organization can persuade managers to achieve corporate objectives. The primary questions concerning senior management are: Do contingent variables affect the usage of BSC in compensation plans? Does BSC usage in compensation plans have an impact on organizational performance? This study involves a search for systemic relationships between BSC usage in managerial remunerations and contextual variables, such as strategy, industry, size, structure, quality, culture and ownership. Empirical evidence will be provided to test performance impacts of BSC, in terms of effectiveness and efficiency. The association between contingency-fit and firm performance will be also investigated. To sum up, this section provides a conceptual model and associated hypotheses that are central to this thesis.

IV. Research Design and Methods

This chapter discusses research design, data collection, measurement of variables and data preparation and reliability check in the study.

1. Research Design

Sound research design is essential for valid research. A good study should ensure readers a clear understanding of the underlying concepts and research quality. There is no universal standard against which a research methodology can be judged, rather it must be evaluated in terms of its ability to produce reliable and valid research results. Ideally, a multi-method approach is highly desirable, but not often feasible.

1.1 Quantification vs. Qualification

In almost every management research project, there is value in consciously combining both qualitative and quantitative methods in what is referred to as a "mixed methods" approach. Because the natural sciences have had such resounding success with quantitative methods, qualitative methods are sometimes looked down upon as less scientific. Like quantitative approaches, qualitative research fits nicely with some research questions and contexts. Lee, Mitchell and Sablynski (1999) and Morgan and Smircich (1980) pointed out that tensions between qualitative and traditional research derive from differing philosophies of science. At one extreme, an objectivist view of the

social world holds a strict “natural science model.” It assumes an empirical world with a single objective reality, and the research process should be to understand that world via valid scientific laws. At the other extreme, the highly subjectivist view assumes that there are as many “realities” as there are people. In the world of work, then, organizational members must engage in the constant social construction of reality, the research process must strive to understand the realities of each and every person, and ideas about valid scientific laws are misnomers. As we pass from assumption to assumption along the subjective-objective continuum, the nature of what constitutes adequate knowledge changes.

Both approaches are contingent on the nature of the phenomena to be studied. Quantitative study is well suited for issues of prevalence, generalizability and calibration, whereas qualitative research is good for the purposes of description, interpretation, and explanation (Lee et al., 1999). Qualitative methods emphasize a broad class of empirical procedures designed to describe and interpret the experiences of research participants in a context-specific setting and are useful in exploratory research. Theoretical arguments, narratives and case studies often fall under this category. Quantitative methods usually incorporate large-scale sampling and use statistical methods. They stress the measurement and testing of causal or correlational hypothesized relationships between variables. In this study, the methods that best fits the theoretical questions and analytical situations are used. The quantitative approach is selected because this research, using a sample of publicly available, quantitative data, tests the relationships between BSC usage

in compensation plans and firm characteristics and their association with firm performance.

1.2 Deduction vs. Induction

Scientific inquiry often involves two broad methods of reasoning—deductive and inductive analyses (Trochim, 2006). Deduction is sometimes called a "top-down" approach that begins with a theory, then narrows down into more specific hypotheses, observations collection, and hypotheses testing—a confirmation of the original theories. Inductive reasoning, a "bottom up" approach, works the other way, moving from specific observations to broader generalizations and theories. In theory, deductive and inductive analyses can be considered as a spectrum from confirmatory to exploratory methods (Bordens and Abbott, 1996). Inductive reasoning is more open-ended and exploratory, especially at the early stage of research in an area. It can be thought of as a technique for data reduction. Deductive reasoning is narrower in nature and is concerned with testing or confirming hypotheses corresponding to prior theoretical notions.

In practice, however, what gets done is often somewhere in between the two ends of the spectrum. Even though a particular study may look like it is purely deductive, most social research involves both inductive and deductive reasoning processes at some time in the project. By using descriptive statistics, the inductive approach analyzes data for the purpose of formulating hypotheses worth testing and complements the tools of conventional statistics for testing hypotheses. By relying on inferential statistics,

deductive analysis is used for testing hypotheses, using statistics, such as p-values, confidence intervals, etc. (McClave, Benson and Sincich, 2001). This thesis primarily employs deductive reasoning, supplemented with inductive pre- and post-analysis and interpretation.

1.3 Reliability and Validity

Quality assurance should be one of the researchers' key objectives. Among the dimensions of research quality, reliability and validity occupy important places (Statistics Canada, 2000). Reliability is a matter of whether a particular technique, applied repeatedly to the same object, would yield the same result each time (Babbie, 1992). In its everyday sense, it is the "consistency" or "repeatability" of the measures. Measurement error plays a key role in degrading reliability. One cannot calculate reliability—only estimate it. Several important steps have been taken to increase reliability of this study:

- *First, the firms that have adopted BSC are matched with a group of control firms that use other performance measures in executive compensation.* This paired design removes important sources of variation, because the two groups are similar on several characteristics. A dichotomy of studying treatment and matched control groups has gained popularity in finance and accounting research. Barber and Lyon (1996) noted that the choice of an appropriate control sample was critical to investigating the operating performance following corporate events, such as the adoption of BSC. To assess whether a firm is performing unusually

well or poorly, one must specify expected performance in the absence of an event, thus providing a benchmark against which sample firms can be compared. A researcher should also control for various factors, unrelated to an event, that affect operating performance. Wallace (1997) studied the actions of a sample of 40 firms that have adopted compensation plans based on residual income, and compared these actions to a matched-pairs control sample of firms where incentive compensation continued to be based on traditional accounting earnings. He identified treatment firms that have used EVA through Lexis over a ten-year period. In order to lessen validity threats, he selected a matched-pair control group of firms that used earnings-based measures in their compensation plans. The firms were matched on four-digit Standard Industry Codes (SIC). Similarly, Said et al. (2003) matched their control firms using two-digit SIC code. Both Wallace and Said et al. then chose the closest control firms based on total assets in the year prior to the treatment firm's adoption of residual income-based performance plans. Barber and Lyon (1996) found that, while matching based on industry and size worked well in certain circumstances, matching based on pre-event performance proved a superior methodology.

- *Second, appropriate techniques are employed to ensure this study meets statistical assumptions.* Linear regression, for example, generally assumes normality, homoscedasticity, non-multicollinearity and exogeneity (McClave et al., 2001). Using both parametric and nonparametric tests, this study has paid particular attention not violating these assumptions. The database will be discussed later in Section 4.

It is important to integrate the idea of reliability with the other major criteria for the quality of measurement—validity—and develop an understanding of the relationships between reliability and validity in measurement. Validity refers to the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration (Babbie, 1992). It is often related to generalizability and the approximate truth of propositions, inferences or conclusions. There are numerous types of validity, such as construct validity, translation validity, criterion-related validity, etc. The type of validity the researcher needs to show depends on the nature of specific research questions (Cone and Foster, 1993). Several important steps have been taken to increase validity of this study:

- *First, the data are carefully prepared, screened and verified to ensure they are free of errors and potential bias due to missing data.* These procedures will be discussed in detail later in Section 4.
- *Second, multiple indicators are carefully selected to measure the underlying constructs.* Sound variable measurement, which is discussed in Section 3, is critical for data validity.
- *Finally, sample size and type are optimized to increase its representativeness of the population.* How a sample is selected from a population is of vital importance in statistical inference, because the probability of an observed sample will be used to infer the characteristics of the population (McClave et al., 2001). The observations used in this study are a good representative sample, which should exhibit characteristics typical of those possessed by the population of interest,

because all firms that reported their adoption of BSC in proxy filings during the research period are included in the sample. As the firms are required by law to make public their compensation plans, the chance of missing some BSC firms is low. Compared to similar studies, this project has a relatively large sample. If the researcher knows in advance the desired level of significance, the acceptable width of confidence intervals, and the approximate variability of data, it is possible to compute the size of the sample that will be desired (Diekhoff, 1992). Using the values of 5%, 0.5 and 3, respectively, which are common in practice, a sample size of 138 would be desired. With 330 firms identified, this thesis has a sample large enough for valid analysis.

Reliability and validity often impinge on each other. These qualities might conflict, requiring a trade off between various degrees of reliability and validity. The best solution is to use several different measures and tests, tapping the different aspects of these concepts (Diekhoff, 1992).

1.4 Levels of Measurement

The level of measurement refers to the relationships among the values that are assigned to the attributes for a variable. Knowing the level of measurement helps to decide how to interpret the data from that variable and what statistical analysis is appropriate on the values that were assigned. There are typically four levels of measurement that are defined

(Diekhoff, 1992; Trochim, 2006; Heiman, 1992; Babbie, 1992). This section presents a general example for each of them:

1. In nominal measurement, the numerical values just "name" the attribute uniquely, which have only the characteristics of exhaustiveness and mutual exclusiveness. No ordering of the cases is implied. An example could be clusters or categories of "culture".
2. In ordinal measurement, the scores indicate logical rank order. Distances between attributes do not have any meaning, and the intervals between values are not interpretable. For example, "firm independence" could be measured from high to low, which is coded as A, B, C or D in this study.
3. In interval measurement, each score indicates an actual amount, and there is an equal unit of measurement separating each score. The distances between attributes do have meaning and are interpretable. Because of this, it makes sense to compute an average of an interval variable, whereas it does not make sense to do so for ordinal variables. This study has no interval measurement for its variables. An example would be asking customers to estimate their satisfaction in a survey study.
4. In ratio measurement, the scores measure an actual amount, using a distinct unit of measurement, i.e., a scale to measure weight, distance, time, money, etc., and for which an absolute zero is meaningful. Unlike interval measures, meaningful fractions, ratios and other statistics can be calculated, such as ROE in this study.

The majority of variables in this study belong to ratio or nominal levels of measurement. Appropriate analytical techniques are selected to meet the requirements of these scales. It is important to recognize that there is a hierarchy implied in the levels of measurement (Figure 4.1). At lower levels of measurement, assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level in the hierarchy, the current level includes all of the qualities of the one below it and adds something new (scaling up). In general, it is desirable to have a higher level of measurement (e.g., interval or ratio) rather than a lower one (nominal or ordinal).

Figure 4.1: Levels of measurement

Source: Trochim (2006)

2. Data Collection

Unlike most related empirical studies that rely on survey methods, data for this study are collected from public databases.

2.1 Sampling Design

One of the major barriers hampering BSC and performance measurement research is the lack of consistent data about what firms do internally. This is a longstanding excuse for the lack of progress in literature. Whereas financial accounting researchers can obtain data from financial statements, firm disclosures, and databases such as Compustat and CRSP, public information on management control and performance measurement practices is rarely available. Instead, researchers must conduct questionnaire surveys, personal interviews, and field observations that tend to vary from study to study (Abernethy and Stoelwinder, 1995; Ittner and Larcker, 2001; Jermias and Gani, 2005). However, the issue of self-assessment is often raised as a concern. Surveys suffer from well-known problems such as response rate, non-random and surveyor biases. Chenhall (2003), for example, suggested that a subordinate's self-assessment correlated with objective assessments and with a superior's subjective assessment. This may have introduced bias into the data, which could bring up potential concerns regarding generalizability, reliability and validity. Further, the mean response rate is only 32%, and is expected to decrease to 27% by 2010 and 4% by 2050, according to Cychota and Harrison (2006, p. 145). This is worrisome because response rates remain an important indicator of survey success and pose a methodological concern in research. When non-response is high, there is a risk that the data are biased. These limitations require researchers to be more careful in drawing inferences.

Thereby, Jermias and Gani (2005) urged that future research uses publicly available data to measure companies' performance in areas such as ROI, net income, or cash flow from operations, as well as to measure the contextual variables, such as R&D expenses and market share to enhance the reliability of the constructs. Replacing self-reported organizational measures with actual results can enhance the validity of results, increasing the reputation of survey-based research. Chenhall (2003) also supported the use of public data. In response to this call, public data is used in this thesis.

This study limits treatment firms to only those that can be identified as having incentive compensation plans explicitly based on BSC. Although some firms have claimed to experiment with BSC for some internal decision making, they have not yet explicitly included performance measures based on BSC in their incentive compensation. Consequently, they are excluded from consideration. The sample of firms consists of all firms that have adopted BSC in executives' compensation plans (also called annual incentives) from January 1992 (when BSC was first published by Kaplan and Norton) to December 2009 (the data collection date). The initial company list was generated by searching company filings (EDGAR Online and SEC Online Database in particular) contained in Lexis-Nexis using key words "'balanced scorecard' and 'compensation or incentive or bonus'". This generated several thousand records. The existence of a BSC-based compensation plan was also confirmed by reading the firms' proxy statements (Appendix 1). The records that encompass multiple filings by the same company and the mutual fund family were excluded. After looking into the remaining reports, 185 companies that explicitly employed BSC in compensation plans were identified from

their proxy reports (Appendix 2). Systematic missing data (to be discussed in Section 4.2) reduced the sample to 165, accounting for 91.2% of total firms. Henceforth, this group of firms is referred as Treatment Group. Lexis-Nexis has become increasingly important in management accounting research. In their widely cited study of performance measures in annual bonus contracts, for example, Ittner et al. (1997) generated their sample from the 1993 and 1994 proxy text files.

Following the matched-pair research design, a matching sample of firms that do not use BSC performance measure was selected, based on same industry and pre-event size. The procedure used consists of choosing the control firm within the same 4-digit SIC code with the closest match on total assets for the year prior to the treatment firm's adoption of BSC-based performance measures. The process was operated through Compustat. Many companies' proxy or annual reports provide a list of peer companies for determining executive performance and compensation (Appendix 1). The database BVDEP-Osiris also provides major competitors for each company. If no peer companies in the same 4-digit SIC industry were identified, this study then moved to a higher level with 3-digit SIC code. In order to achieve the best match, the proxy filings of control firms, which were selected from the same industries as treatment firms with similar total assets, were further checked to ensure that they indeed used at least some other performance measurement models instead of BSC. Using this step-by-step strategy, 148 control firms were matched at the 4-digit SIC industry, while only 17 control firms were chosen with the 3-digit SIC code.

2.2 Sample Characteristics

This study is able to identify the year of the BSC adoption from the proxy statements. Table 4.1 and Figure 4.2 shows firms are increasingly implementing BSC to measure executives' performance. Since the data collection ended in April 2009 for statistical analysis, it is possible that some firms that used BSC had not filed their proxy reports to the SEC, resulting in the low number of adoption for 2009. An alternative explanation is that the widespread financial crisis and economic recession forced firms to re-consider their incentive programs, causing the disruption in BSC adoptions.

Table 4.1: The year of BSC adoption

<u>Year</u>	<u>Frequency</u>	<u>Percent</u>	<u>Cum.</u>
1992	2	1.21	1.21
1994	2	1.21	2.42
1995	1	0.61	3.03
1996	5	3.03	6.06
1997	6	3.64	9.70
1998	4	2.42	12.12
1999	7	4.24	16.36
2000	7	4.24	20.61
2001	9	5.45	26.06
2002	9	5.45	31.52
2003	16	9.70	41.21
2004	13	7.88	49.09
2005	17	10.3	59.39
2006	13	7.88	67.27
2007	19	11.52	78.79
2008	31	18.79	97.58
2009	4	2.42	100
Total	165		100

Figure 4.2: The distribution of BSC adoption by year

The companies were matched on industry and size, assuming that operating performance varies by industry and size (Fama and French, 1996). Although some matched-pairs slightly differ in size, overall the treatment and control firms show no significant size difference in the year prior to adoption, based on total assets (p -value=0.7983) in the two-sample two tailed t -tests reported below (Table 4.2). A complete list of firms in both treatment and control groups is shown in Appendix 2, which also includes their operating industry SIC number and the year that treatment firms first added BSC into their incentive compensation plans.

Table 4.2: Two-sample t-test with equal variances (total assets)

	Group	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Total assets(t-1) (\$mi.)	Treatment	50063	16381	211694	17721	82406
	Control	44704	13065	168839	18909	70499
Ho: mean (treatment group) – mean (control group) = diff = 0 Ha: diff $\neq$ 0 Degrees of freedom: 328						
				T = -0.2558 p-value = 0.7983		

Although much effort is made to match firms as closely as possible, treatment firms are different by definition in that they have self-selected to adopt BSC-based compensation plans. The factors that lead treatment firms to choose BSC may not be present in control firms, e.g., control firms may have found other means of mitigating potential agency costs. All research is plagued by the presence of confounding variables. The problem will be minimized by various kinds of controls or be estimated and taken into account by randomization processes. In Section 3 of this chapter, possible control variables are hence identified.

Having nested down the company list of both treatment and control groups, the data of variables were collected from Standard & Poor's Compustat, which covers approximately 98% of the world's market capitalization with data for 90,000 global securities. Missing data were supplemented using Bureau van Dijk's global database Osiris, which combines information from around 100 sources and covers 65 million companies.

2.3 The Sub-sample of Firms with Non-financial Weights

One potential source of error in measuring the use of BSC is the use of a dummy variable, which does not fully capture the actual reliance on, and relative importance of, non-financial measures. Not all firms in the treatment sample use identical measures of BSC in evaluating performance. Among these firms, weights assigned to non-financial/financial measures may vary significantly across firms. In addition, a dichotomous variable may pose a statistical problem, which will be discussed in detail in

the next section. Therefore, in this study, additional tests were conducted to check the robustness of the results for possible specification errors. Estimating the relative weights placed to non-financial measures may provide a better understanding of the degree of the usage of BSC on firm performance. Thus, the BSC measure was also replaced with the relative weights assigned to the non-financial measures as a robustness check on the main results, for the firms for which this information is available.

From the compensation committee report included in the proxy statements for BSC firms, this study was able to identify a sub-sample of 58 firms that provided the explicit weights placed on non-financial measures of BSC. The non-financial weights were calculated by 1) either summing the weights of non-financial BSC components, such as customers, internal process, learning and growth, etc., or 2) subtracting the weights of financial components from 100 per cent. These firms, which are identified with asterisk (*) in Appendix 2, are increasingly incorporating BSC into their executives' compensation plans, as showed in Figure 4.3.

Figure 4.3: The distribution of BSC adoption by year (sub-sample)

For this sub-group, using relative weights of non-financial components as a proxy of the extent of BSC usage, this study also investigates the relationships 1) between the extent of using BSC and contingency variables, 2) between firm performance and the extent of using BSC, and 3) between firm performance and contingent fit. Consistent with Ittner et al. (1997) and Said et al. (2003), this study predicts a positive relation between non-financial weights of BSC usage and its contingent factors. Yet, there is no prior research on the association between non-financial weights of a measurement scheme and their performance consequences. It is logical to assume that such an association is weak across sections, because a positive association would lead to 100% use of non-financial measures, while a negative association would result in 100% use of only financial measures.

3. Measurement of Variables

3.1 BSC Usage in Executive Compensation Plans

The usage of BSC is measured by a dummy variable, which takes on value of 1 for firms that use BSC (treatment group) and 0 for matched firms that do not use BSC (control group). This allows for testing of potential systematic differences in these two groups on the variables of interest.

Binary dependent data are a common feature in many areas of empirical economics, finance and accounting. There is a long history of binary choice models (Matyas and

Sevestre, 2008). Dummy variables can be useful because they enable us to analyze differences between multiple groups in a single regression analysis. This means that the researchers do not need to write out separate equation models for each subgroup. The dummy variables act like “switches” that turn various parameters on and off in an equation. Another advantage of a 0,1 dummy-coded variable is that, even though it is a nominal-level variable with two categories, we can treat it statistically like an interval-level variable. For instance, if we take an average of a 0,1 variable, the result is the proportion of 1s in the distribution.

The researchers can deal with dummy variables as independent or dependent variables (Baltagi, 2008). The use of dummy variables usually increases model fit (coefficient of determination), but at a cost of fewer degrees of freedom and loss of generality of the model. Logistic regression is the approach to prediction for the dichotomous dependent variable of BSC usage and the researcher is predicting a dichotomous outcome. Yet, it is often possible that the variable is multinomial. This situation poses problems for the assumptions of Ordinary Least Squares (OLS) regression that the error variances (residuals) are normally distributed. When using the logistic distribution, the researchers usually need to make a conversion of the dichotomous relationship into a linear regression equation (Agresti, 2002). One other difference between OLS and logistic regression is that there is no R^2 and F-test to gauge the model integrity. Instead, pseudo- R^2 and χ^2 test are used to indicate how well the logistic regression model fits the data.

For the sub-sample, the binary BSC variable was replaced by the relative weights placed on non-financial measures, as explained in Section 2.3 of this chapter.

3.2 Contingent Variables

Table 4.3: Measurement of contingent variables

<u>Variables</u>	<u>Predicted sign</u>	<u>Definitions</u>	<u>Data source</u>
Strategy			
R&D ratio	+	R&D expense divided by total sales	Compustat
MV/BV	+	Market-to-book value ratio	Compustat/Osiris
Labour cost	+	Total labour costs divided by total sales	Compustat
Industry			
2-digit SIC code	+/-	The Standard Industry Classification	Compustat
Size			
Market cap.	+	Market value of the firm	Compustat/Osiris
Total assets	+	Total assets from balance sheet	Compustat
Total revenues	+	Total sales from income statement	Compustat
Total employees	+/-	Total employees from financial reports	Compustat
Quality			
Quality	+	1 for winning a major quality award, 0 otherwise	Quality institutes' websites/Lexis-Nexis
Structure			
# segments	+	The number of business segments	Osiris
Foreign operations	+	Coded 1 if a firm reports foreign operations, and 0 otherwise	Osiris
Culture			
Culture	+	Anglo cultural clusters (See Table 4.5)	Compustat/Osiris
Ownership			
Inst. ownership	+	% shares held by institutional holders divided by the number of institutions that own the stock	Compustat
Independence	-	A,B,C,D	Osiris

Strategy. Table 4.3 presents the measurement of contingent variables. Strategy orientation is plotted on a spectrum from cost leader (defender) to differentiator

(prospector). Higher ratios represent firms that are closer to the right end of the strategy continuum. Following Ittner et al. (1997), the variable “strategy” is measured with two indicators: 1) the ratio of R&D to sales, which provides a measure of a firm’s propensity to search for new products. Since prospectors engage in a greater amount of innovative activity, they are expected to carry out more research and development than defenders; and 2) the market-to-book ratio, which is a proxy for a firm’s growth or investment opportunities, with prospectors expected to have greater growth potential than defenders. In addition, this study uses the ratio of labour costs to sales to reflect the firms’ orientation to hire more skilled workers to produce innovative, differentiated goods and services. Prospector firms should have higher score for all three ratios. These indicators are the respective annual ratios at the year of BSC adoption.

Industry. The variable “industry” is classified by 2-digit SIC (Standard Industry Classification) codes (Table 4.4).

Table 4.4: Industry classification by SIC

01-09	agri_forestry_fishing
10-19	mining_construct
20-29	manufacturing_nondurables
30-39	manufacturing_durables
40-49	utility_transport
50-59	wholesale_retail
60-69	financials
70-89	services

Size. The variable “size” is measured using four metrics —market capitalization, total assets, total revenues, and the number of employees (Table 4.3). While accounting researchers tend to use total assets (HassabElnaby et al., 2005), financial economists

prefer market capitalization or equity (Fama and French, 1996). As each side has its own arguments, this study uses both of them.

Quality. The variable “quality” is binary, coded as 1 if the company has won a major quality award, and zero otherwise (Table 4.3). The intuition behind this measure is that the quality award criteria require the firm to demonstrate how quality programs fit into the firm’s overall business strategy. Major quality awards are U.S. Malcolm Baldrige National Quality Award, European EFQM Excellence Award, Japan’s Deming Application Prize, Canada Quality Awards, and ISO 9000 Certification. The profiles of the activities and accomplishments for each award recipient are usually posted on these quality organizations’ websites at the time when the awards were presented. After the firms that won certain awards have been identified from these websites, this project also double-checked that event through an extensive keyword search of publications in Lexis/Nexis. Canada Quality Award and ISO 9000 Certification are added as additional awarding agencies to minimize the omission of some important recipients, compared to those in Ittner et al. (1997) and Said et al. (2003).

Structure. This study posits that firms with greater complexity and scope of operations are more likely to adopt BSC. The complexity of firm operations, and consequently, the intricacy of its transactions increase as a firm operates in international markets. The more complicated the firm’s transactions, the more urgent the need to retain adequate internal controls. This study uses “# segments” (defined as the number of reported business segments) and “foreign operations” as proxies for structure (Table 4.3). The variable

“foreign operations” is coded as 1 if a firm forms foreign subsidiaries and 0 otherwise.

Both variables are identified using Bureau van Dijk’s Osiris database.

Table 4.5: Global society/culture clusters

<u>Anglo (1)</u>	<u>Latin Europe (2)</u>	<u>Nordic Europe (3)</u>	<u>Germanic Europe (4)</u>
Australia	France	Denmark	Austria
Canada	Israel	Finland	Germany
England	Italy	Sweden	Netherlands
Ireland	Portugal		Switzerland (German speaking)
New Zealand	Spain		
South Africa (White sample)	Switzerland (French-speaking)		
United States			
<u>Eastern Europe (5)</u>	<u>Latin America (6)</u>	<u>Sub-Saharan Africa (7)</u>	<u>Middle East/Arab (8)</u>
Albania	Argentina	Namibia	Egypt
Georgia	Bolivia	Nigeria	Kuwait
Greece	Brazil	South Africa (Black sample)	Morocco
Hungary	Colombia	Zambia	Qatar
Kazakhstan	Costa Rica	Zimbabwe	Turkey
Poland	Ecuador		
Russia	El Salvador		
Slovenia	Guatemala		
	Mexico		
	Venezuela		
<u>Southern Asia (9)</u>	<u>Confucian Asia (10)</u>		
India	China		
Indonesia	Hong Kong		
Iran	Japan		
Malaysia	Singapore		
Philippines	South Korea		
Thailand	Taiwan		

Adapted from Chhokar et al. (2008, p. 13 & 1027)

Culture. The variable “culture” is classified using cultural clusters, where each of the ten groups assigned a nominal value from 1 to 10 (Table 4.5). One line of research has attempted to establish clusters of countries based on their relative similarity (Ronen and Shenkar, 1985). The Global Leadership and Organizational Behaviour Effectiveness Research Project led by Chhokar et al. (2008, p. 13) studied culture topics afresh and in a

simultaneous breadth and depth that has never before been attempted. With the participation of 17,300 middle managers from 950 organizations, the scale of this project rivals a large multinational corporation. This ground-breaking project involved 170 co-investigators from 61 countries. On the basis of the data on cultural practices and values from these countries, the project identified ten cultural clusters around the world that correspond highly with previously published attempts to identify distinguishable cultural regions. Countries from the same cluster share characteristics such as geographic proximity and climate zone, mass migration and ethnic social capital, and religious and linguistic roots. The ten clusters serve as a useful framework for measuring a firm's "culture".

Ownership. Regarding the variable "ownership", this study uses the percentage of shares held by institutions divided by the number of institutions that own a firm's stock, as a measure of concentrated institutional ownership (Table 4.3). This thesis assumes that institutional owners' influence on the usage of BSC is an increasing function of concentrated institutional ownership. This study also uses independence indicators for a firm to characterise the degree of independence of a company with regard to its shareholders. The independence indicator, which exhibits an inverse relationship to the usage of BSC, is assigned to each company by Bureau van Dijk's database Osiris, according to the logic defined below:

- A: Attached to any company with known recorded shareholders, none of which have more than 25% of direct or total ownership;

- B: Attached to any company with known recorded shareholders, none of which have an ownership percentage (direct, total or calculated total) over 50%, but have one or more shareholders with an ownership percentage above 25%;
- C: Attached to any company with recorded shareholders with a total or calculated total ownership over 50%;
- D: Attached to any company with a recorded shareholder with a direct ownership of over 50%.

3.3 Performance Variables

Organizational performance is measured through two key dimensions: effectiveness and efficiency. They lie at the heart of the task of controlling an organization. Efficiency is the relationship of output to given sets of inputs. Efficiency gains will occur either if the value of the outputs rises per unit of input, or if the cost of the inputs falls per unit of output (Berry et al., 2005). Effectiveness is defined as a measure of the achievement of the organization's purpose. They can be expressed in several ways. For example, Wallace (1997) measured organizational performance using asset dispositions, new investments, share repurchase, dividends, total asset turnover, inventory turnover, accounts receivable turnover, and accounts payable turnover, and found that firms adopting residual income-based incentive program utilized existing assets more intensively. In order to mitigate likely measurement error, multiple indicators are used in this thesis to measure effectiveness and efficiency (Table 4.6). The time frame for firm performance is one year. BSC is designed to evaluate managers' both short-term and long-term performance.

A performance evaluation period longer than one year could be exposed to too much “noise” other than the effects of BSC. And a short-window less than one year could be also problematic, given that managers usually take longer time to improve efficiency, such as productivity, growth and asset utilization.

Table 4.6: Measurement of performance variables

<u>Variables</u>	<u>Predicted sign</u>	<u>Definitions and data source</u>	<u>Data source</u>
Effectiveness	+		
ROE	+	Net income divided by equity	Compustat
ROA	+	Net income divided by assets	Compustat
ROI	+	Net income divided by investments	Compustat
Gross margin	+	Gross profit / sales	Compustat
EBIT margin	+	EBIT / sales	Compustat
Net profit margin	+	Net income / sales	Compustat
Dividend payout	+	Dividend / EPS	Compustat
P/E ratio	+	Stock price / EPS	Compustat/Osiris
Price to CF	+	Stock price / cash flow from operations	Compustat/Osiris
Stock return	+	1-year total return	Compustat/Osiris
Dividend yield	+	Dividend / stock price	Compustat/Osiris
Efficiency	+		
Rev. per employee	+	Total revenue / total employees	Compustat
Labour cost	-	Labour costs per employee	Compustat
NI per employee	+	Net income / total employees	Compustat
Sales growth	+	1-year sales growth rate	Compustat
EPS growth	+	1-year earnings growth rate	Compustat
AR turnover	+	Sales / average AR	Compustat
Inventory turnover	+	COGS / average inventory	Compustat
Asset turnover	+	Sales / total assets	Compustat
Operating cycle	-	Days of operating cycle	Compustat
Quick ratio	+	(cash+market. securities)/current liabilities	Compustat
Interest coverage	+	EBIT / interest expense	Compustat
D/E ratio	-	Total long-term debt / total equity	Compustat

This study uses both accounting-based and market-based measures to capture effectiveness. Accounting measures have been extensively criticized as being misleading

or inadequate indicators of the economic rate of return. In particular, they may be distorted by the failure to consider differences in systematic risk, temporary disequilibrium effects, tax laws, and accounting procedures (Said et al., 2003). Accounting-based measures are subject to noise effects from generally accepted accounting principles (GAAP) or managers' manipulation and may not necessarily reflect real performance. For various reasons, GAAP measures may fail to reflect relevant values or managers may smooth income. Sloan (1996) found that firms with extreme accruals tend to revert to the mean more quickly than firms with moderate levels of accruals. If managers smooth income or manipulate earnings downward, over time this may result in no or insignificant changes in accounting returns from one year to another. This may explain the weak association between accounting performance and the weight of non-financial performance measures.

Stock prices are often argued to be superior measures of performance compared to accounting measures because accounting policies prevent accounting performance measures from reflecting the entire value-relevant information set that is impounded in stock prices (Bushman, Indjejikian and Smith, 1996). However, while market-based performance (i.e., stock price) indicators efficiently aggregate information about firm performance, it inefficiently aggregates information about managers' actions. Using these indicators only as evaluation criterion motivate managers to over-emphasize measurable activities fully impounded in price at the expense of other activities that take longer to affect stock prices (HassabElnaby et al., 2005).

Therefore, this project combines accounting-based with market-based measures to examine the effects of BSC measures on performance, using stock return, price-to-earnings ratio, price-to-cash flow ratio, and dividend yield. This approach reveals whether adoptions of BSC compensation plans are reflected in increased shareholder wealth. The market-adjusted return, also called abnormal return, is not a suitable indicator for this thesis. Using classic “event study”, it is typically calculated from daily stock returns adjusted by a benchmark index. Unlike “classic events”, such as mergers and acquisitions, which are announced timely by companies, it is difficult to determine exactly when the market first learned of the firm’s decision to adopt BSC, and therefore, when this information is impounded in the firm’s share price. For many firms, the first public information on BSC is probably the firm’s proxy statements. These statements are usually released between three and four months after the firm’s fiscal year end. Because of this uncertainty, annual return is a better measure of performance for BSC. Meanwhile, the matched-pair design is sufficient to assess return disparity between BSC firms and control firms by removing sources of variation, because the two groups are similar on size and industry.

Effectiveness, as manifested in indicators, such as return on equity, stock return and EPS growth, is expected to be higher for firms following adoption of a BSC-based compensation plan, relative to firms in which the manager is evaluated and compensated based on other measures. Similarly, one would expect higher efficiency for firms following the adoption of BSC-based performance measures, as compared to the control

group. There should be an increase in indicators such as revenue per employee, asset turnover and current ratio and a decrease in debt-to-equity ratio.

3.4 Control Variables

Corporate performance is measured with error. The incorporation of non-financial measures in the annual bonus contract is only one means for motivating executives to take actions that are in shareholders' best interests. Individual firms' performance results at the time of an event, such as the adoption of BSC, might be affected for reasons unrelated to the event, and this component of the abnormal return does not average to literally zero in the cross-section (Khotari and Warner, 2007). Further, although the effort was made to match firms as closely as possible, treatment firms are different by definition in that they have self-selected to adopt BSC-based compensation plans. The factors that lead to higher performance in treatment firms may not be present in control firms, e.g. control firms may have found means of enhancing performance other than BSC. As a result, in testing the performance consequences of using BSC, this study introduces five control variables – merger and acquisition, new Chief Executive Officer (CEO), financial distress, product cycles and economic events (Table 4.7).

Mergers and Acquisitions. Because an announcement of a merger or acquisition can have significant impact on the stock price behaviour of the companies involved, merger and acquisition (M&A) is used as a control variable. Also, acquiring firms face significant management control challenges when integrating their operations, systems, and cultures with those of acquired firms. Failure to develop adequate controls over accounting for

acquired assets can increase the risk for acquiring firms and reduce efficiency. Thus, M&A can be equally critical for accounting-based performance. Following Ashbaugh-Skaife et al. (2007), M&A is coded as 1 if the firm has been involved in a merger or acquisition at year t or $t-1$, and 0 otherwise. The intuition is that the transactions at year t can extend their effects into next year. This study predicts a positive association between M&A and firm performance. M&A are identified from Bureau van Dijk's Zephyr, a database specialized in M&A.

Table 4.7: Measurement of control variables

<u>Variables</u>	<u>Predicted sign</u>	<u>Definitions</u>	<u>Data source</u>
M&A	+	Coded 1 if a firm is involved in a merger or acquisition at t or $t-1$, and 0 otherwise	Zephyr/Osiris
New CEO	-	Coded 1 if a firm hires a new CEO at t or $t-1$, and 0 otherwise	Osiris/Lexis-Nexis
Financial health	+	Altman z-score measure of bankruptcy	Compustat
Product cycle			
Product develop cycle	-	Coded 1 if long, and 0 otherwise	U.S. National Academy of Engineering
Product life cycle	+	Coded 1 if long, and 0 otherwise	
Economic events			
Industry ROA	+	ROA of the S&P industry groups	Compustat
Industry volatility	-	Standard deviations of ROI of the S&P industry groups	Compustat

New CEO. Similarly, a firm's performance may also be a function of the influence of the CEO. This variable is coded as 1 if the firm has hired a new CEO at year t or $t-1$, and 0 otherwise. The intuition is that organizational changes initiated by new CEOs at year t may also influence firm performance next year. It is likely that changes in the executive position are a sign that the compensation agreements are in a state of flux (HassabElnaby

et al., 2005). Meanwhile, in the case of the retiring CEO, agreements on compensation that were applicable to the predecessor executive may not be appropriate for his/her successor. As a new CEO signals a significant interruption to the continuity of management and operations, a negative relationship to firm performance is expected.

Financial health. This study also posits that firms facing bankruptcy risk are more likely to under-perform, compared to peer companies. This study uses the Altman Z-score to capture distress risk, with higher z-scores indicating lower financial distress and higher firm performance. Z-score is a bankruptcy prediction model developed by Edward Altman at New York University in the 1960s. This model combines several financial ratios to determine the likelihood of bankruptcy. Generally speaking, the lower the score, the higher the odds of bankruptcy. If the value is less than 1.81, there is a high probability of bankruptcy. If the value is greater than 3.0, there is a low probability of bankruptcy. This is a relatively accurate model—real world application of the Z-score successfully predicts 72% of corporate bankruptcies two years prior to their filing for bankruptcies (www.wikipedia.org).

Product cycles. Previous studies have indicated that product cycles are potentially important determinants of firm performance. Following Said et al (2003), this thesis adopts the classification scheme from National Academy of Engineering (1992) to categorize firms into those with shorter versus longer time horizons. It uses the two-digit SIC code and the business description in the firm's 10-K reports and Compustat files to classify the companies according to National Academy of Engineering's classifications.

Both product development and life cycles are dummy variables, which take on the value of 1 if the cycle of a firm is long, and 0 otherwise (Table 4.8). This study predicts that firm performance is negatively associated with product development cycle (due to higher expenses) and positively with product life cycle (due to higher revenues).

Table 4.8: Product development and life cycles

	<u>Product development cycle</u>	<u>Product life cycle</u>
aircraft	long	long
autos	short	long
biotechnology	long	short
chemicals	long	long
cigarettes	short	long
communication systems	long	long
computers	short	short
fashion, apparel	short	short
financial services	short	short
foods	short	long
lodging, hotels	short	long
metal products	short	short
mining	short	long
oil and gas	long	long
paper	long	long
pharmaceuticals	long	long
publishing	short	long
retail, wholesale, entertainment	short	short
software	short	short
timber products	long	long
transportation systems	long	long

Source: Adapted from Said et al.(2003, pp.204-205)

Economic events. As industry events and general economic condition can have effects on firm performance, this thesis also includes industry ROA (return on assets) and industry volatility (the standard deviation of industry return on investment) as control variables. These two variables are used as proxies for the noise in the usage of BSC measures that is beyond the executive's control (Ittner et al., 1997). It is not clear how these two variables affect firm performance.

3.5 Measurement of Fit

Following HassabElnaby et al. (2005), this study adopts the fitted factor approach to measure contingent fit between BSC and firm characteristics. This study first runs the regressions on BSC usage and firm variables. The fit is captured in the predicted values, using the Equation: $BSC_dummy = f(Firm_characteristics)$, where BSC is a dummy variable and Firm_characteristics includes the organization's strategy, industry, size, quality, structure, culture and ownership. Prior research indicates that firms that better match executive incentive schemes with firm characteristics observe better performance (Said et al., 2003).

4. Data Preparation, Verification and Testing

Data preparation and verification involves checking or logging the data in; checking the data for accuracy; entering the data into the computer; handling missing values; transforming the data; and developing and documenting a database structure that integrates the various measures.

4.1 Database Construction

A database for this project was set up to record incoming data for all variables and some general information. This is a critical component in good research record-keeping. The

data were collected from multiple sources including Compustat and Osiris. The dataset enables the researcher to assess at any time what data are already in and what data are still required. It was accomplished by using standard statistical programs in Stata and Excel, which allowed the production of timely reports on data status. It is also critical to retain the original data records for a reasonable period of time in order to trace a result from a data analysis back to the original forms on which the data were collected. There are no research guidelines on how long the data should be retained. Rather, it depends on the nature of specific research and individual preference. The dataset for this project will be maintained to be available for future use. Also, unlike in survey design, confidentiality is not a concern, as this project uses only public information. However, multiple copies of the dataset are maintained in order to minimize data loss due to technological failure.

The database structure is the manner in which the data are stored for the subsequent analyses. It describes the data and indicates where and how it can be accessed. This document includes the items for each variable, including variable name, variable description, variable format (number, data, text), instrument/method of collection, date collected, variable location (database), and notes. In most research, quality of data is a major issue in order to ensure reliability. Assuring that the data collection process does not contribute inaccuracies helps assure the overall quality of subsequent analyses. In order to ensure data accuracy, records were further spot checked against original sources on a random basis. A good way to detect entry errors is to have the computer printout a copy of the data file and then look carefully at the pattern of numbers.

4.2 Missing Data

Once the data have been entered it is necessary to transform the raw data into variables that are usable in the analyses (Trochim, 2006). Yet, proper handling of missing values is important in all research. Improper treatment will distort analysis. One of the popular methods involves the deletion of cases with missing data. This is attractive because it is easy to implement. However, this strategy suffers serious drawbacks, because, until proven otherwise, the researcher must assume that missing cases differ in analytically important ways from cases where values are present. That is, the problem with the treatment of missing values is not so much the reduced sample size as it is the possibility that the remaining data set is biased (Little and Rubin, 1987). Another common method available in most statistical software is mean substitution, which replaces missing data with the average of valid data. Because the same value is being substituted for each missing case, however, this method artificially reduces the variance of the variable in question, in addition to diminishing relationships with other variables. Graham, Cumsille and Elek-Fisk (2003) referred to these traditional methods as “unacceptable methods.” Instead, missing values for any variable can be predicted using existing values from other variables. The predicted values are substituted for the missing values, resulting in a full data set. This exercise is called “imputation”, which has developed significantly since the 1980s. Wayman (2003) showed that imputation provided adequate results in the presence of low sample size or high rates of missing data.

Categorical data. For the categorical variables in this study, there are no missing values. All firms can be classified by industry (SIC codes) and a cultural cluster (headquarters)

from their proxy files. All firms can be identified whether or not they explicitly have the features of interest, including a quality award, a merger or acquisition, a new CEO, and a foreign operation. If they do, they are coded as 1 and 0 otherwise. Similarly, all firms can be assigned into a long or short product life and development cycle, based on their industry codes.

Continuous data. For the continuous variables, the missing values only accounted for less than 5% of the data. And they are minimized because, if a data point is missing for a variable in the public databases (Compustat and Osiris), it was verified from a firm's financial statements in SEC filing or company websites. The missing values were then imputed using either interpolation or extrapolation. The former is an estimation of a value within two known values in a sequence of values; and the latter is an estimation of a value based on extending a known sequence of values or facts beyond the area that is certainly known. The missing values filled using imputation ranged from 0.3% to 4.6%, depending on individual variables. Once all missing values were imputed, the dataset could then be analysed using standard techniques for complete data. Imputation restores the natural variability of the sample due to missing data. Yet, the downside is that imputation might incorporate uncertainty caused by estimating missing data. However, since the missing values in this study are random and merely comprise a small proportion of the sample, the imputed substitutions are unlikely to cause any large bias in the results. Wayman (2003) argued that imputation is an attractive choice as a solution to missing data problems, because it represents a good balance between the quality of results and the ease of use. Not long ago, missing data were viewed as something to discard. Researchers

knew the bias problems it presented, but there were no methods available to account for missing data bias. Today, researchers are not bound by such constraints. Methods such as imputation are available and usable for researchers, so there is no need to publish studies that suffer from sample bias.

Final sample size. In this project, 185 companies that explicitly employ BSC in compensation plans were initially identified from their proxy reports. Yet, 20 of these firms, which account for 10.8% of the total, had systematically missing values for continuous data that could not be obtained. These firms were excluded from the sample (see the notes in Appendix 2), bringing the treatment group to 165 companies. Together with the matched-pair control firms, the final sample of this project is thus 330. Missing data is not a problem for control firms because usually more than one matched firms could be identified. If one of the matched firms had a large number of missing data points, another one was used. For the sub-sample of firms with relative non-financial weights, 58 firms were identified from 165 treatment firms. As a result of this methodology, data analyses for all variables in the full sample (330 observations) and in the sub-sample (58 observations) have no “blank” or missing data points.

4.3 Normality Assessment

Many statistical methods are based on various underlying assumptions. One common assumption is that variables are random and normally distributed (Park, 2008). Normality is critical in ordinary least square regression (OLS), one statistical technique used in this

study. When this assumption is violated, interpretation and inference may not be reliable or valid. Logistic regression, the other main statistical technique used in this study, has many analogies to OLS. But unlike OLS, logistic regression does not require normally distributed variables, does not assume equal variances (homoscedasticity), does not assume linearity of relationship between the independent and dependent variables, and in general has less stringent requirements (Baltagi, 2008). It does, however, require that the independent variables be linearly related to the logit (log transformation of probability) of the dependent variable. Goodness-of-fit tests, such as the likelihood ratio and χ^2 tests, can be used as indicators of model appropriateness.

Continuous variables. Normality tests can be used to determine whether each of these variables is well-modeled by a normal distribution or not, or to compute how likely an underlying random variable is to be normally distributed. Most abnormalities can be detected first using descriptive statistics that show the mean tendency of a variable, including median and standard deviation. A standard deviation higher than mean value, for instance, indicates some extreme values. If the problems of the minimum and maximum values, for example, outliers, are corrected, the standard deviation will be smaller. It is important to know how to handle them appropriately, so that the data can be analyzed without losing their integrity or robustness. Using Stata, normality in this study has been tested by graphical and numerical methods. Graphical methods visualize the distributions of random variables or differences between an empirical distribution and a theoretical distribution (e.g., the standard normal distribution). Numerical methods present summary statistics such as skewness and kurtosis or conduct statistical tests of

normality. Graphical methods are intuitive and easy to interpret, while numerical methods provide objective ways of examining normality (Park, 2008).

As the first data screening technique, Kernel density estimation was employed in this study to inspect the normal distribution of continuous variables. It constructs a smooth probability density function of a variable, which in general can accurately reflect the underlying variable. Given some data about a sample, Kernel density estimation makes it possible to extrapolate the data to the entire population (Park, 2008). Since large values can cause skewed distributions, total assets, market capitalization and total sales (“size” variable) are first transformed to their logarithmic values, following Said et al. (2003) and HassabElnaby et al. (2005). This normalization stabilizes the sample variance (Bordens and Abbott, 1996).

Overall, the results show that the variables of this study have approximately normal distribution (Appendix 3). Even when a distribution may not be exactly normal, Baltagi (2008) and McClave et al. (2001) argued that, it may still be convenient to use normal distribution as a good approximation by relying on the Central Limit Theorem, if the sample size is large. This study contains 330 firms, a sample that is relatively larger than most of the similar studies in the area. Hence, many statistical procedures, such as the t-test can be used with reasonable assurance.

Categorical variables. Normal distribution is not a concern for this type of data. Histograms, plots of data density, are used to display their tabular frequencies (Appendix

4). However, the unequal group size for some variables could affect the significance of some statistical tests.

The detailed descriptive statistics for both continuous and categorical variables, as well as the specific requirements of each statistical test, are documented in Chapters V and VI.

4.4 Parametric vs. Non-parametric Statistics

Choosing between parametric and non-parametric tests can be a bit tricky. Many statistical tests are based upon the assumption that the data are sampled from a normal distribution. These tests are referred to as parametric tests. Tests that do not make assumptions about the population distribution are referred to as non-parametric tests. Motulsky (1995) recommended that a parametric test should be chosen if the data are sampled from a population that follows a normal distribution (at least approximately). The researcher should select a non-parametric test if 1) the scale for a variable is ordinal or nominal, 2) the distribution is clearly not normal, or 2) some values are “off the scale”, that is, too high or too low to measure (Bordens and Abbott, 1996). Yet, it is not always easy to decide whether a sample comes from a normally distributed population. Does it matter whether one chooses a parametric or nonparametric test? The answer depends on the sample size. The central limit theorem ensures that parametric tests work well with large samples even if the distribution is not normal. When there are at least 24 data points and the variables use interval or ratio scales, parametric and non-parametric tests are both powerful and robust (Motulsky, 1995).

This study has 330 observations, and the distributions of most variables are approximately normal. Yet, several indicators, such as dividend yield and operating cycle, appear to be slightly skewed, as shown by K-density estimations and descriptive statistics, and many independent and control variables are nominal. For these reasons, this study focuses on parametric tests (i.e., pair t-test and Pearson correlation), supplemented with non-parametric tests (i.e., Wilcoxon test and Spearman correlation). Each test used is introduced in the appropriate results sections.

5. Summary

Research design provides the glue that holds the research project together. It is used to structure the research, to show how all of the major parts of the research project—the samples, measures, treatments and methods of assignment—work together to address the central research questions. This study makes use of public databases. The dataset consists of all firms listed in the chosen databases that have adopted BSC in executives' compensation plans from 1992 to 2009. The treatment firms are matched with control firms that do not use BSC. This design allows for testing whether systematic differences exist between the two groups. In order to mitigate likely measurement errors, preliminary steps were taken to check data for reliability and validity and to prepare them for analysis, using both parametric and non-parametric tests. The empirical results are presented in the next two chapters.

V. Results: Full Sample of Firms Adopting BSC

The objectives of this chapter are threefold. First, it investigates the association between the usage of BSC and contingent factors (H1). Second, it investigates whether corporate performance is positively related to the inclusion of BSC in compensation contracts (H2). Third, it investigates the related relationships between the use of BSC measures and economic performance as a “fit” or “match” between a firm’s operational and competitive circumstances and its choice of measurement schemes (H3).

Covering the period of 1992–2009, a sample of firms that used BSC measures is compared to a matched sample of firms that based their executive compensation on non-BSC measures. Appropriate parametric and non-parametric tools, as discussed in the previous chapter, are selected to match the type and amount of data available for each test. Logistic regression is used for categorical dependent variables (H1) and OLS regression for continuous dependent variables (H2 and H3). All analyses are run in Stata.

Consistent with predictions, the findings indicate that the use of BSC is positively and significantly associated with innovation-oriented strategy, utility and financial industries, larger size, quality management, complexity of corporate structure, Anglo culture, and institutional ownership. Organizational performance is positively related to the adoption of BSC measures. As expected, firm performance is a function of the match of BSC with certain firm characteristics.

1. Model 1—BSC Usage and Firm Characteristics

Descriptive statistics are presented first to provide important information about the variables to be analyzed. Data distributions are then examined through univariate analyses, including two-sample difference tests, pair-wise correlations and multicollinearity. The hypotheses (H1A to H1G) are then tested using logistic regression.

1.1 Descriptive Statistics and Univariate Tests

Descriptive statistics. The results are summarized in Table 5.1. Mean and median measure central tendencies of a continuous variable, while standard deviation, skewness and kurtosis assess its dispersion (Park, 2008). For categorical variables, their frequency and distributions are examined as well. Regarding industry distribution, financials (27.9%) and utility (19.4%) account for almost half of the total firms. Approximately one third of the firms come from the durable-goods (16.4%) and nondurable-goods (13.9%) manufacturing industries. Services, mining/construction and wholesale/retail consist of 10.3%, 6.1% and 6.1% respectively. In terms of culture, the Anglo cluster is the dominant business culture in the sample (74.5%). In contrast, other clusters as a whole account for only 25.5%.

The data meet the needs of logistic regression for H1. Although certain indicators appear skewed, they cause no problem for statistical analysis. Logistic regression makes no

assumption about the distribution, which does not have to be normally distributed, linearly related or of equal variance within each group (Agresti, 2002; Diekhoff, 1992).

Table 5.1: Descriptive statistics (contingent variables)

		<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>Skewness</u>	<u>Kurtosis</u>
Strategy	R&D ratio	0.08	0	0.77	15.18	250
	MV/BV ratio	3.82	1.41	36.8	16.69	296
	labour costs/revenue	0.08	0	0.12	1.61	5.66
Size	total assets (\$mi.)	53156	3798	208482	7.63	75.1
	market_cap (\$mi.)	10136	1466	26458	5.31	36.4
	#employees (000)	25.17	5	60.09	6.10	57.1
	Total sales (\$mi.)	7976	1575	22374	9.31	118
Structure	#segments	2.33	2	1.52	1.50	5.76
Ownership	institutional ownership	0.38	0.20	0.91	11.26	161

	<u>Category</u>	<u>Frequency</u>	<u>Percent</u>
Industry	financials	92	27.9%
	manufacturing_durables	54	16.4%
	manufacturing_nondurables	46	13.9%
	mining_construct	20	6.1%
	services	34	10.3%
	utility_transport	64	19.4%
	wholesale_retail	20	6.1%
Culture	Anglo	246	74.5%
	Confucian_Asia	25	7.6%
	Germanic_Europe	14	4.2%
	Latin_America	25	7.6%
	Latin_Europe	14	4.2%
	Nordic_Europe	5	1.5%
	South_Asia	1	0.3%
Quality	No	278	84.2%
	Yes	52	15.8%
Structure (foreign operations)	No	164	49.9%
	Yes	165	50.1%
Ownership (Independence)	A	221	67.0%
	B	29	8.8%
	C	3	0.9%
	D	77	23.3%

Number of observations: 330 for all variables

Two-sample difference. The univariate tests (paired t-test and Wilcoxon signed-rank test) statistically examine the differences between the BSC and control firms (Table 5.2). This

study predicts that significant differences exist because the usage of BSC in executive compensation plans depends on contingent factors. It is important to check whether or not these two groups of firms are indeed dissimilar. The t-test, a parametric test, assumes that the means of both datasets are equal. The increased power associated with this matched-pair test results from the elimination of sources of variability in the sampling distribution of the means. A non-parametric alternative to t-test is Wilcoxon signed test for medians. Because it makes no assumptions of equal variance and normal distribution like t-test, Diekhoff (1992) argued, it provides a useful, and often, powerful alternative to evaluate the differences between two groups of data.

With few exceptions, the test results support the conjecture about the determinants of BSC usage, as the two groups are systematically different. The firms adopting BSC tend to be differentiator firms (prospector) instead of cost leader firms (defender), as evidenced in higher R&D ratio, market-to-book-value ratio, and labour-costs-to-revenue ratio. As predicted, the average firm size in the treatment group is larger than in the control group, with the former showing significantly greater market capitalization values. Yet, total assets of the two samples are close, as expected, because they were matched by industry and size. The treatment firms are more likely to receive quality awards, and have more complicated structure in terms of greater number of segments and foreign operations. Institutional shareholders also seem to play a greater monitoring role in these firms, as shown in higher block ownership and lower independence.

Table 5.2: Two-sample difference tests (contingent variables)

Variable	Group	T-test			Wilcoxon test
		Mean	Std. Err.	Std. Dev.	Median
Strategy	<i>R&D ratio (%)</i>				
	Treatment	0.153*	0.085	1.096	0.000***
	Control	0.019	0.005	0.061	0.000
	<i>MV/BV</i>				
	Treatment	7.270*	3.970	50.991	1.903***
	Control	0.376	0.759	9.748	1.100
	<i>Labour Cost/revenue</i>				
	Treatment	0.096**	0.010	0.130	0.000***
	Control	0.070	0.009	0.114	0.000
Size	<i>Total assets (\$mi.)</i>				
	Treatment	59667	18699	240196	4738***
	Control	46645	13352	171510	2804
	<i>Market Cap (\$mi.)</i>				
	Treatment	13896***	2598	33268	2162***
	Control	6399	1283	16486	1065
	<i>Total employees (th)</i>				
	Treatment	26.6	3.9	50.2	6.0***
	Control	23.8	5.3	68.7	3.7
	<i>Total sales (\$mi.)</i>				
	Treatment	9787	2234	28695	2107***
	Control	6164	1028	13202	1232
Quality	<i>Quality awards</i>				
	Treatment	0.212***	0.032	0.410	0.000***
	Control	0.103	0.024	0.305	0.000
Structure	<i>Segments</i>				
	Treatment	2.78***	0.13	1.62	3.00***
	Control	1.88	0.10	1.26	2.00
	<i>Foreign operations</i>				
	Treatment	0.624***	0.038	0.486	1.000***
	Control	0.378	0.038	0.486	0.000
Ownership	<i>Inst ownership</i>				
	Treatment	0.404	0.095	1.225	0.174**
	Control	0.359	0.033	0.419	0.235
	<i>Independence</i>				
	Treatment	3.02***	0.10	1.34	4.00**
	Control	3.37	0.09	1.14	4.00

- 1) T-test Ho: mean(treatment group) - mean(control group) = diff = 0
2) W-test Ho: median(treatment group) - median(control group) = diff = 0
Ha: diff $\neq$ 0

Degrees of freedom: 328

*** 1% significance level

** 5% significance level

* 10% significance level

Pair-wise correlation. The correlations between dependent and independent variables are summarized in Table 5.3. As predicted, the usage of BSC in executives' compensation plans shows significant relationships to firm characteristics. The non-parametric Spearman rank–order correlations are qualitatively similar to, yet even more significant than, the parametric Pearson correlations. The dichotomy of BSC firms and matched control firms appear to have captured the distinct features of the two groups.

Table 5.3: Correlations between dependent (BSC) and independent (contingent) variables

		<u>Pearson</u>		<u>Spearman</u>	
		corr. coeff.	p-value	corr. coeff.	p-value
Strategy	R&D ratio	0.0864	0.0974*	0.1139	0.0387**
	MV/BV	0.0938	0.0890*	0.3948	0.0000***
	Labour/revenue	0.1078	0.0500**	0.1015	0.0655*
Size	Total assets	0.0313	0.5713	0.0966	0.0797*
	Market cap	0.1419	0.01***	0.1947	0.0004***
	#employees	0.0233	0.6739	0.128	0.0201**
	Total sales	0.0811	0.1417	0.1351	0.014**
Quality	Quality awards	0.1497	0.0064***	0.1497	0.0064***
Structure	Segments	0.2958	0.0000***	0.3346	0.0000***
	Foreign operations	0.2462	0.0000***	0.2462	0.0000***
Ownership	Inst ownership	0.0249	0.6525	-0.1278	0.0203**
	Independence	-0.1401	0.0109**	-0.1327	0.0159**

Number of observations = 330 for all variables

*** 1% significance level

** 5% significance level

* 10% significance level

Multicollinearity. If two or more predictor (independent) variables in a multiple regression model are highly correlated, the coefficient estimates may change erratically in response to small changes in the model or the data. In order to detect whether the model has multicollinearity, the pair-wise correlation matrix between independent variables was analyzed (Table 5.4). The lower left-hand portion of Table 5.4 shows significant Pearson

correlations and the upper right-hand portion significant Spearman correlations. This study concludes that multicollinearity is not a concern based on the following six reasons:

1. Most significant correlations only occur between measures of the same variable, which is good as they are intended to measure the same construct from diverse aspects. The variable “size”, for example, has four indicators including total assets, market value, employees and total sales, which show significant correlations. In contrast, the correlations across different sets of variables are much weaker.
2. Although some correlations are significant, the vast majority of them fall between -0.20 and +0.20, which suggests that the variables included in the determinant model capture rather distinct features of firms’ characteristics. This pattern is consistent with prior research (Ittner et al., 1997; Ashbaugh-Skaife et al., 2007).
3. Since most of the measures are based on continuous data, this study focuses on Pearson correlations, which includes far fewer significant correlations among independent variables than Spearman correlations. The Spearman test does convey useful information, but it is likely to over-estimate correlations due to its less stringent assumptions (Diekhoff, 1992).

Table 5.4: Correlation matrix (contingent variables)

<u>Variables</u>	<u>Indicators</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>
Strategy	A. R&D_ratio	1	0.1986**	-0.2261**	-0.1322	0.0741	0.0647	0.0240	0.2647**	0.3378**	0.3299**	-0.0886	-0.0156
	B. MV/BV ratio	-0.0065	1	-0.0172	0.1734**	0.4697**	0.2693**	0.2910**	0.2228**	0.3352**	0.3744**	-0.2122**	0.1302*
	C. labour/revenue	-0.0002	-0.0347	1	0.3114**	0.1550**	0.1996**	0.1426**	-0.0560	-0.0404	-0.0178	-0.1109*	0.0062
Size	D. total assets	-0.0203	-0.0119	0.0866	1	0.8079**	0.7285**	0.8757**	0.1102*	0.2620**	0.2554**	-0.4503**	0.0351
	E. market cap	0.0183	0.0046	0.062	0.6637**	1	0.7657**	0.8575**	0.1696**	0.4757**	0.4605**	-0.4872**	0.0748
	F. #employees	-0.0112	-0.0101	0.0618	0.2952**	0.3906**	1	0.8820**	0.2868**	0.4810**	0.4583**	-0.4702**	-0.0703
Quality	G. total sales	-0.0024	-0.0078	0.0016	0.4657**	0.8308**	0.3796**	1	0.2452**	0.4512**	0.4245**	-0.4976**	-0.0160
	H. quality	-0.0277	-0.0036	-0.0578	0.0572	0.2145**	0.2948**	0.2706**	1	0.2177**	0.2320**	-0.1167*	0.0243
	I. #segments	0.0379	0.0479	0.0141	0.0403	0.3186**	0.3184**	0.294**	0.2231**	1	0.4992**	-0.2593**	0.0692
Ownership	J. foreign operations	0.065	0.0902	0.0006	0.1146*	0.2655**	0.3024**	0.2266**	0.2311**	0.4407**	1	-0.2605**	0.0583
	K. inst ownership	-0.0225	-0.0164	-0.0268	0.1365*	0.0002	-0.0968	-0.0064	-0.061	-0.1534**	-0.1624**	1	0.3726**
	L. independence	-0.0206	0.0192	0.0475	0.109*	0.1235*	0.0058	0.1029	0.0036	0.0944	0.072	0.1509**	1

The lower left-hand portion: Pearson correlations.

The upper right-hand portion: Spearman rank-order correlations.

Number of observations = 330 for all variables

** 1% significance level

* 5% significance level

4. Multicollinearity does not reduce the predictive power or reliability of the model as a whole; it only produces large standard errors in the related independent variables. With the increase in data, these errors decrease. With 330 observations, larger number than in many similar studies in performance measurement, standard errors are not expected to be unreasonably large in this study.
5. Almost any correlation, no matter how small, is statistically significant, if it is based on a large enough sample. Heiman (1992) pointed out that, with larger samples and thus larger degrees of freedom, the critical values are smaller, and thus given coefficients are more likely to be significant. Diekhoff (1992) found that with samples larger than 100, correlations that approach zero could be statistically significant! The absolute magnitude of the correlation should always be considered when interpreting the meaning and importance of a statistically significant correlation. With the sample of 330 in this study, statistical significance of correlations across variables could be over-emphasized.
6. A formal non-parametric test is also conducted using the variance inflation factor ($VIF=1/\text{tolerance}$) for multicollinearity (O'Brien 2007). This test quantifies the severity of multicollinearity in a regression analysis by providing an index that measures how much the variance of an estimated regression coefficient increases because of multicollinearity. A common rule of thumb is that if $VIF>5$ (or some have argued 10), multicollinearity is high and could be problematic. Model 1 of logistic regression has a $VIF=1.42$, which clearly rules out serious multicollinearity.

1.2 Composite Measures of Contingent Variables

The principal component analysis (PCA) was used to create composite measures for four contingent variables—*strategy*, *size*, *structure* and *ownership*. Composite measures comprise weighted contributions of their respective individual indicators. As a data-reduction procedure, this technique is useful when the researcher has obtained data on a large number of measures, and believes that there is some redundancy in these measures (Diekhoff, 1992). In this case, redundancy means that some measures are correlated with one another, because they measure the same construct. Therefore, it is possible to reduce the observed variables into a smaller number of principal components that account for most of the variance in the observed variables.

PCA is normally conducted in a sequence of steps, with somewhat subjective decisions being made at several steps. It can be performed on covariances or correlations, depending on whether the variables are expressed in the same units. Because the indicators in this study are measured in diverse units, i.e. ROE (%) and total assets (\$1000), the correlation matrix is used to derive the principal components. To obtain reliable results, the minimal size of sample for providing usable data for the analysis should be the larger of 100 observations or five times the number of indicators being analyzed. With 330 firms, this study satisfies this requirement.

The number of eigenvalues (weights) extracted is equal to the number of indicators being analyzed. An eigenvalue represents the amount of variance that is accounted for by a

given component. Under typical conditions, only the first one accounts for a maximal amount of the total variance in the observed variables, implying that it is correlated the most with the observed measures. In other words, the later components tend to account for only trivial variance. Table 5.5 show the eigenvalues and their proportions of variance. This study retains the first eigenvalue for each variable, as it explains the largest proportion of the variance in the original set of measures.

Table 5.5: Eigenvalues for composite contingent variables

	Component	<u>Eigenvalue</u>	<u>Proportion</u>	<u>Cumulative</u>
Strategy (H1A)	1	1.06	0.35	0.35
	2	0.98	0.33	0.68
	3	0.96	0.32	1.00
Size (H1C)	1	2.85	0.71	0.71
	2	0.79	0.19	0.91
	3	0.23	0.06	0.97
	4	0.13	0.03	1.00
Structure (H1E)	1	1.44	0.72	0.72
	2	0.56	0.28	1.00
Ownership (H1G)	1	1.15	0.58	0.58
	2	0.85	0.42	1.00

Using Stata, for example, a composite measure of “size” was developed. It is technically defined as a linear combination of the 4 optimally-weighted indicators (total assets, market value, total sales, total employees). Similarly, the composite measures were developed for “strategy” (R&D ratio, MV/BV, labour cost to revenue), “structure” (segments, foreign operations) and ownership (institutional ownership, firm independence). There is no need to create composites for industry (nominal scale), culture (nominal scale) and quality (single measure).

1.3 Logistic Regression for BSC Usage

The relationships between the usage of BSC and contingent variables were tested using logistic regressions. This method has, in recent years, become the analytical technique of choice for the modeling of categorical dependent variables (DeMaris, 1995). Almost all research in the social and behavioural sciences, and also in economic and marketing research, criminological research, and social medical research deals with the analysis of categorical data (Chintagunta and Prasad, 1998). The explanatory or independent variables may be quantitative, qualitative or both. Logistic regression is popular in part because, as a non-parametric version of regression analysis, it enables the researcher to overcome many of the restrictive assumptions of OLS regression (McClave et al., 2001; Baltagi, 2008):

- Logistic regression does not assume a linear relationship between dependent and independent variables.
- The dependent variable need not be normally distributed.
- The dependent variable need not be homoscedastic for each level of independent variables; that is, there is no homogeneity of variance assumption.
- Normally distributed error terms are not assumed.
- Logistic regression does not require that independent variables be measured on interval or ratio scales.

The hypothesized relationships (H1A to H1G) between BSC in executive compensation and contingent variables were tested using the following regression model:

$$\begin{aligned} \text{Logit (BSC_dummy)} = & \alpha + \beta_1 * \text{Strategy} + \beta_2 * \text{Industry} + \beta_3 * \text{Size} + \beta_4 * \text{Quality} \\ & + \beta_5 * \text{Structure} + \beta_6 * \text{Culture} + \beta_7 * \text{Ownership} + \text{error} \end{aligned} \quad (1)$$

The model assesses the extent to which a firm's contingent factors are associated with BSC adoption. The composite measures developed in Section 1.2 for the contingent variables were used in this analysis.

Table 5.6: Test of H1 - BSC and firm characteristics

	<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Strategy (H1A)	+	0.492	2.13	0.033**
Industry (H1B)	+/-			
Utility_transport	+	1.209	2.30	0.022**
Financials		0.966	1.82	0.069*
Manuf_durables		-1.45	-0.67	0.505
Manuf_nondurables		0.242	0.46	0.644
Mining_construct		-0.165	-0.24	0.809
Services		0.307	0.49	0.625
Wholesale_retail		0.593	0.85	0.393
Size (H1C)	+	0.219	1.18	0.238
Quality (H1D)	+	0.907	1.99	0.047**
Structure (H1E)	+	1.025	5.74	0.000***
Culture (H1F)	+/-			
Anglo	+	4.042	3.49	0.000***
Germanic_Europe		1.420	1.06	0.287
Latin_America		3.158	2.50	0.012**
Nordic_Europe		2.023	1.33	0.185
Ownership (H1G)	+	0.839	3.68	0.000***
Intercept		-4.118	-3.57	0.000***

LR $\chi^2 = 101$

Pseudo $R^2 = 0.2469$

VIF=1.33

Log likelihood = -155

$\chi^2 = 0.000***$

No. of observations = 330 for all variables

*** 1% significance level

** 5% significance level

* 10% significance level

The hypothesis testing results in Table 5.6 can be interpreted based on the sign and statistical significance of the regression coefficients. Pseudo R^2 and χ^2 are the

counterparts of R^2 and F-test in linear regressions, which indicate the model adequacy and significance, respectively. The significance of the individual regression weights are tested using t-statistics (p-value). They are useful in evaluating the relative importance of each predictor variable when combined with the other predictors in multiple regression equations. The null hypothesis in each test is that the regression weight in the equation is zero. A significant regression weight means that the corresponding independent variable adds predictive power to the model.

All of the estimated coefficients on contingent attributes have the expected signs, and most of them are significant at conventional levels. The estimated parameters indicate that a one-unit change in the probability of adopting BSC results from a one-unit change in the relevant predictor variables. For example, an increase of 1 unit in strategy increases the probability that the firm will use BSC measures by 0.49 units.

Using composite measures represents progress made in this study, as previous research tends to use only a single measure for each variable, failing to draw the full picture of the contingent attributes. Therefore, the regression was re-run using individual measures, for the purpose of comparison with previous studies that used individual measures (Table 5.7). They can also reveal which individual indicators are better measures of contingent characteristics. The results for each sub-hypothesis are presented next.

Strategy (H1A). The results suggest that firms engaged in differentiation strategy (prospectors) rather than cost leadership strategy (defenders) are more likely to

implement BSC at 5% significance level (Table 5.6). This is also evidenced in the positive and significant coefficients of R&D ratio and MV/BV ratio significant at the 1% and 5% levels, respectively (Table 5.7). The ratio of labour costs as a percentage of revenue has the expected sign although it is not significant. Taken together, the results suggest that the firms following growth-oriented prospector strategy view BSC measure as valid indicators of managerial performance.

Table 5.7: Test of H1 - BSC and firm characteristics (individual measures)

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Strategy (H1A)	R&D_ratio	+	6.314	2.55	0.010***
	MV/BV	+	0.106	1.91	0.040**
	Labour/revenue	+	1.924	1.33	0.183
Industry (H1B)	Utility_transport	+	1.805	2.85	0.004***
	Financials	+/-	2.000	2.64	0.008***
	Manuf_durables	+/-	-1.276	-0.59	0.554
	Manuf_nondurables	+/-	0.305	0.52	0.603
	Mining_construct	+/-	0.409	0.52	0.606
	Services	+/-	0.925	1.31	0.191
	Wholesale_retail	+/-	1.966	2.54	0.011**
Size (H1C)	Total assets	+	0.001	0.00	0.983
	Market_cap	+	0.305	0.96	0.337
	#Employees	+/-	-0.007	-2.45	0.014**
	Total sales	+	0.045	0.09	0.926
Quality (H1D)	Quality awards	+	1.310	2.35	0.019**
Structure (H1E)	Segments	+	0.755	4.19	0.000***
	Foreign_operations	+	0.993	2.53	0.011**
Culture (H1F)	Anglo	+	4.335	3.62	0.000***
	Germanic_Europe	+/-	2.180	1.59	0.111
	Latin_America	+/-	3.095	2.40	0.016**
	Nordic_Europe	+/-	1.992	1.18	0.239
Ownership (H1G)	Inst ownership/holders	+	0.562	1.77	0.077*
	Independence	-	-0.637	-4.14	0.000***
	Intercept	+/-	-6.850	-4.37	0.000***

LR $\chi^2 = 122.94$	Pseudo $R^2 = 0.3099$	VIF=1.42
Log likelihood = -142.71	$\chi^2 = 0.000$ ***	No. of observations = 330 for all variables
*** 1% significance level	** 5% significance level	* 10% significance level

Industry (H1B). The coefficient on utility industry is both positive and statistically significant at the 5% (Table 5.6) and 1% (Table 5.7) levels, supporting the conjecture that regulatory pressures may lead to many utility and telecommunication firms to measure and reward performance based on BSC measure that consists of safety and customer satisfaction, as well as financial results. As expected, the extent of using BSC varies across industries. Firms operating in the financial industry also have interest in using BSC, significant at 10% (Table 5.6) and 1% (Table 5.7) levels, as does the wholesale and retail industry at 5% level (Table 5.7).

Size (H1C). The coefficients of the composite and log-transformed total assets, market value and total sales have positive signs (Table 5.6 and 5.7), but they are not significant, as the BSC firms are matched by total assets with control firms. Interestingly, there is a negative and significant relationship between the number of employees and the usage of BSC (Table 5.7). It might reflect the possibility that firms with greater number of employees do not necessarily have higher efficiency and effectiveness, both of which are assumed to positively correlate with BSC.

Quality (H1D). Involvement in quality management is positively related to the use of BSC at the 5% level (Table 5.6 and 5.7). This result is consistent with the hypothesis of a positive relationship between the use of quality management practices and the relative importance of non-financial measures in reward systems.

Structure (H1E). Firms with more complex operations, as reflected in the number of business segments and engagement in foreign operations, tend to adopt BSC in their executives' compensation contracts. All coefficients are significant at the 1% or 5% levels (Table 5.6 and 5.7).

Culture (H1F). BSC adoption is more likely in the Anglo cluster (1% significance), followed by the Latin America cluster (5% significance). The results for the Nordic Europe cluster and the Germanic Europe cluster are not significant (Table 5.6 and 5.7).

Ownership (H1G). The usage of BSC significantly correlates with institutional ownership, at the 1% level (Table 5.6). The concentration of institutional holders is positively related to BSC. In other words, highly independent firms are less likely to adopt BSC. Both indicators are significant at the 10% and 1% levels, respectively (Table 5.7).

Both models appear to be well specified. Table 5.6 shows a pseudo R^2 (0.25) and a LR χ^2 statistic (101). χ^2 is significant at the 1% level, and a VIF of 1.33 implies no serious multicollinearity issue. Table 5.7 conveys similar results.

2. Model 2—BSC Usage and Firm Performance

The analytical techniques used in the previous section are repeated for Model 2. But unlike for Model 1, parametric OLS regression is employed to test H2A and H2B.

2.1 Descriptive Statistics and Univariate Tests

Table 5.8: Descriptive analysis (performance and control variables)

	<u>Variable</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>Skewness</u>	<u>Kurtosis</u>
Effectiveness	ROA(%)	-0.4	1.1	13.1	-3.2	19.4
	ROE(%)	-8.6	7.1	152.3	-13.5	222
	ROI(%)	-4.8	3.7	67.5	-9.2	111
	gross_profit_margin(%)	31.8	33.9	73.0	-14.7	248
	EBIT_margin(%)	-2.8	10.1	171.4	-12.6	166
	net_profit_margin(%)	-18.5	4.4	213.2	-13.0	180
	dividend_payout(%)	53.2	0	251.6	8.5	83.1
	stock_return(%)	-0.3	0	51.1	3.0	26.9
	P/E_ratio	9.5	11.5	39.2	-5.4	69.2
	dividend_yield (%)	2.1	0.9	3.2	2.6	12.9
	price/cash_flow	4.9	0	20.6	12.3	194
Efficiency	rev_per_employee(\$1000)	558	302	1026	5.8	44.9
	laborCost_per_employee	26.8	0	41.7	2.8	17.7
	NI_per_employee('000 \$)	-37.2	10.1	664.	-12.9	199
	sales_growth_rate	6.1	0.3	37.9	2.3	52.8
	EPS_growth_rate	-54.2	0.04	446	-8.6	96.3
	inventory_turnover	278	5.21	3977	17.1	303
	receivables_turnover	6.8	4.9	11.6	5.8	45.9
	total_asset_turnover	20.2	0.47	255	13.3	183
	operating_cycle(days)	883	117	1722	2.3	8.4
	quick_ratio	0.99	0.65	2.8	14.8	250
	interest_coverage	8.0	1.4	179	1.5	76.8
	D/E(%)	211	69.5	913	12.1	182
Control variables	bankruptcy_Zscore	1.9	1.3	4.8	7.9	95.1
	industry_ROA%	3.7	3.0	4.0	-0.6	5.3
	industry_ROI_SD	2.9	1.8	2.9	2.3	8.7

	<u>Category</u>	<u>Frequency</u>	<u>Percentage</u>
Merger and acquisition	No	188	57%
	Yes	142	43%
New CEO	No	266	81%
	Yes	64	19%
Product devt cycle	Short	190	58%
	Long	140	42%
Product life cycle	Short	170	52%
	Long	160	48%

Number of observations: 330 for all variables.

Descriptive statistics. Descriptive statistics for performance and control variables are summarized in Table 5.8. Most indicators are normally distributed.

Two-sample difference. The univariate parametric t-tests suggest that the treatment firms significantly outperform control firms in many performance measures (Table 5.9).

Table 5.9: Two-sample difference test (performance and control variables)

Panel A: Effectiveness					
	Group	T-test			Wilcoxon test
		Mean	Std. Err.	Std. Dev.	Median
ROA(%)	Treatment	1.77***	0.84	10.81	2.21***
	Control	-2.70	1.15	14.75	0.65
ROE(%)	Treatment	11.33**	3.40	43.57	10.15***
	Control	-28.69	16.35	209.38	4.12
ROI(%)	Treatment	2.21*	4.33	55.66	5.23***
	Control	-11.93	6.05	77.26	2.04
gross_profit_margin(%)	Treatment	31.2	7.7	99.1	35.35***
	Control	32.6	2.3	29.6	33.04
EBIT_margin(%)	Treatment	1.7	11.5	147.9	12.60***
	Control	-7.5	15.0	192.6	8.68
net_profit_margin(%)	Treatment	-9.7	13.0	167.2	5.89***
	Control	-27.5	19.6	251.2	2.35
dividend_payout(%)	Treatment	54.9	17.7	226.9	13.64
	Control	51.6	21.4	274.8	0.00
stock_return(%)	Treatment	12.78***	4.6	59.6	6.21***
	Control	-13.5	2.9	36.7	-12.03
P/E_ratio	Treatment	19.7***	2.2	28.8	15.65***
	Control	-0.5	3.5	45.3	5.30
dividend_yield(%)	Treatment	3.0***	0.3	3.9	2.03***
	Control	1.3	0.2	2.2	0.00
price/cash_flow	Treatment	8.9***	2.2	27.7	1.77***
	Control	1.0	0.6	7.6	0.00
Panel B: Efficiency					
	Group	T-test			Wilcoxon test
		Mean	Std. Err.	Std. Dev.	Median
sales_per_employee(th)	Treatment	532	78	1004	316**
	Control	585	82	1052	287
LaborCost_Average	Treatment	31*	3	43	0.00**
	Control	23	3	40	0.00

NI_per employee(th)	Treatment	28*	18	232	17.1***
	Control	-103	71	908	4.91
sales_growth	Treatment	4.3	1.7	22.2	0.13
	Control	7.9	3.8	49.0	4.35
EPS_growth	Treatment	16.8***	11.1	143.1	0.18***
	Control	-125.3	47.3	607.5	-9.06
inventory_turnover	Treatment	18.4	3.2	41.0	5.51
	Control	538.9	437.6	5621.4	5.11
receivables_turnover	Treatment	6.6	0.9	11.2	4.86
	Control	7.0	1.0	12.2	5.13
total_asset_turnover	Treatment	39.9	28.1	360.5	0.51
	Control	0.6	0.1	0.7	0.44
operating_cycle(days)	Treatment	840	127	1629	110
	Control	927	141	1815	130
quick_ratio	Treatment	1.2	0.3	3.8	0.73**
	Control	0.8	0.1	1.1	0.59
interest_coverage	Treatment	4.7	13.1	168.7	2.12***
	Control	11.4	14.9	190.9	0.50
D/E(%)	Treatment	128*	21	265	65*
	Control	246	114	1461	74

Panel C: Control variable

	Group	T-test			Wilcoxon test
		Mean	Std. Err.	Std. Dev.	Median
M&A	Treatment	0.527***	0.039	0.501	1.00***
	Control	0.333	0.037	0.473	0.00
new_CEO	Treatment	0.224	0.033	0.418	0.00
	Control	0.164	0.029	0.371	0.00
bankruptcy Z-score	Treatment	2.751***	0.490	6.291	1.83***
	Control	1.125	0.196	2.522	1.02
industry_ROA(%)	Treatment	3.670	0.315	4.047	3.08
	Control	3.824	0.323	4.155	3.00
industry volatility	Treatment	2.894	0.233	2.993	1.88
	Control	3.046	0.235	3.013	1.95
prod_devt_cycle	Treatment	0.424	0.039	0.496	0.00
	Control	0.424	0.039	0.496	0.00
prod_life_cycle	Treatment	0.485	0.039	0.501	0.00
	Control	0.485	0.039	0.501	0.00

- 1) T-test H_0 : mean(treatment group) - mean(control group) = diff = 0
2) W-test H_0 : median(treatment group) - median(control group) = diff = 0
 H_a : diff $\neq$ 0

Degrees of freedom: 328

*** 1% significance level

** 5% significance level

* 10% significance level

Panel A in Table 5.9 presents the results for effectiveness measures of the two sample groups. The treatment firms significantly outperform the control firms on accounting

measures, including the average ROA, ROE and ROI. BSC firms also perform better in EBIT margin, net profit margin, and dividend payout, but the differences are not significant. Regarding market-based performance, the means of all four measures, including stock return, P/E ratio, dividend yield, and price/cash flow, are significantly higher for firms using BSC at the 1% level.

Panel B in Table 5.9 compares the efficiency ratios between firms using BSC and not. For labour cost per employee, net income per employee, EPS growth, and D/E ratio, BSC firms' performance is significantly better than that for the control group at the 1% or 10% level. In terms of total assets turnover, operating cycle, and quick ratio, treatment firms outperform control firms, although the differences are not significant. In contrast, the means of sales per employee, sales growth, inventory turnover, receivables turnover, and interest coverage, are a bit higher for control firms, but none of them is significant.

Panel C in Table 5.9 summarizes the two sample tests for control variables. Compared with control firms, the firms using BSC are unlikely to suffer financial distress, as shown in significantly higher Z-score (2.751). As expected, M&A (significant at the 1% level) seems to play a role in performance between the two samples. There is, however, no significant difference for new CEO, industry ROA, the standard deviation of industry ROI, product development cycle, and product life cycle between the two groups of firms.

This study found that the measures, which are significant using t-tests, are also significant using the non-parametric Wilcoxon median tests. But the reverse is not necessarily true.

The median Wilcoxon tests have six more significant indicators including gross profit margin, EBIT margin, net profit margin, sales per employee, interest coverage and quick ratio. This disparity could occur because non-parametric tests usually require less stringent assumptions. Nevertheless, both tests reveal that the performance of the two groups of firms is indeed different for most measures, implying that the BSC firms generally outperform those firms that do not adopt BSC measures.

One general observation is that effectiveness indicators appear to better capture performance than efficiency indicators, as the former have more significant results. In addition, composite measures for effectiveness and efficiency are considered in Section 2.2 of this chapter.

Pair-wise correlations. The pair-wise correlations between dependent (firm performance) and independent variable (BSC usage) are summarized in Table 5.10. In both Pearson and Spearman correlations, significant relations are found between the usage of BSC and performance measures, including ROA, ROE, ROI, stock return, P/E ratio, dividend yield, price/cash flow, labour cost per employee, net income per employee, and EPS growth rate. These measures seem better than others to capture the effect of BSC on firm performance.

Table 5.10: Correlations between dependent (firm performance) and independent variables (BSC)

		<u>Pearson</u>		<u>Spearman</u>	
		<u>corr. coeff.</u>	<u>p-value</u>	<u>corr. coeff.</u>	<u>p-value</u>
Effectiveness	ROA(%)	0.1709	0.0018***	0.2802	0.0000***
	ROE(%)	0.1316	0.0171**	0.3274	0.0000***
	ROI(%)	0.1049	0.0578*	0.311	0.0000***
	gross_profit_margin(%)	-0.0098	0.8594	0.0973	0.0775*
	EBIT_margin(%)	0.0269	0.6258	0.177	0.0012***
	net_profit_margin(%)	0.0418	0.4488	0.2628	0.0000
	dividend_payout(%)	0.0066	0.9052	0.1532	0.0053***
	stock_return(%)	0.257	0.0000***	0.2563	0.0000***
	P/E_ratio	0.2583	0.0000***	0.3736	0.0000***
	dividend_yield (%)	0.2562	0.0000***	0.2713	0.0000***
	price/cash_flow	0.1922	0.0004***	0.2928	0.0000***
Efficiency	rev_per_employee(\$1000)	-0.0258	0.6411	0.032	0.5624
	laborCost_per_employee	0.0938	0.0891*	0.0967	0.0795*
	NI_per_employee('000 \$)	0.0985	0.0739*	0.2302	0.0000***
	sales_growth_rate	-0.0476	0.3885	-0.0793	0.1508
	EPS_growth_rate	0.1594	0.0037***	0.2024	0.0002***
	inventory_turnover	-0.0655	0.2351	-0.0155	0.7795
	receivables_turnover	-0.0153	0.7812	-0.0111	0.8412
	total_asset_turnover	0.077	0.1629	0.0702	0.2035
	operating_cycle(days)	-0.0254	0.6461	-0.028	0.6128
	quick_ratio	0.0634	0.2505	0.0759	0.169
	interest_coverage	-0.0185	0.7378	0.2423	0.0000***
	D/E(%)	-0.092	0.0953*	-0.0597	0.2795

Number of observations = 330 for all variables

*** 1% significance level

** 5% significance level

* 10% significance level

Multicollinearity. The pair-wise correlations for the independent and control variables are demonstrated in Table 5.11, where the lower left-hand portion of the table presents Pearson correlations and the upper right-hand portion presents the Spearman rank-order correlations. The discussions of multicollinearity in previous section also apply here. Once again, the Pearson test shows a similar pattern to, but much fewer significant correlations than, the Spearman test. Most correlations, however, fall between +0.20 and -0.20, indicating that the variables included in the model capture distinct features of the

two groups of firms. One exception is a Pearson correlation of 0.8146 for the product development and life cycles, both of which measure the same concept of “product cycle” and are expected to highly correlate with each other. Again, the correlations can be over-estimated in a large sample being used. Finally, the VIF statistics for all regressions are close to 1, showing that multicollinearity is not a major concern.

Table 5.11: Correlation matrix (independent and control variables)

	A	B	C	D	E	F	G	H
A. BSC dummy	1	0.1959 **	0.0766	0.1981 **	0.0000	0.0000	-0.0143	-0.0522
B. M&A	0.1959 **	1	0.0071	0.1633 **	0.0589	0.0876	0.0413	0.0868
C. new_CEO	0.0766	0.0071	1	0.0462	-0.0023	-0.0158	-0.0198	0.0255
D. bankruptcy_Zscore	0.1677 **	0.0648	0.0304	1	0.1352 **	0.2826 **	0.5185 **	0.0938
E. prod_devt_cycle	0.0000	0.0589	-0.0135	0.011	1	0.8112 **	0.2466 **	0.0538
F. prod_life_cycle	0.0000	0.0876	-0.0238	0.0504	0.8146 **	1	0.2972 **	0.0029
G. industry_ROA	-0.0188	0.0238	-0.0059	0.2051 **	0.1783 **	0.2164 **	1	0.0613
H. industry_volatility	-0.0254	0.1149 *	-0.0316	0.0502	-0.2196	-0.067	-0.2196 **	1

The lower left-hand portion: Pearson correlations.
The upper right-hand portion: Spearman rank-order correlations.
Number of observations = 330 for all variables

** 1% significance level
* 5% significance level

2.2 Composite Measures of Firm Performance

Principal component analysis is employed to create two composite measures of firm performance, effectiveness and efficiency. Table 5.12 shows the eigenvalues and their proportions of variance. This study retains 2.88 for effectiveness, as the first eigenvalue explains 26% of the variance, which is significantly larger than the other ten. Following the procedure explained in Section 1.2, “effectiveness” is a linear combination of the 11 optimally-weighted indicators (ROA, ROE, ROI, gross profit margin, EBIT margin, net profit margin, dividend payout, stock return, P/E ratio, dividend yield and price/cash

flow). Similarly, this study keeps 2.53 for efficiency in order to maximize the explained proportion of the variance in the original set of measures. “Efficiency” was developed from the 12 indicators (revenue per employee, labour cost per employee, net income per employee, sales growth rate, EPS growth rate, inventory turnover, receivables turnover, total asset turnover, operating cycle, quick ratio, interest coverage and D/E ratio).

Table 5.12 Eigenvalues for composite firm performance

	Component	Eigenvalue	Proportion	Cumulative
Effectiveness	1	2.88	0.26	0.26
	2	1.60	0.15	0.41
	3	1.26	0.11	0.52
	4	1.22	0.11	0.63
	5	0.93	0.08	0.72
	6	0.84	0.08	0.80
	7	0.71	0.06	0.86
	8	0.66	0.06	0.92
	9	0.49	0.04	0.96
	10	0.38	0.03	1.00
	11	0.00	0.00	1.00
Efficiency	1	2.53	0.21	0.21
	2	1.80	0.15	0.36
	3	1.44	0.12	0.48
	4	1.12	0.09	0.57
	5	0.97	0.08	0.65
	6	0.95	0.08	0.73
	7	0.80	0.07	0.80
	8	0.77	0.06	0.87
	9	0.57	0.05	0.91
	10	0.48	0.04	0.95
	11	0.39	0.03	0.99
	12	0.18	0.01	1.00

2.3 OLS Regressions for Performance

OLS regression is employed to assess the direct effects of the usage of BSC in executive compensation plans on corporate performance (H2A & H2B). This technique is a straight-line linear model that derives from the fact that the deterministic portion of the model graphs as a straight line. Key assumptions of the OLS regression include 1) the variables are normally distributed (normality), the error terms have the same variance (homoscedasticity), the independent variables are linearly independent (no multicollinearity), and the errors in the regression are zero (exogeneity) (McClave et al., 2001).

The data were carefully prepared, verified and tested, as discussed in Section 4 of Chapter IV, to satisfy the OLS assumptions, and the descriptive statistics in the preceding Section 2.1 show that the data are suitable for OLS regressions. F-test is used to determine whether the predictor set is globally significant. If it is, then at least one beta is non-zero, and t-tests determine which betas are non-zero. The R^2 represents the proportion of the total sample variability in the dependent variable that can be explained by the relationship between the dependent and independent variables. This sample statistic indicates how well the model fits the data and thereby stands for a measure of the validity of the entire model.

The hypotheses H2A and H2B were tested for the two composite dimensions of performance separately using the following model:

$$\text{Performance (effectiveness)} = \alpha + \beta * \text{BSC_dummy} + \text{error} \quad (2a)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{BSC_dummy} + \text{error} \quad (2b)$$

The regression results using these two composite variables are displayed in Table 5.13. Firm effectiveness is significantly and positively associated with the usage of BSC at the 1% level and firm efficiency at the 10% level, implying that BSC compensation plans add value to firms. The model appears to be well designed as both F-tests are significant at the conventional levels and VIFs are small.

Table 5.13: Test of H2 - BSC and firm performance

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness (H2A)	BSC dummy	+	0.523	2.82	0.005***
	Intercept	+/-	-0.263	-2.00	0.046**
Efficiency (H2B)	BSC dummy	+	0.199	1.86	0.064*
	Intercept	+/-	-0.994	-1.31	0.189
H2A: F-test = 0.0051*** VIF = 1.02 $R^2 = 0.0239$ adj. $R^2 = 0.0209$					
H2B: F-test = 0.0638* VIF = 1.01 $R^2 = 0.0104$ adj. $R^2 = 0.0074$					
No. of observations = 330 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

Firm performance could be confounded by factors other than the usage of BSC. Hence, the regressions were re-run with control variables to ensure the robustness of the results, using the following models:

$$\begin{aligned} \text{Performance (effectiveness)} = & \alpha + \beta * \text{BSC_dummy} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} \\ & + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 \\ & * \text{Ind_volatility} + \text{error} \end{aligned} \quad (2c)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{BSC_dummy} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 * \text{Ind_volatility} + \text{error} \quad (2d)$$

Table 5.14: Test of H2 - BSC and firm performance (with control variables)

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness (H2A)	BSC dummy	+	0.356	1.90	0.059*
	M&A	+	0.366	1.94	0.054*
	New CEO	-	-0.227	-1.00	0.320
	Bankruptcy score	+	0.052	2.69	0.008***
	Product develop cycle	-	-0.607	-1.95	0.053*
	Product life cycle	+	0.294	0.95	0.345
	Industry ROA	+	-0.037	-1.53	0.127
	Industry volatility	-	-0.104	-3.27	0.001***
	Intercept	+/-	0.168	0.78	0.435
Efficiency (H2B)	BSC dummy	+	0.236	2.11	0.036**
	M&A	+	-0.056	-0.50	0.617
	New CEO	-	0.094	0.69	0.493
	Bankruptcy score	-	-0.018	-1.62	0.107
	Product develop cycle	-	0.015	0.08	0.935
	Product life cycle	+	0.053	0.29	0.774
	Industry ROA	+	0.011	0.79	0.429
	Industry volatility	-	-0.004	-0.19	0.848
	Intercept	+/-	-0.139	-1.09	0.275
H2A: F-test = 0.0001*** VIF = 1.10 R ² = 0.0945 adj. R ² = 0.0727					
H2B: F-test = 0.4665* VIF = 1.02 R ² = 0.0234 adj. R ² = -0.0009					
No. of observations = 330 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

Both tests for H2A and H2B are significant, as the coefficients of BSC dummy are significant with predicted signs (Table 5.14). This indicates that the control variables do not have many major confounding effects on firm performance. For the effectiveness regression, M&A and bankruptcy z-score, however, are positive and significant, consistent with the assumptions that firms engaged in M&A and with stronger financial health tend to achieve higher returns. Product development cycle and industry volatility are negatively associated with firm performance. For the efficiency regression, no control variables are significant at the conventional levels.

The bankruptcy score and industry volatility, which pose the greatest potential confounding effects (significant at the 1% level), warrant closer analysis. Confounding, although always a matter of concern, does not necessarily present a serious threat to internal validity. It is less problematic when the confounding variables are known to have little or no effect on the dependent variable, or when their known effects can be taken into account in the analysis (Bordens and Abbott, 1996). Therefore, additional analyses were conducted to examine whether they indeed have strong effects on firm performance. The sample was first divided into two groups by their financial health (bankruptcy z-score)—strong (165 firms) and weak (165 firms). The regressions were then run separately for each group. The results are reported in Appendix 5-Panel A. Neither of the two groups is significant, implying that financial health does not confound the association between firm performance and the usage of BSC. Similarly, the sample was classified into two groups by industry volatility—high (165 firms) and low (165 firms) and the regressions were re-run for each group. Once again, none of the regressions for either group is significant, further showing that the intervening effects of industry volatility is negligible.

2.4 Additional Tests on Individual Measures

For comparison with studies that only used individual instead of composite measures, regressions were also run using 23 individual measures. A brief overview is provided in this section, with detailed statistics in Appendix 5. This analysis gives further evidence

that firm performance is indeed a function of their use of BSC measures, but the significance varies, depending on the individual indicators. In Panel A—Effectiveness, the indicators have the estimated signs on BSC, of which, seven are significant, including ROA, ROE, ROI, stock return, P/E ratio, dividend yield, and price to cash flow. In Panel B, the estimations of efficiency-based indicators also have expected signs, but only four of them are significant, including labour cost per employee, net income per employee, EPS growth rate and D/E ratio. This is an indication of the value relevance of BSC to investors.

It is noteworthy that all four market-based effectiveness measures (stock return, P/E ratio, dividend yield, price/cash flow ratio) are significant at the 1% level, in contrast to accounting-based measures. A possible explanation for this result is that accounting-based measure may be subject to noise effects from generally accepted accounting principles (GAAP) or managers' manipulation and may not necessarily reflect firms' real performance. The second general observation is that effectiveness (Panel A) appears to be a more valid performance indicator than efficiency (Panel B), as the former have much more significant results. It is probably why previous studies almost unanimously prefer the former over the latter in measuring firm performance.

The regressions were also re-run using individual indicators with control variables, with statistics shown in Appendix 6. The pattern is similar to Table 5.14. In Panel A—effectiveness, firm performance can be positively affected by M&A and bankruptcy, depending on individual indicators. Interestingly, it is negatively associated with new

CEOs. It is possible that poor performance brings new CEOs into the troubled firms. Economic noise mostly has a negative effect on firm performance. In addition, firms with shorter product development cycle and longer product life cycle are more likely to perform better, which are consistent with predictions. Labour costs per employee and EPS growth are significant for both regressions. Net income per employee and D/E ratio are significant for regressions without control variables but insignificant with control variables. The positive coefficient of labour cost per employee may simply imply higher expenses related to research and development.

In Panel B-efficiency, the effects of control variables, in general, are weak, as the signs and values of coefficients for *BSC_dummy* are similar for the regressions with and without control variables. The majority of coefficients for control variables are insignificant. It implies that control variables do not have much confounding effects on firm performance, and the usage of BSC contributes to firm performance. Nevertheless, the coefficients of M&A for operating cycle, and new CEO for D/E ratio, are positive and significant. The significant negative coefficients of bankruptcy Z-score for operating cycle and positive one for quick ratio indicate that firms with less financial distress are able to achieve higher efficiency. Economic noise, as shown in industry ROA and the standard deviation of industry ROI, has a widespread impact on many indicators, such as labour costs per employee and operating cycle. The coefficient of product development cycle for labour costs per employee is negative and significant. Longer product life cycle is likely to be negatively associated with operating cycle.

3. Model 3—Contingent Fit and Firm Performance

This section investigates the fit hypotheses (H3A & H3B) that the relationship between the use of BSC and performance is a function of the “fit” or “match” of a firm’s operational and competitive circumstances with its choice of BSC scheme.

3.1 Descriptive Statistics

Table 5.15: Two-sample t-test (fit)

	<u>Group</u>	<u>Mean</u>	<u>Std. Err.</u>	<u>Std. Dev.</u>	<u>[95% Conf.</u>	<u>Interval]</u>
Fit_BSC	Treatment	0.714***	0.019	0.241	0.677	0.752
	Control	0.338	0.020	0.230	0.299	0.377
Ho: mean(treatment group) - mean(control group) = diff = 0						
Ha: diff ≠ 0						
Degrees of freedom: 328						
*** 1% significance level ** 5% significance level * 10% significance level						

Table 5.15 reports fitted values of the treatment (0.714) and control (0.338) groups. The mean differences are significant at the 1% level, indicating that firms adopting BSC tend to attain a higher fit with their contingent characteristics.

3.2 OLS Regressions for Fit

The following model is used to capture the relationship between performance and the fit of BSC with firm characteristics:

$$\text{Performance (effectiveness)} = \alpha + \beta * \text{Fit_BSC} + \text{error} \quad (3a)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Fit_BSC} + \text{error} \quad (3b)$$

Performance is measured by the composite measures of effectiveness and efficiency respectively, and Fit_BSC is the extent of match of BSC with contingent variables (see Chapter IV, Section 3.5).

Table 5.16: Test of H3 - BSC fit and firm performance

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness (H3A)	BSC fit	+	0.893	2.65	0.008***
	Intercept	+/-	-0.478	-2.29	0.023**
Efficiency (H3B)	BSC fit	+	0.498	2.20	0.028**
	Intercept	+/-	-0.287	-2.05	0.042**
H3A: F-test = 0.0085*** VIF = 1.02 $R^2 = 0.0233$ adj. $R^2 = 0.0200$					
H3B: F-test = 0.0284* VIF = 1.02 $R^2 = 0.0161$ adj. $R^2 = 0.0128$					
No. of observations = 330 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

Table 5.16 reports the results for *effectiveness* and *efficiency*. The evidence shows that firm performance is a function of the match of the firm's BSC measure with its contingent characteristics. The parameters estimate a one-unit change in firm performance resulting from a one-unit change in the fit of BSC. An increase of 1 unit in *Fit_BSC* increases firm performance by 0.893 (effectiveness) and 0.498 (efficiency) units. Both constructs are significant at 1% and 5% levels, respectively.

Once again, the intervening effect of control variables was also tested, using the following models:

$$\begin{aligned} \text{Performance (effectiveness)} = & \alpha + \beta * \text{Fit_BSC} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + \\ & Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + \\ & Z_7 * \text{Industry_volatility} + \text{error} \end{aligned} \quad (3c)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Fit_BSC} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 * \text{Industry_volatility} + \text{error} \quad (3d)$$

The control variables are the covariates for performance discussed earlier. Firm performance should be better (worse) when the firm's use of BSC measure does (does not) fit the firm's particular economic characteristics. A positive relationship between performance and fit is predicted.

Table 5.17: Test of H3 – BSC fit and firm performance (with control variables)

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness (H3A)	BSC fit	+	0.561	1.64	0.100*
	M&A	+	0.407	1.97	0.050**
	New CEO	-	-0.115	-0.47	0.640
	Bankruptcy score	+	0.077	2.64	0.009***
	Product develop cycle	-	-0.610	-1.80	0.073*
	Product life cycle	+	0.251	0.74	0.459
	Industry ROA	+	-0.053	-1.90	0.058*
	Industry volatility	-	-0.102	-2.79	0.006***
	Intercept	+/-	0.040	0.15	0.883
Efficiency (H3B)	BSC fit	+	0.476	2.11	0.036**
	M&A	+	0.294	2.15	0.033**
	New CEO	-	-0.036	-0.22	0.825
	Bankruptcy score	-	-0.037	-1.93	0.055*
	Product develop cycle	-	-0.207	-0.93	0.353
	Product life cycle	+	-0.246	-1.10	0.270
	Industry ROA	+	-0.028	-1.54	0.124
	Industry volatility	-	-0.093	-3.89	0.000***
	Intercept	+/-			
H3A: F-test = 0.0002*** VIF = 1.11 R ² = 0.0994 Adj. R ² = 0.0743					
H3B: F-test = 0.0000* VIF = 1.14 R ² = 0.1224 Adj. R ² = 0.0982					
No. of observations = 330 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

The results are presented in Table 5.17. Both coefficients of BSC fit are significant, similar to the case without control variables, but only at the 10% level for effectiveness.

Although the influence from control variables is not strong enough to change the conclusions, several variables do intervene the effects of fit on firm performance. Firm performance is positively affected by M&A and bankruptcy score, and negatively by product development cycle and industry volatility.

Similar to the analyses in Section 2.3, the influence of financial health and industry volatility on firm performance was also investigated. The sample was divided into two groups and separate regressions were run. Again, most regressions for both groups are insignificant at the conventional levels (Appendix 5-Panel B), except for financial health, indicating only a weak influence of control variables. Bordens and Abbott (1996) argued that a study may include confounding variables and still maintain a fair degree of internal validity, if the effects of confounding variables under scrutiny are known. This is fortunate, because it is often impossible to eliminate all sources of confounding in a study.

3.3 Additional Tests on Individual Measures

As in Section 2.4, regressions were also re-run using individual measures. The results show that all 11 effectiveness indicators are positively associated with the fit and six of them are significant (Appendix 7). The efficiency indicators also have expected signs, but only three are significant. It appears that firms appropriately matching their BSC usage with firm characteristics can achieve higher performance, despite that some indicators are

not significant at the conventional levels. In other words, mismatched firms are likely to be associated with poor performance.

The results with control variables are reported in Appendix 8. As expected, a small number of control variables are significant, depending on which performance indicator is used. ROA for healthy firms, for example, positively correlates with long product life cycle, and negatively with economic volatility and long product development cycle.

The F-statistics for effectiveness and efficiency are significant at the 1% and 10% levels (Tables 5.16 and 5.17), respectively, indicating that the overall results of the regressions are strong. The modest R^2 tests are consistent with prior research. The hypothesis (H3) predicts that the contingent-fit of BSC with its contextual variables is positively related to firm performance. The positive and significant coefficients of BSC Fit support this hypothesis. The firms seem to benefit from aligning their BSC with their contextual variables.

4. Summary

Differences are observed between BSC firms and closely matched non-BSC firms. Consistent with expectations, this study finds that prospector firms rely more heavily on BSC than defender firms. Likewise, the use of BSC is greater in quality-oriented firms. The results support the claims that firms in certain industries and culture clusters are more likely to use BSC than others. Other variables, including size, structure and

ownership, also confirm their associations with BSC usage in executive incentive programs.

The overall evidence suggests that firm performance is positively related to the use of BSC, and that performance is maximized when the choice of BSC fits into the company's operational and competitive environment.

VI. Results: Sub-sample of Firms with Non-financial Weights

Using the sub-sample of 58 firms, this chapter examines whether the relative weights of non-financial measures are contingent on firm characteristics, whether firm performance is associated with the relative weights of non-financial measures, and whether performance is related to the predicted fit between non-financial weights and firm characteristics. By substituting the BSC dummy with the relative weights of non-financial measures, this chapter repeats the analytical procedures used in Chapter V.

Contextual factors vary across firms and, in turn, adopting appropriate non-financial measures could determine the performance consequences of such measures. The findings support the contention that firms with certain characteristics place more weights on non-financial measures. Yet, this study only finds a weak association between firm performance and 1) the relative weights of non-financial measures and 2) the fitted weights of non-financial measures. These results suggest that the relative weights of non-financial measures only work if they are used in an optimal way. Over-relying or under-relying on non-financial measures do not necessarily lead to better firm performance.

1. Model 1—Non-financial Weights and Firm Characteristics

This section investigates the relationships between the relative weights of non-financial measures and firm characteristics.

1.1 Descriptive Statistics and Univariate Tests

Table 6.1: Descriptive statistics (sub-sample contingent variables)

Panel A:						
	<u>Variable</u>	<u>Mean</u>	<u>Median</u>	<u>Std. Dev.</u>	<u>Skewness</u>	<u>Kurtosis</u>
Dependent	nonfin_weight	0.48	0.5	0.18	0.06	2.15
Strategy	R&D_ratio	0.03	0	0.11	4.14	20.1
	MV/BV_ratio	3.16	2.02	6.78	6.74	49.1
	labour/rev	0.12	0.12	0.13	0.93	4.39
Size	total_assets	74662	6693	214160	3.97	18.6
	market_cap	20991	3739	44054	3.27	13.9
	#employees	24.07	4.82	43.11	2.74	10.7
	total_sales	16867	2732	45213	5.46	35.9
Quality	quality_awards	0.24	0	0.43	1.20	2.46
Structure	#segments	2.84	2	1.97	1.45	4.95
	foreign_oper_dum	0.56	1	0.49	-.27	1.07
Ownership	inst ownership	0.48	0.18	1.84	7.19	53.8
	independence	3.31	4	1.14	-1.34	3.10
Panel B:						
	<u>Category</u>	<u>Freq.</u>		<u>Percent</u>		
Industry	financials	19		32.8%		
	utility_transport	14		24.1%		
	manufacturing_nondurables	9		15.5%		
	manufacturing_durables	8		13.8%		
	mining_construct	6		10.3%		
	services	2		3.5%		
Culture	Anglo	54		93.1%		
	Germanic_Europe	3		5.2%		
	Sub_Sahara_Africa	1		1.7%		
Quality	No	44		75.8%		
	Yes	14		24.1%		
Structure (foreign operat.)	No	25		43.1%		
	Yes	33		56.9%		
Ownership (independence)	A	39		67.2%		
	B	8		13.8%		
	C	1		1.7%		
	D	10		17.2%		

No. of observations: 58 for all variables

Table 6.1 exhibits the descriptive statistics. The mean weight placed on non-financial measures of BSC is 48.6% in the sub-sample. This is remarkable, compared to only

37.1% in Ittner et al.'s (1997) study with 114 firms that employed non-financial measures. It may be that firms adopting BSC more consciously emphasize non-financial measures. Compared to the corresponding table in Chapter V, the data have much smaller skewness and kurtosis, implying more normal distributions. Most firms are concentrated in financial and utility industries (33% and 24%) and Anglo cluster (93%), 24% of the firms have won quality awards, 57% engage in foreign operations, and 67% are highly independent.

Table 6.2 reports bivariate correlations between the dependent and independent variables. For Pearson correlations, the number of employees, segments and foreign operations are significant at the 5% or 10% level. For Spearman correlations, R&D ratio, segments and foreign operations are significant at the 5% or 10% level.

Table 6.2: Correlations between dependent (non-financial weights) and independent (contingent) variables

		<u>Pearson</u>		<u>Spearman</u>	
		<u>Coefficients</u>	<u>p-value</u>	<u>Coefficients</u>	<u>p-value</u>
Strategy	R&D ratio	0.1508	0.2584	-0.2155	0.1042*
	MV/BV	-0.0917	0.4934	-0.2019	0.1285
	Labour/revenue	0.0375	0.7797	0.0995	0.4574
Size	Total assets	-0.0081	0.9516	0.1598	0.2309
	Market cap	-0.1111	0.4063	-0.0241	0.8575
	#employees	-0.2600	0.0487*	-0.1465	0.2725
	Total sales	-0.0507	0.7054	-0.0601	0.6542
Quality	Quality award	-0.1022	0.4452	-0.0992	0.4586
Structure	Segments	-0.2858	0.0296**	-0.2183	0.0998*
	Foreign operations	-0.2809	0.0327**	-0.3032	0.0207**
Ownership	Inst ownership	0.0378	0.7782	0.1773	0.1830
	Independence	-0.0473	0.7245	-0.0817	0.5420

No. of observations = 58 for all variables
 *** 1% significance level ** 5% significance level 10% significance level

Table 6.3: Correlation matrix (sub-sample contingent variables)

<u>Variables</u>	<u>Indicators</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>
Strategy	A. R&D_ratio	1	0.283*	-0.2092	-0.2824*	-0.0464	0.0253	-0.0490	0.1050	0.3094*	0.4599**	-0.1644	-0.0921
	B. MV/BV_ratio	-0.0227	1	-0.0645	0.1268	0.4478**	0.3382**	0.3668**	0.2214	0.1898	0.4815**	-0.2219	0.0474
	C. labour/revenue	-0.2069	0.0178	1	0.2157	0.1138	0.2631*	0.0538	0.0538	-0.0231	-0.0727	-0.0236	-0.0750
Size	D. total_assets	-0.1056	-0.0105	0.0577	1	0.8410**	0.7719**	0.8664**	0.2455	0.3293*	0.1861	-0.3770**	0.0987
	E. market_cap	-0.1081	-0.008	-0.0956	0.5534**	1	0.8155**	0.9157**	0.2190	0.5568**	0.4752**	-0.5166**	0.1466
	F. #employees	-0.0826	-0.0173	0.2159	0.4298**	0.6822**	1	0.9077**	0.4380**	0.5603**	0.5386**	-0.4967**	0.0129
Quality	G. total_sales	-0.0938	-0.0208	-0.1145	0.3294*	0.8712**	0.502**	1	0.3321*	0.5488**	0.4440**	-0.4591**	0.0896
	H. quality_dummy	-0.0462	0.2345	-0.0647	0.0184	0.2979*	0.437**	0.2891*	1	0.1655	0.2469	-0.2241	0.0261
	I. #segments	-0.081	-0.0224	0.0014	0.1334	0.3853**	0.6019**	0.2682*	0.1478	1	0.5870**	-0.4402**	0.1118
Ownership	J. foreign_operation	0.1104	0.1954	-0.0243	0.1041	0.295*	0.3846**	0.2372	0.2469	0.4831**	1	-0.3977**	-0.0827
	K. inst_ownership	-0.017	0.0204	-0.1111	0.5509**	0.1089	-0.1112	0.0691	-0.0929	-0.181	-0.2029	1	0.3955**
	L. independence	-0.0329	-0.2894*	0.0127	0.1339	0.128	0.152	0.0579	-0.0123	0.1541	-0.0689	0.112	1

The lower left-hand portion: Pearson correlations.

The upper right-hand portion: Spearman rank-order correlations.

Number of firms: 58 for all variables

** 1% significance level

* 5% significance level

Table 6.3 shows the correlation matrix for the independent variables. Some variables appear to correlate with each other, but most correlations fall between +0.20 and -0.20. Also, a VIF of 1.36 further rules out the possibility of serious multicollinearity. Like for the full sample, Pearson correlations are, qualitatively similar to, but much less significant than, Spearman correlations

1.2 Composite Measures of Contingent Variables

Following the principal component analysis in Chapter V, four composite measures of firm characteristics were also generated from their respective individual measures. Table 6.4 shows that the first eigenvalue of each variable accounts for a significant proportion of total variance: 41%, 68%, 74% and 56% respectively. They are used to develop the eigenvectors for regressions.

Table 6.4: Eigenvalues for composite contingent variables (sub-sample)

	Component	<u>Eigenvalue</u>	<u>Proportion</u>	<u>Cumulative</u>
Strategy (H1A)	1	1.22	0.41	0.41
	2	0.99	0.33	0.73
	3	0.79	0.26	1.00
Size (H1C)	1	2.72	0.68	0.68
	2	0.71	0.18	0.86
	3	0.50	0.12	0.98
	4	0.08	0.02	1.00
Structure (H1E)	1	1.48	0.74	0.74
	2	0.52	0.26	1.00
Ownership (H1G)	1	1.11	0.56	0.56
	2	0.89	0.44	1.00

1.3 OLS Regressions for Non-financial Weights

The hypotheses H1A to H1G were re-tested for the sub-sample, using the following equation:

$$\begin{aligned} \text{Non-financial_weights} = & \alpha + \beta_1 * \text{Strategy} + \beta_2 * \text{Industry} + \beta_3 * \text{Size} + \beta_4 * \text{Quality} \\ & + \beta_5 * \text{Structure} + \beta_6 * \text{Culture} + \beta_7 * \text{Ownership} + \text{error} \end{aligned} \quad (1)$$

The results using composite measures for strategy, size, structure and ownership are presented in Table 6.5, and using individual measures in Table 6.6. The patterns in both analyses are similar.

Table 6.5: Non-financial weights and firm characteristics (sub-sample)

	<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Strategy (H1A)	+	0.061	1.92	0.061*
Industry (H1B)	+/-			
utility_transport	+	0.324	2.21	0.032**
financials		0.322	2.24	0.030**
manuf_durables		0.217	1.50	0.142
manuf_nondurables		0.199	1.41	0.165
mining_construct		0.225	1.54	0.130
Size (H1C)	+	0.036	1.69	0.098*
Quality (H1D)	+	0.035	0.54	0.593
Structure (H1E)	+	0.059	1.74	0.088*
Culture (H1F)	+/-			
Anglo	+	0.293	1.55	0.128
Germanic_Europe		0.137	0.63	0.533
Ownership (H1G)	+	0.014	0.58	0.056*
Intercept		0.015	0.06	0.949
<div> <div> $R^2 = 0.3187$ adj. $R^2 = 0.1370$ *** 1% significance level </div> <div> F-test= 0.0866* VIF=1.47 ** 5% significance level </div> <div> No. of observations = 58 for all variables * 10% significance level </div> </div>				

Strategy. The prospector firms rely more heavily on non-financial measures than defender firms, as the coefficient is significantly positive at the 10% level (Table 6.5). R&D ratio is positive and significant at the 1% level. Both labour-to-revenue ratio and market-to-book ratio are not significant (Table 6.6). R&D ratio appears to be a better indicator to measure strategy than the other two.

Table 6.6: Non-financial weight and firm characteristics (sub-sample individual measures)

Variable	Indicator	Predicted sign	Coefficient	t-stat	p-value
Strategy	R&D_ratio	+	0.810	2.96	0.005***
	MV/BV_ratio	+	0.002	-0.6	0.554
	labour/revenue	+	0.155	0.72	0.478
Industry	utility_transport	+	0.407	2.57	0.014**
	financials	+/-	0.348	2.07	0.045**
	manuf. durables	+/-	0.199	1.19	0.240
	manuf. nondurables	+/-	0.173	0.98	0.331
	mining	+/-	0.262	1.72	0.094*
Size	total_assets	+	0.195	1.72	0.094*
	market_cap	+	0.031	0.68	0.499
	#employees	+/-	-0.003	-2.88	0.007***
	total_sales	+	0.078	-0.64	0.529
Quality	quality_award	+	0.063	0.89	0.382
Structure	#segments	+	0.001	0.06	0.953
	foreign_oper_dum	+	0.011	0.15	0.883
Culture	Anglo	+	0.359	1.78	0.082*
	German. Europe	+/-	0.106	0.40	0.693
Ownership	inst/holders	+	0.001	0.08	0.940
	independence	-	-0.009	-0.36	0.721
	intercept	+/-	-0.399	-1.45	0.154

No. of observations=58 for all variables

F-test=0.0267** $R^2=0.5102$ adj. $R^2=0.2653$ VIF = 1.36

***1% significance level **5% significance level 10% significance level

Industry. The significant positive coefficient for utility industry (5%) supports the claim that regulated industries are more likely to use non-financial measures than non-regulated ones (Table 6.5 and 6.6). Regulated industries face greater pressures to improve non-

financial dimensions, such as safety and customer satisfaction. In addition, financial industry is also significantly associated with the use of non-financial measures at the 5% level (Table 6.5 and 6.6).

Size. Larger firms focus more on non-financial measures, as demonstrated in Table 6.5. The total assets measure has the expected sign and is significant at the 10% level. In contrast, the number of employees is negative and significant at the 1% level, indicating that firms with fewer employees tend to place greater weights on non-financial measures (Table 6.6). Combining evidence from Chapter V, this study finds that total assets, market capitalization and total sales appear to be better measures of size than the number of employees.

Quality. The weight of non-financial components is positively, yet insignificantly, correlated with quality-oriented firms (Table 6.6 and 6.6). Even so, the expected sign may still indicate that firms seek to increase their competitiveness by employing innovative quality-oriented management initiatives and utilizing performance measurement systems that include a broad range of financial and forward-looking non-financial measures.

Structure. Table 6.5 shows that firms with more complicated structures are likely to place more weight on non-financial measures. Yet, neither the number of segments nor foreign operations is significant although positive (Table 6.6).

Culture. Firms in the Anglo cluster appears to place more weight on non-financial measures at the 10% significance level (Table 6.6), implying only a weak relationship between culture and non-financial measures.

Ownership. Firms with greater concentration of block ownership and with lower independence appear to put more emphasis on non-financial measures (Table 6.5 and 6.6), although they are not significant.

Overall, the results are similar to the full-sample analyses using the BSC dummy in Chapter V, although the magnitude of significance levels varies. The model appears reasonably well specified. The R^2 in Table 6.6, for example, is 51%, indicating that the variables explain a substantial proportion of the variance in the weights placed on non-financial measures in BSC. F-test is significant at the 5% level. Replacing BSC dummy in Chapter V with non-financial weights substantially increases the robustness of the model.

2. Model 2— Firm Performance and Non-financial Weights

This section examines whether firm performance is associated with the relative weights placed on non-financial measures.

2.1 Descriptive Statistics and Univariate tests

Table 6.7: Descriptive analysis (sub-sample performance and control variables)

	<u>Variable</u>	<u>Mean</u>	<u>Median</u>	<u>Std.Dev.</u>	<u>Skewness</u>	<u>Kurtosis</u>
Effectiveness	gross_profit_margin(%)	39.7	35.4	26.1	-1.2	8.7
	EBIT_margin(%)	14.2	17.2	29.0	-2.6	15.0
	net_profit_margin(%)	3.5	8.0	25.4	-4.2	21.4
	ROA(%)	2.6	1.7	8.8	-1.6	12.2
	ROE(%)	19.3	12.2	62.9	3.7	20.4
	ROI(%)	9.7	5.1	39.4	5.8	42.1
	dividend_payout(%)	79.6	21.9	329	7.1	53.1
	stock_return(%)	9.7	5.7	48.1	2.0	10.9
	P/E_ratio	17.3	15.6	25.2	2.2	18.9
	dividend_yield (%)	3.1	2.7	3.6	3.0	16.6
	price/cash_flow	4.9	2.7	6.6	1.2	3.8
Efficiency	rev_per_employee(\$1000)	811	407	1524	5.1	31.8
	laborCost_per_employee	43.1	43.6	45.3	1.5	7.2
	NI_per_employee('000 \$)	74.3	37.4	259	3.6	23.2
	sales_growth_rate	0.04	0.05	9.1	-2.7	25.8
	EPS_growth_rate	-14.8	0.06	121.2	-7.2	54.3
	inventory_turnover	10.4	4.8	19.3	4.5	26.8
	receivables_turnover	4.7	4.0	5.5	2.5	12.8
	total_asset_turnover	0.4	0.3	0.4	1.3	4.8
	operating_cycle(days)	1099	117	1690	1.3	3.4
	quick_ratio	0.6	0.6	0.6	0.9	3.8
	interest_coverage	13.4	1.9	59.1	5.4	34.6
	D/E(%)	172	87	392	5.0	29.5
Control variables	bankruptcy_Zscore	1.6	1.3	1.6	1.2	4.9
	industry_ROA%	2.9	2.7	3.8	0.1	3.3
	industry_volatility	3.1	1.9	2.8	1.8	6.7

	<u>Category</u>	<u>Frequency</u>	<u>Percentage</u>
Merger and acquisition	No	24	41.4%
	Yes	34	58.6%
New CEO	No	44	75.9%
	Yes	14	24.1%
Product development cycle	Short	33	56.9%
	Long	25	43.1%
Product life cycle	Short	27	46.6%
	Long	31	53.4%

No. of observations: 58 for all measures

Table 6.7 summarizes descriptive analyses for performance and control variables, such as mean, median, standard deviation, skewness and kurtosis and the frequencies for nominal data. Compared to the corresponding table in Chapter V, this table reports much smaller skewness and kurtosis, implying that the data is more normally distributed. Fifty-nine percent of the firms had one or more M&A and 24% hired new CEOs. More than half of them have long product development cycles and short life cycles.

Table 6.8: Sub-sample correlations between dependent (firm performance) and independent (non-financial weights) variables

		<u>Pearson</u>		<u>Spearman</u>	
		<u>Coeff.</u>	<u>p-value</u>	<u>Coeff.</u>	<u>p-value</u>
Effectiveness	ROA(%)	-0.2867	0.0291**	-0.2232	0.0921*
	ROE(%)	-0.1928	0.1472	-0.0477	0.7222
	ROI(%)	-0.2057	0.1214	-0.1492	0.2636
	gross_profit_margin(%)	0.0123	0.9269	0.0384	0.7748
	EBIT_margin(%)	-0.0731	0.5855	0.1308	0.3277
	net_profit_margin(%)	-0.1765	0.185	-0.0252	0.8509
	dividend_payout(%)	0.1063	0.427	0.0016	0.9906
	stock_return(%)	-0.1196	0.3711	-0.1033	0.4403
	P/E_ratio	-0.0126	0.9253	-0.1477	0.2686
	dividend_yield (%)	0.0718	0.5921	0.0301	0.8225
	price/cash_flow	-0.3643	0.0049***	-0.3870	0.0027***
Efficiency	rev_per_employee(\$1000)	0.1091	0.4149	0.1686	0.2058
	laborCost_per_employee	0.1147	0.3912	0.1839	0.1671
	NI_per_employee('000 \$)	0.073	0.5863	0.0931	0.4869
	sales_growth_rate	-0.1221	0.3612	-0.0536	0.6894
	EPS_growth_rate	-0.0336	0.8022	-0.0627	0.6402
	inventory_turnover	0.1693	0.2039	0.2481	0.0604*
	receivables_turnover	0.0389	0.772	-0.0055	0.9675
	total_asset_turnover	-0.3178	0.015**	-0.2960	0.0241**
	operating_cycle(days)	0.1617	0.2253	0.0180	0.8933
	quick_ratio	-0.1943	0.1438	-0.2544	0.0540*
	interest_coverage	-0.1081	0.4192	-0.1816	0.1726
	D/E(%)	0.0187	0.8892	0.1476	0.2689

Number of observations = 58 for all measures

*** 1% significance level

**5% significance level

*10% significance level

Table 6.8 summarizes correlations between the dependent and independent variables. Both parametric and non-parametric tests show that ROA, price-to-cash flow and total assets turnover are significant. In addition, inventory turnover and quick ratio are significant for Spearman correlations.

The correlations between the independent variables are listed in Table 6.9. The only significant Pearson correlation is between non-financial weights and bankruptcy z-score. The pattern is consistent with the postulation that firms that suffer financial distress emphasize more non-financial measures. Noticeably, this matrix has many fewer significant correlations than the parallel table in Model 2 of Chapter 5. This confirms a previous discussion that a smaller sample typically results in fewer significant correlations.

Table 6.9: Correlation matrix (sub-sample independent and control variables)

	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>
A. non-fin weight	1	-0.0925	0.1815	-0.4068**	0.2123	-0.0831	-0.1061	0.1100
B. M&A	-0.0915	1	-0.0987	0.2289	-0.117	0.1283	0.1579	-0.0031
C. new_CEO	0.189	-0.0987	1	0.0404	0.0786	-0.0390	0.0795	-0.0734
D. bankrupt_Zscore	-0.3573**	0.2516	-0.0519	1	0.2731*	0.5444**	0.5632**	0.5632
E. prod_dev_cycle	0.209	-0.117	0.0786	0.2441	1	0.7425**	0.3526**	0.0947
F. prod_life_cycle	-0.0775	0.1283	-0.039	0.4736	0.7425	1	0.4585**	0.0083
G. industry_ROA %	-0.0521	0.1796	0.0664	0.5309	0.3102	0.3394	1	-0.0263
H. SD_volatility	0.2084	0.0036	-0.039	-0.0383	-0.0649	-0.1308	-0.1232	1

The lower left-hand portion: Pearson correlations.
The upper right-hand portion: Spearman rank-order correlations.
Number of observations = 58 for all variables ** 1% significance level *5% significance level

The VIFs range from 1.1 to 2.4, indicating that no serious multicollinearity is detected in the data.

2.2 Composite Measures of Firm Performance

Relying on the principal component analysis, two composite measures for firm performance were created, as shown in Table 6.10. As discussed in Chapter 5, the largest eigenvalues are used for *effectiveness* (3.67) and *efficiency* (3.33). Both components account respectively for 33% and 28% of the total variance, which are significantly higher than any others.

Table 6.10: Eigenvalues for composite firm performance (sub-sample)

	<u>Component</u>	<u>Eigenvalue</u>	<u>Proportion</u>	<u>Cumulative</u>
Effectiveness	1	3.67	0.33	0.33
	2	1.94	0.18	0.51
	3	1.46	0.13	0.64
	4	1.24	0.11	0.76
	5	0.97	0.09	0.84
	6	0.65	0.06	0.90
	7	0.35	0.03	0.93
	8	0.30	0.03	0.96
	9	0.22	0.02	0.98
	10	0.14	0.01	1.00
	11	0.06	0.01	1.00
Efficiency	1	3.33	0.28	0.28
	2	2.70	0.23	0.50
	3	1.51	0.13	0.63
	4	1.16	0.10	0.73
	5	1.00	0.08	0.81
	6	0.62	0.05	0.86
	7	0.56	0.05	0.91
	8	0.38	0.03	0.94
	9	0.34	0.03	0.97
	10	0.21	0.02	0.99
	11	0.16	0.01	1.00
	12	0.02	0.00	1.00

2.3 OLS Regressions for Performance

Hypotheses H2A and H2B were re-tested for the sub-sample in this section, using the following equations:

$$\text{Performance (effectiveness)} = \alpha + \beta * \text{Non-financial_weights} + \text{error} \quad (2a)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Non-financial_weights} + \text{error} \quad (2b)$$

The regression results are reported in Table 6.11. Firms that place greater weights on non-financial measures are more likely to achieve greater effectiveness (significant at 1%) and efficiency (positive but insignificant).

Table 6.11: Non-financial weights and firm performance (sub-sample)

		<u>Expected sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness	Non-financial weight	+	2.583	1.93	0.058*
	Intercept	+/-	1.256	1.81	0.076*
Efficiency	Non-financial weight	+	1.811	1.40	0.166
	Intercept	+/-	-0.881	-1.31	0.195
Effectiveness: F-test = 0.0581* VIF = 1.07 $R^2 = 0.0627$ Adj. $R^2 = 0.0459$					
Efficiency: F-test = 0.1660 VIF = 1.04 $R^2 = 0.0340$ Adj. $R^2 = 0.0167$					
No. of observations = 58 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

Similarly, the following OLS regressions were re-run to check the confounding effects of control variables:

$$\begin{aligned} \text{Performance (effectiveness)} = & \alpha + \beta * \text{Non-financial_weights} + Z_1 * \text{M\&A} + \\ & Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \\ & \text{Ind_ROA} + Z_7 * \text{Industry_volatility} + \text{error} \end{aligned} \quad (2c)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Non-financial_weights} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 * \text{Industry_volatility} + \text{error} \quad (2d)$$

With the inclusion of control variables, neither effectiveness nor efficiency is significant, indicating potential confounding effects of control variables (Table 6.12). Effectiveness is positively associated with financial health (higher z-score) and product life cycle while it is negatively related to product development cycle and industry volatility. Efficiency only shows a significant relationship with financial health.

Table 6.12: Non-financial weights and firm performance (Sub-sample with control variables)

		<u>Expected sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness	Non-financial weights	+	1.477	1.21	0.231
	M&A	+	0.339	0.83	0.412
	New CEO	-	-0.615	-1.39	0.172
	Bankruptcy Z-score	+	0.424	2.69	0.010***
	Product develop cycle	-	-1.716	-2.65	0.011**
	Product life cycle	+	1.574	2.40	0.020**
	Industry ROA	+	-0.068	-1.13	0.265
	Industry volatility	-	-0.349	-5.10	0.000***
	Intercept	+/-	-0.246	-0.36	0.723
Efficiency	Non-financial weights	+	1.551	1.11	0.272
	M&A	+	0.094	0.20	0.842
	New CEO	-	-0.721	-1.41	0.164
	Bankruptcy score	+	0.412	2.28	0.027**
	Product develop cycle	-	-0.644	-0.86	0.392
	Product life cycle	+	-0.276	-0.37	0.715
	Industry ROA	+	0.009	0.13	0.894
	Industry volatility	-	-0.129	-1.65	0.105
	Intercept	+/-	0.832	1.05	0.300
Effectiveness: F-test = 0.0077*** VIF = 1.49 R ² = 0.3304 Adj. R ² = 0.2211					
Efficiency: F-test = 0.0000* VIF = 2.17 R ² = 0.5399 Adj. R ² = 0.4647					
No. of observations = 58 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

It is again noteworthy that bankruptcy z-score and industry volatility are significant at the 1% level. Further action was taken to scrutinize their intervening effects. The sub-sample was divided into two equal groups. The results show that no regression is significant for the associations between non-financial weights and firm performance (Appendix 10-Panel A). Hence, the confounding effects of control variables likely are not a major concern in this study.

2.4 Additional Tests on Individual Measures

Taking each indicator as a dependent variable, 23 separate regressions were also run, excluding control variables. Appendix 9 presents the performance results of the weighted non-financial measures. The relative weights placed on non-financial components of BSC appear to have some positive effects on firm performance, as most performance indicators have expected signs. However, only ROA, price-to-cash flow, and total assets turnover are significant at the conventional levels. Compared to Model 2 of Chapter V, the association between non-financial weights and firm performance is much weaker than using the BSC measures.

The regression results with control variables are reported in Appendix 10. In Panel A-effectiveness, the positive coefficients of the indicators indicate that firms which place more weights on non-financial measures are possibly associated with higher accounting and market returns, but none of them is significant at the conventional levels. There is no significant relationship between non-financial measures weights and M&A. In contrast,

firms with a new CEOs tend to be negatively associated with ROA and stock return. As expected, a higher z-score is significantly and positively related to ROA, stock return, P/E ratio and price-to-cash flow ratio. Firms with short product development cycle and long product life cycle tend to outperform in accounting returns. The noise, as shown in industry volatility, has a widespread negative impact to both accounting and market performance. The coefficient of industry ROA is significant only for stock return.

In Panel B-efficiency, the relative weight of non-financial measures exhibits positive but insignificant associations with revenue per employee and net income per employee. The results provide no clear support that firms with higher weights on non-financial measures grow faster, borrow less, or make greater use of assets, as most indicators are not significant, except for total assets turnover. Regarding control variables, M&A shows a negative association with receivables turnover. New CEOs appear to have a positive effect on efficiency, as evidenced in shorter operating cycle and higher quick ratio. Healthier firms exhibit shorter operating cycles, higher quick ratio and interest coverage, and lower D/E ratio. Product development cycle is only significant for quick ratio. Long product life cycle may increase receivables turnover and reduce operating cycle. Industry ROA is only significant for total assets turnover. Once again, volatility dampens firm performance, as exhibited in the significant and negative coefficients of net income per employee, sales growth rate, and EPS growth rate.

Using the relative weights of non-financial measures steadily increases the explanatory power of the models, as demonstrated in the significantly higher R^2 statistics, compared

with those in Model 2 of Chapter V. The scores of R^2 vary from 0.10 to 0.60. F-tests for most regressions are significant. Taken together, the models seem to be well designed.

3. Model 3—Contingent Fit and Firm Performance

This section investigates whether firm performance is associated with the fit between non-financial weights and firm characteristics. Similar to Model 3 in Chapter V, the fit is captured in the predicted values in the regression of Model 1-Non-financial weights and contingent factors.

3.1 OLS Regressions for Fit

The hypotheses H3A and H3B were re-tested using the following OLS regression models:

$$\text{Performance (effectiveness)} = \alpha + \beta * \text{Non-financial_fit} + \text{error} \quad (3a)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Non-financial_fit} + \text{error} \quad (3b)$$

Table 6.13: Non-financial fit and firm performance (sub-sample)

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness	Non-financial fit	+	7.17	3.50	0.001***
	Intercept	+/-	3.49	3.41	0.001***
Efficiency	Non-financial fit	+	2.17	1.75	0.085*
	Intercept	+/-	-1.05	-1.71	0.093*
Effectiveness: F-test = 0.0009*** VIF = 1.22 $R^2 = 0.1794$ Adj. $R^2 = 0.1648$					
Efficiency: F-test = 0.0851* VIF = 1.05 = 0.0520 Adj. $R^2 = 0.0351$					
No. of observations = 58 for all variables					
*** 1% significance level		** 5% significance level		* 10% significance level	

Table 6.13 provides evidence that firm performance is a function of the fit of non-financial measures. Both regressions of effectiveness and efficiency are positive and significant at 1% and 10% respectively.

The confounding effects of control variables were checked using the following equations:

$$\text{Performance (effectiveness)} = \alpha + \beta * \text{Non-financial_fit} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 * \text{Industry_volatility} + \text{error} \quad (3c)$$

$$\text{Performance (efficiency)} = \alpha + \beta * \text{Non-financial_fit} + Z_1 * \text{M\&A} + Z_2 * \text{New_CEO} + Z_3 * \text{Bankruptcy} + Z_4 * \text{Prod_dev_cycle} + Z_5 * \text{Prod_life_cycle} + Z_6 * \text{Ind_ROA} + Z_7 * \text{Industry_volatility} + \text{error} \quad (3d)$$

Table 6.14: Non-financial fit and firm performance (sub-sample with control variables)

		<u>Predicted sign</u>	<u>Coef.</u>	<u>t-stat</u>	<u>P-value</u>
Effectiveness	Non-financial fit	+	2.41	1.01	0.318
	M&A	+	0.43	1.04	0.304
	New CEO	-	-0.49	-1.12	0.267
	Bankruptcy z-score	+	0.28	1.74	0.088*
	Product develop cycle	-	-1.02	-1.35	0.182
	Product life cycle	+	1.19	1.70	0.095*
	Industry ROA	+	-0.065	-1.08	0.287
	Industry volatility	-	-0.33	-4.88	0.000***
	Intercept	+/-	1.61	1.27	0.212
Efficiency	Non-financial fit	+	0.40	0.23	0.817
	M&A	+	-0.17	-0.57	0.574
	New CEO	-	0.26	0.81	0.424
	Bankruptcy score	+	0.28	2.36	0.022**
	Product develop cycle	-	-0.10	-0.19	0.851
	Product life cycle	+	0.993	1.98	0.054*
	Industry ROA	+	-0.001	-0.02	0.981
	Industry volatility	-	0.03	0.66	0.509
	Intercept	+/-	-0.29	-0.32	0.748
H2A: F-test = 0.000*** VIF = 2.15 $R^2 = 0.5357$ adj. $R^2 = 0.4599$					
H2B: F-test = 0.0850* VIF = 1.31 $R^2 = 0.2347$ adj. $R^2 = 0.1098$					
No. of observations = 58 for all variables					
*** 1% significance level ** 5% significance level * 10% significance level					

As compared with the results with control variables, Table 6.14 shows that non-financial fit becomes insignificant for both regressions with the inclusion of control variables, revealing that control variables may interfere with firm performance. In particular, bankruptcy z-score, product life cycle and industry volatility show significant relationship with firm performance.

Again, bankruptcy z-score and industry volatility were singled out for further tests. There is no difference between the significance of the two groups, except for financial health regressions, implying a merely weak confounding effect (Appendix 10-Panel B).

3.2 Additional Tests for Individual Measures

The results, using individual indicators as dependent variables, are displayed in Appendix 11. In Panel A, ROA, ROE, ROI, net profit margin, and price-to-cash-flow have positive and significant relationships with the fit of non-financial weights, consistent with expectations. In Panel B, only total assets turnover is significant.

The results are mixed, with the inclusion of control variables (Appendix 12). In Panel A-effectiveness, only the coefficients of ROI and price-to-cash-flow ratio are significant. In Panel B-efficiency, none of the indicators is significant. Yet, the results appear to provide some evidence that firm performance may be affected by the fit between the weights of

non-financial measures and the firm's contingent characteristics, since most indicators report expected signs. Most control variables are insignificant.

The models appear well designed, as the majority of F-tests are significant. Also, using the relative weights of non-financial measures significantly increases the explanatory power of the models, with the highest R^2 being 0.597. The VIFs range from 1.09 to 2.48, showing no obvious multicollinearity.

Overall, the results provide only limited support for the proposition that corporate performance is positively associated with the fit of non-financial weights. The results imply that the relative weights placed on non-financial measures of BSC only work if they are used in an optimal way, depending on individual firms. Over-relying or under-relying on non-financial measures do not help firm performance.

4. Summary

This chapter finds that the relative weights placed on non-financial components of BSC are significantly and positively associated with certain contingent variables including strategy, certain industries, size and Anglo culture. It provides weak support for the association between firm performance and 1) relative weights placed on non-financial measures and 2) the fitted weights of non-financial measures.

VII. Discussions

Although BSC has been widely advocated and adopted, empirical research has provided little evidence on its usage in executive compensation plans. This project tested three hypotheses: 1) the usage of BSC in compensation plans is contingent on firm characteristics, 2) firm performance is positively associated with the usage of BSC, and 3) performance is positively associated with the fit between BSC and firm characteristics. The empirical results show that the findings of this study, in general, are consistent with the hypotheses and previous work in the area.

1. Analysis of Findings

Full sample. The literature argues that the design and use of performance measurement depends upon organizational and environmental contexts (Otley 1980; Chenhall, 2003; Hansen et al., 2004). This study provides evidence to support contingency theory propositions, specifically with respect to the adoption of BSC in executive compensation contracts:

- Consistent with previous studies (Simons, 1990; Langfield-Smith 1997; Govindarajan and Gupta 1985; Simons 1990; Chenhall, 2003; Jermias and Gani, 2005), the results exhibit that firms pursuing an innovation-oriented prospector strategy put more emphasis on BSC in their annual bonus contracts.
- Ely (1991) and Speckbacher et al. (2003) found that the choice of performance measures varies by industry. The results confirm that BSC is significantly and

positively associated with firms in the financial and utility industries, indicating that BSC measure must be tailored to reflect industry-specific values and surrounding environments. Previous studies suffer a major drawback as they usually exclude the financial industry.

- This study reports a positive association between BSC and firm size. Murphy (1985) showed that firm size yielded information relevant for determining levels of managerial effort. Hoque and James (2000) and Chenhall (2003) argued that performance measurement systems tend to be more specialized in larger firms. This study finds that the number of employees may not be a good indication of firm size, and probably that is why no prior studies have used it as a measure of size.
- Firms emphasizing quality management are more likely to adopt BSC. This result is consistent with previous research that has found a positive relationship between the use of quality management practices and the relative importance of non-financial measures in reward systems (Ittner et al., 1997; Said et al., 2003). These findings support the notion that firms attempt to ensure that management incentives and organizational objectives are aligned.
- Structural arrangements remain an important factor in understanding the design of BSC (Anthony and Govindarajan, 2007; Chenhall, 2003; Waterhouse and Tiessen, 1978). Firms with more sophisticated structure significantly correlate with BSC usage in executive compensation.
- Compared to other contextual variables, previous research into culture is still exploratory in executive performance management. Yet, national culture remains

an important factor in multinationals' use of compensation practices (Schuler and Rogovsky, 1998; Bourguignon et al., 2004). Chhokar et al. (2008), Carr and Tomkins (1998), and Chow et al. (2002) showed differences in organizational cultures. As one of the first studies, this thesis suggests that BSCs tend to be adopted by firms in the Anglo cultural cluster.

- This study also supports the hypothesis that more concentrated institutional ownership provides motivation for managers to adopt BSC. Institutional owners that hold large blocks of shares have both the incentives and voting power to monitor management (Ashbaugh-Skaife et al., 2007; Jermias and Gani, 2005).

Ideally, the performance measures selected for management compensation contracts are measures that best motivate management to maximize the value of the firm. It has been broadly accepted that management compensation should be tied to BSC measures, (Kaplan and Norton, 2001; Banker et al., 2000). Although one cannot conclude from this research that BSC-based performance measures “cause” managers to make decisions that are valued by shareholders, a strong association is found in this study between the adoption of BSC and firm performance. Relative to control firms, treatment firms report higher effectiveness and efficiency. This conclusion is supported by prior research. Hoque and James (2000) reported that usage of BSC was significantly correlated with organizational performance. Davis and Albright (2004) found that a group of branches that used BSC outperformed a group that did not use BSC.

A match between BSC and contextual variables has been postulated to result in increased corporate performance, while poor fit implies diminished performance (Chenhall and Langfield-Smith, 1998; Waterhouse and Tiessen, 1978; Fisher, 1998; Hussain and Hoque, 2002; Sila, 2007). Jermias and Gani (2005) found that the contingent-fit was positively related to business-unit performance. Yet, there are few empirical tests on how to measure fit and how this fit affects a firm's performance. HassabElnaby et al. (2005) recently used fitted factors to approximate the vague fit. This thesis extends this method by investigating the interfaces between the fit of BSC and its performance consequences. Consistent with expectations, the results indicate that firm performance is an increasing function of the level of fit of BSC and its surrounding factors.

Sub-sample. For the first time, this study also examined the relationships between the relative weights placed on non-financial measures and 1) firm characteristics, 2) firm performance, and 3) the concept of fit. The results suggest that firm characteristics explain the reliance on non-financial performance measures in compensation contracts, in agreement with Govindarajan and Gupta (1985), Ittner et al. (1997) and Simons (1990). The findings are similar to those using BSC instead of non-financial weights. The prospector firms and larger firms rely more heavily on the use of non-financial measures. Firms operating in financial and utility industries and in Anglo cultural cluster are more likely to focus on non-financial measures. The results regarding the connection between the relative weights of non-financial measures and firm performance are positive, but are relatively weaker than using the BSC dummy variable measure.

2. Major Contributions

While the work of Kaplan and Norton has added to the knowledge of how BSC can provide managers with an integrative framework to manage organizational activities, little empirical evidence is available on current practice. Its benefits have been asserted largely through anecdotal, self-reported survey data, and case study evidence. No large-scale studies have documented the prevalence of BSC in executive bonus contracts and its impacts. Accordingly, it is not surprising that many researchers question the inherent value of BSC system. This study is an attempt to contribute to the limited body of knowledge in this area. It provides a solid, large-scale empirical analysis with benefits for both researchers and practitioners by addressing several unanswered questions.

This research is most closely related with the studies of Ittner et al. (1997) and Said et al. (2003). Yet, their studies examine only non-financial measures in compensation contracts and their performance consequences. This thesis expands these two studies by including BSC that comprises both financial and non-financial measures, and provides valuable and concrete evidence using multiple measures for key variables. This thesis consists of a thorough analysis of the empirical data collected from public sources and thus avoids common pitfalls of traditional survey research. In particular, this study enriches the existing knowledge of performance measurement and BSC by making significant contributions to researchers and practitioners:

Implications for researchers. To my best knowledge, this is the first study that examines BSC compensation plans in a broader theoretical context. The concept of fit has been used in management accounting literature to investigate contingent relationships between performance measures, contextual variables, and organizational performance. A number of major attempts have been made to probe if and how BSC is affected by contingent variables. This study provides concrete evidence regarding the validity of contingency theory. The results indicate that the association between BSC measures and firm performance is contingent on whether the use of those measures fits the firm's specific characteristics.

Furthermore, by replacing the BSC with the relative weights assigned to non-financial measures of BSC, this study re-examines the hypotheses using alternate measures. Said et al. (2003) reported that firms that employ a combination of financial and non-financial measures had higher mean levels of ROA and stock return, and that performance is contingent on several firm characteristics. Yet, a limitation of their study is that they considered the use of non-financial measures in aggregate terms only—1 for using non-financial measures and zero otherwise (dummy variable). It may be the case that the relative weights of non-financial measures affect performance, contingent on certain situations. Accordingly, this study, by using the sub-sample of 58 firms, addresses this unanswered question by investigating the relative weights of non-financial measures and their relationships with contingent factors and firm performance. By replacing a nominal scale (BSC dummy) with a ratio scale (non-financial weights), the robustness of the model also increases, as shown in a much higher squared multiple correlations.

In addition to the traditional descriptive, univariate and regression analyses, this thesis used a variety of data collection and analysis methods. A matched pair of control firms improves the validity of the results by controlling for the intervening variables. This design decreases validity threats, since both treatment and control firms are subject to common omitted variables and primarily differ based on the partitioning variable. Further, unlike survey data, this study is based on more objective public datasets. The identification of each firm as a BSC adopter is based on publicly available information. Unlike prior research (Ittner et al., 1997; Said et al., 2003), this study uses multiple measures to represent each major contingent, performance and control variable, which improves measurement reliability. Said et al. (2003), for example, used only ROA and stock return to measure firm performance. In this study, both accounting- and market-based metrics are used to measure organizational performance. Some of the variables, such as structure and culture, are introduced into an empirical study on this topic for the first time. This study also examines firm performance in regressions with several control variables. It can reveal whether the addition of control variables makes a significant contribution to the predictive model. Both parametric and non-parametric tests are employed to examine the hypotheses. Non-parametric tests are particularly useful if the variables are ordinal or nominal, the distribution is clearly not normal, or the data have extreme values.

Implications for practitioners. This study provides empirical evidence on executive incentives. Despite an increasing usage of BSC measures in managerial compensation

plans (Ittner et al., 1997), virtually no empirical evidence of performance impacts of such plans exists. The results of this study demonstrate that the installation of BSC and its fit are associated with enhanced corporate performance. Shareholders and managers may seek to improve the alignment between their interests by implementing a better incentive package. Also, this study differs significantly from other works published in the area in that the market effect of adopting BSC is assessed.

Yet, there is no clear support that a firm's performance is associated with the relative weights placed on non-financial measures. Over-relying or under-relying on non-financial measures (or financial measures) does not necessarily lead to better performance. The findings should, at a minimum, caution organizations currently considering the adoption of BSC to carefully consider the relative weight of each component of BSC and the cost/benefit aspects of designing measurement systems.

3. Ex-post Model Development

The priori model proposed in Chapter III appears well designed as the results in general are consistent with expectations and with prior research. Even so, with new findings from the study, the model still has room for improvement in future studies. These issues are discussed in this section and presented in Figure 7.1.

The contingent hypotheses. All of the contingent variables have significant relationships, which support the hypotheses of Model 1 from H1A to H1G, as discussed in Section 1.

Strategy, industry, size, quality, structure, culture and ownership are valid components of contingent context in the model. As these factors account for only part of the variance in the model, however, other contingent factors not examined in the study may exist, as the analysis using additional control variables demonstrated.

Figure 7.1 Ex-post Model

The performance hypotheses. The results are consistent with Hypotheses H2A and H2B, indicating that effectiveness and efficiency are valid dimensions of firm performance. Yet, one general observation is that effectiveness appears to be a better performance indicator than efficiency, as the results for the former are more significant. Probably this

is why previous studies prefer the former over the latter. Also, it is striking that market-based measures tend to be more significant than accounting-based measures. The strong evidence suggests that market participants respond favourably to the adoption of BSC-based compensation plans. Future studies could delineate accounting- and market-based performance instead of effectiveness. Yet, using financial returns alone, which is very common in the existing literature, does not draw a full picture of firm performance, because non-financial measures also provide value for a firm's internal management. Similar to effectiveness, the variable "efficiency" could be decomposed further into productivity, asset utilization, growth, and leverage.

Table 7.1 Measurement Levels for Firm Performance

<u>Level 1</u>	<u>Level 2</u>	<u>Level 3</u>
Effectiveness	Accounting return	ROA
		ROE
		ROI
	Market return	Stock return
		P/E ratio
		Dividend yield
		Price/cash flow ratio
	Profit Margin	Gross profit margin
		EBIT margin
		Net profit margin
		Dividend payout
Efficiency	Productivity	Revenue per employee
		Cost per employee
		Net income per employee
	Growth	Sales growth rate
		EPS growth rate
	Asset utilization	Inventory turnover
		Receivables turnover
		Total asset turnover
		Operating cycle
	Leverage	Quick ratio
		Interest coverage
		D/E ratio

Table 7.1 shows the new grouping consisting of three levels. Since Level 1 might be somewhat vague and Level 3 unmanageable and fragmented, Level 2 could provide valid measures as well as capture firm performance in a manageable manner.

The confounding effects. Firm performance could be confounded by factors other than the usage of BSC. Hence, this study compared the regressions including control variables with those excluding control variables to ensure the robustness of the study:

- For the full sample, the coefficients of BSC in both tests are significant with predicted signs, indicating that the effects of control variables are not strong enough to disturb firm performance. However, certain control variables are significant at conventional levels. M&A and bankruptcy z-score are positive and significant, consistent with the assumptions that firms engaged in M&A and with higher financial health tend to achieve higher returns. Product development cycle and industry volatility are negatively associated with firm performance.
- For the sub-sample, compared to the regressions without control variables, the coefficients of non-financial weights with control variables become insignificant, revealing that the intervening effects of control variables on firm performance are more likely. Firm performance is positively associated with bankruptcy z-score and product life cycle, while it is negatively related to product development cycle and industry volatility. However, these findings could relate to the small sub-sample size.

- The confounding effects of other variables, including new CEO, product life cycle and industry ROA, are much weaker, and do not pose major validity concerns for the study.

As a result, the revised model would retain M&A, bankruptcy, product development cycle and industry volatility, while removing new CEO, product life cycle and industry ROA. However, additional control variables not identified in this study still remain possible.

The fit hypothesis. Recent studies have proposed that the contingent-fit of BSC with its contextual variables is likely to have positive effect on firm performance. This presumption is confirmed in H3A and H3B, implying that the measurement of fit has successfully captured the congruence between BSC and the set of contingent variables (strategy, industry, size, quality, structure, culture and ownership).

The financial vs. non-financial argument. The BSC was developed as a consequence of the intense argument regarding the merits of financial-versus-non-financial measures. Compensation plans based solely on profits and other financial accounting measures have been criticized for encouraging managers to sacrifice long-run performance to increase short-term financial results, and thereby to maximize their bonuses. To overcome the short-run orientation of accounting-based reward systems, many firms are supplementing financial metrics with non-financial measures in order to assess performance dimensions that are not captured in short-term financial results (Ittner et al., 2003b). The BSC

approach has gained prominence in management research as a way of systematically integrating financial and non-financial performance measures. BSC views organizational performance from four dimensions: financial (or shareholders), customers, internal business processes, and learning and growth. Atkinson et al. (1997, pp. 93-94) suggested that BSC has the potential to provide planners with a way of expressing and testing a sophisticated model of cause-and-effect relationships in the organization—a model that provides managers with a basis on which to manage the drivers of desired outcomes. This study supports this assertion.

4. Methodological Considerations

Extensive efforts have been made to ensure the reliability and validity of the study from data collection, variable measurement and statistical tests to reporting results. Yet, the research methods could be improved with the following considerations for future studies:

Factor analysis. Principal component analysis used in this study is a mathematical procedure that transforms a number of possibly correlated indicators into a smaller number of uncorrelated variables (Diekhoff, 1992). It is sometimes confused with factor analysis because there are many important similarities between the two procedures: both are variable reduction methods that can be used to identify groups of observed variables that tend to “hang together” empirically. And in most cases, these two methods usually yield very similar results. Nonetheless, there are some important conceptual differences between them (Stevens, 1986). Perhaps the most important deals with the assumption of

an underlying causal structure: the researchers use factor analysis when they believe that certain latent variables (factors) exert causal influence on the observed variables they are studying. In contrast, PCA makes no assumption about an underlying causal model. It is simply a variable reduction procedure that results in a small number of components that account for most of the variance in a set of observed variables. Instead of principal component analysis, future research could use factor analysis to develop latent constructs for contingent variables (strategy, ownership, etc.) and firm performance (effectiveness and efficiency).

Multicollinearity and sample size. Researchers usually prefer a large size of sample. This is of vital importance in parametric statistical inference because the probability of an observed sample will be used to infer the characteristics of the population (McClave et al., 2001). The observations used in this study constitute a large representative sample, which should exhibit characteristics typical of those possessed by the population of interest. However, using large samples would result in more significant correlations among independent variables. A correlation that approaches zero could be statistically significant, if it is based on a large enough sample (Heiman, 1992; Diekhoff, 1992). Multicollinearity in such cases should be interpreted with caution, and a balance is needed between statistical significance and substantial significance, i.e., meaningful theory-based relationships.

Stepwise regression. Multiple regression analysis entails a combination of two or more independent variables predicting the value of a single dependent variable. A direct

solution, as is the case in this study, includes all available predictors in the regression equation and is particularly useful when the researcher wishes to quantify the relationships between an entire set of predictor variables and the dependent variable. Another procedure, called stepwise regression, builds one step at a time by sequentially adding or dropping predictors to the equation, based on theoretical grounds. A compelling advantage of this method is that it enables the researcher to include only those variables that add significant predictive power and can mitigate the problem of multicollinearity (Diekhoff, 1992). Future research can use this technique to investigate which contingent variables predict best the usage of BSC in executive compensation plans.

Structure equation modeling. Tests of fit as mediation can also be carried out using structural equation modeling (SEM), a technique that is also known as path analysis with latent variables. The relative proportion of the two effects – indirect versus direct – provides an index of the relative effect of fit (namely, indirect effects) versus the direct effects (Venkatraman, 1989). SEM is a method for representing dependency (arguably “causal”) relations in multivariate data in business, behavioural and social sciences (McDonald and Moon-Ho, 2002). Generally, this model consists of two main parts, the measurement model and the path model. The former represents a set of observable variables as multiple indicators of a smaller set of latent variables, which are usually common factors. The latter describes relations of dependency—usually accepted to be in some sense causal—between endogenous and exogenous latent variables that cannot be

directly observed but that must be inferred by means of manifest indicator variables, i.e., variables measured without error (Thomas, Lu and Cedzynski, 2006).

SEM techniques, which are based on fitting the model-implied covariance matrix of the manifest variables to the empirically determined covariance matrix, are typically implemented through software packages, such as LISREL. The covariance-based approach in SEM is of particular interest, because it explicitly models the measurement error associated with each latent variable and therefore ensures that estimates of the SEM parameters are consistent. By explicitly modeling measurement error, SEM users seek to derive unbiased estimates for the relationships between latent constructs. As compared to regression and factor analysis, SEM is a relatively young analytical tool. As such, the methodology is still developing, and even fundamental concepts are subject to challenge and revision.

5. Summary

This chapter analyzes the empirical results, comparing them to existing literature of performance measurement. This study has implications for both researchers and managers. The research model appears well designed, as the hypotheses are consistent with the expectations and the tests are robust.

VIII. Conclusions and Future Directions

This chapter concludes the study with key points, limitations and future directions.

1. Summary

The Balanced Scorecard (BSC) has been argued to be a superior combination of non-financial and financial measures of performance. The “real power” of this concept can be used only if BSC measures are linked to reward systems. Despite an increasing usage of BSC, however, virtually no large-scale empirical studies have documented the prevalence of BSC in executive compensation plans. This thesis attempts to address this deficiency. Three primary questions are central to the study: Do contextual factors affect the usage of BSC in compensation plans? Does BSC have an impact on organizational performance? Is this performance associated with the fit between BSC and firm characteristics?

The hypotheses are tested using a three-stage regression model including logistic and OLS regressions, based on a sample of 165 firms that have adopted BSC and 165 control firms over the period of 1992–2009.

The results underscore the importance of considering firm characteristics prior to adopting BSC. This study finds that 1) the usage of BSC in executive compensation plans is related to strategy, certain industries, size, quality, structure, Anglo culture and ownership; 2) there is a strong association between firm performance and the adoption of

BSC; and 2) firm performance is an increasing function of the level of fit between BSC and its contingent factors.

In addition, this study provides evidence that firm characteristics explain the relative weights placed on non-financial components of BSC. Yet, it provides only limited evidence regarding the relationships between firm performance and the non-financial weights and the fit.

This research enriches the existing knowledge of performance measurement by making contributions to limited literature. As one of the first empirical studies that systematically examines BSC compensation plans in the context of contingency theory, this thesis provides solid evidence on the validity of this theory. It also suggests that firms should seek to improve the alignment between BSC incentive packages and firm characteristics to ensure better performance. Over-relying or under-relying on the relative weights placed on financial or non-financial measures may not help firm performance.

2. Limitations

This study is by no means perfect. The results require careful interpretation because they may be influenced by selection bias, as the sample does not constitute a random sample. The firms also chose whether to use and whether to disclose the use of BSC and its weighted components, whether to link them to managers' incentives and determine their own corporate governance structures. The treatment firms may be ones whose board or

senior management determined that certain problems existed and needed new performance measurement models. It is also possible that only successful treatment firms were selected because they promote their compensation plans in proxy statements.

This study has assumed a causal relationship between the adoption of BSC and corporate performance, as in the cases of Kaplan and Norton (1993, 1996), Ittner et al. (1997), Ittner et al. (2003b), Said et al. (2003), and HassabElnaby et al. (2005). Yet, reverse causality from performance to BSC is possible, since it may be the case that only those firms with better performance had chosen to use BSC in their executive incentive programs.

Another limitation is the exclusive reliance on the use of BSC in compensation, as opposed to other potential uses of BSC, which this study does not address. Further, this study does not consider the potential and likely impact of target setting with regard to both financial and non-financial performance measures, i.e., variances from target measures used in executive compensation and how.

A further constraint is the relatively small size of the full sample, as thousands of public firms are listed in the U.S. stock exchanges and the results of this study imply that all firms could benefit from BSC. Only 185 companies that explicitly implemented BSC in compensation plans were identified from their proxy reports. It is possible that many firms that used BSC did not mention this measurement scheme, or they used other

variations of BSC, a combination of both financial and non-financial measures, without naming “BSC”.

Regarding the sum-sample, this study only considers the relative weights of non-financial measures in aggregation. It may be the case that a specific non-financial component (e.g., cost control, quality, customer satisfaction, etc.) affects performance in a different fashion than an alternate non-financial measure. Ittner and Larcker (2003), for example, found variation in the specific factors influencing the use of different types of non-financial measures, suggesting that the aggregate performance measure classification commonly used in compensation research can provide somewhat misleading inferences regarding performance measurement choices. Similarly, the performance implications associated with the use of different non-financial measures in different circumstances are also likely to vary.

Although this thesis has tested whether the adoption of BSC-based compensation plans is reflected in increased shareholder wealth, tests of changes in shareholder wealth bear some caveats. It is difficult to determine exactly when the market first learns of a firm’s change in its incentive compensation plan and therefore when this information is impounded in the firm’s share price. Because of this uncertainty, it is not possible to use a traditional short-window event study methodology. Instead, this study looks at firms’ annual stock return, P/E ratio, dividend yield, and price/cash flow ratio, which are a combination of stock price and accounting measures. Regarding the interpretation of the results, as pointed by Wallace (1997), any change in shareholder wealth can reasonably

be attributed to a multitude of causes. In the present case, a change in behaviour associated with the adoption of BSC-based compensation is one possible cause of increases in shareholder wealth. Another plausible explanation is that the adoption of BSC measure signals further policy changes. The adoption can also signal management's private information regarding future firm performance.

Finally, even though this study has incorporated seven variables as a control, such as M&A and new CEO, there may be other variables that intervene with corporate performance besides the adoption of BSC-based compensation.

3. Future directions

This thesis has attempted to address many unexplored issues regarding BSC compensation policy; however, many fruitful research avenues remain, as we are still left with many important topics to resolve and challenges to overcome.

There is an absence of empirical evidence that identifies and examines why some firms discard the use of BSC measures after adopting them, even though BSC demonstrates benefits to shareholders. Future studies can investigate the firm's decision to retain or discard the use of BSC measures included in compensation contracts. In particular, the following questions are interesting to researchers: Why do firms that had adopted BSC at one point subsequently decided to discard them from their compensation contracts? Does

poor performance or the mismatch of BSC measures with firm characteristics play a role in that decision?

The relative merits of stock price versus accounting numbers as measures of performance continue to be debated, and we have witnessed a tremendous increase in the use of stock-based compensation. Hence, a fruitful venue is stock market reaction to the adoption of BSC-based compensation plans. The indicators used in this study, including P/E ratio, MV/BV ratio, price/cash flow ratio, and dividend yield, are a combination of accounting- and market-based measures, and accordingly the market reaction is moderated by the accounting values in denominators. Future studies have the opportunity to investigate excess returns using the classic “event study” method, which typically tries to examine return behaviour for a sample of firms experiencing a common type of event (e.g., the adoption of BSC, stock split, etc). Such an exercise can be conducted in many ways. One could, for example, examine the entire distribution of abnormal returns. This is equivalent to comparing the distributions of actual returns with the distribution of predicted returns and asking whether the distributions are the same. Typically, the specific null hypothesis to be tested is whether the mean abnormal return at time t is equal to zero. The focus on mean effects makes sense if one wants to understand whether a BSC event is, on average, associated with a change in security holder wealth.

In general, compensation policy is an endogenous process, arising simultaneously with corporate investing, financing and operating decisions. Is the driving force behind all of these policy decisions the nature of the firm's opportunity or the adoption of BSC?

Continued progress in understanding the causal relation between BSC compensation policy and other corporate policy decisions requires careful consideration.

Research is also needed to capture different dimensions of the contextual impact on firm performance. A host of implementation issues concerning the adoption of BSC has not been addressed in the literature. Although prior survey research has identified several contingent factors, this study is limited to the measures for which data were publicly available.

This study indicates that the use of BSC performance measures is positively associated with performance contemporaneously, which is consistent with prior studies (Chenhall, 2005; Said et al. 2003). A natural extended question is: will the usage of BSC bring up firm performance prospectively?

Finally, a combination of survey and public data can potentially provide a richer understanding of the impact of BSC measures on firm performance. This study has found that effectiveness measures appear more significant and important than efficiency measures. The former tends to be publicly available and come from external sources whereas the latter are usually related to internal management, such as productivity, asset utilization, and leverage, and often have incomplete records. The market-based performance is determined by collective forces, unlike efficiency that is easier to measure and are subject to management manipulations. Future research has the potential to draw

from the relative strengths of each method by collecting both internal and external data for similar study.

4. Concluding Remarks

A small, but growing number of firms are selecting BSC-based measures. Ideally, the performance measures selected for management compensation contracts are measures that best motivate management to maximize the value of the firm. Despite widespread practitioner interest in the subject, little research has been conducted on the implementation consequences of BSC. Kaplan and Norton (1996) do not provide specific recommendations regarding how to link BSC to compensation. The results of this research allow researchers and managers to predict which firms are likely to adopt BSC-based measures, and to test whether BSC leads to better performance.

The new century has witnessed a massive transformation in business organizations, which have been forced by intense competition and the globalization of markets to redefine their strategies, structures and processes. High-performance organizations, whether public or private, are and must be interested in developing and deploying effective and efficient performance measurement systems. It is only through such systems, that they can remain high-performance organizations. Consequently, BSC has received considerable attention in academia and industry. In order for research to continue to make substantive managerial and research contributions, this study provides significant empirical evidence and suggests future research directions.

References

- Abernethy, M.A. and Chua, W.F., "A field study of control system 'redesign': the impact of institutional processes on strategic choice", *Contemporary Accounting Research*, 13, 1996, 569-606.
- Abernethy, M.A. and Stoelwinder, J.U., "The role of professional control in the management of complex organizations", *Accounting, Organizations and Society*, 20, 1995, 1-17.
- Agresti, A., *Introduction to categorical data analysis*, Wiley, New York, 2002.
- Anthony, R. N., "Planning and control systems: a framework for analysis", Harvard Business School Division of Research, 1965.
- Anthony, R.N. and Govindarajan, V., "Management control systems", 12th edition, McGraw-Hill/Irwin, Boston, 2007.
- Atkinson, A., Balakrishnan, R., Booth, P., Cote, M., Grout, T., Mali, T., Roberts, H., Ulan, E. and Wu, A., "New directions in management accounting research", *Journal of Management Accounting Research*, 9, 1997, 80-108.
- Ashbaugh-Skaife, H., Collins, D.W. and Kinney, W.R.Jr., "The discovery and reporting of internal control deficiencies prior to SOX-mandated audits", *Journal of Accounting and Economics*, 44, 2007, 166-192.
- Ax, C. and Bjørnenak, T., "Bundling and diffusion of management accounting innovations—the case of the balanced scorecard in Sweden", *Management Accounting Research*, 16, 2005, 1–20.
- Babbie, E. "The practice of social research", Wadsworth Publishing Company, Belmont, CA, 1983.
- Baker, G., Gibbons, R. and Murphy, K J., "Subjective performance measures in optimal incentive contracts", *Quarterly Journal of Economics*, 109 (4), 1994, 1125-1156.
- Baltagi, B.H., "Econometrics", 4th Edition, Springer, Berlin Heidelberg, 2008.
- Banker, R., Lee, S. and Potter, G., "A field study of the impact of a performance-based incentive plan", *Journal of Accounting and Economics*, 21, 1996, 195-226.
- Banker, R., Potter, G. and Srinivasan, D., "An empirical investigation of an incentive plan that includes non-financial performance measures", *The Accounting Review*, 75(1), 2000, 65-92.

- Barber, B.M. and Lyon, J.D., "Detecting abnormal operating performance: the empirical power and specification of test statistics", *Journal of Financial Economics*, 41, 1996, 359-399.
- Berry, A.J., Broadbent, J. and Otley, D., "Management control: theories, issues and performance", 2nd edition, Palgrave Macmillan, New York, 2005.
- Bititci, U.S., Turner, T. and Begemann, C., "Dynamics of Performance Measurement Systems", *International Journal of Operations and Production Management*, 20 (6), 2000, 692-704.
- Baltagi, B.H., *Econometrics*, 4th Edition, Springer, Berlin Heidelberg, 2008.
- Bordens, K.S. and Abbott, B.B., *Research design and methods: a process approach*, 3rd edition, Mayfield Publishing Company, Toronto ON, 1996.
- Bourguignon, A., Malleret, V. and Nørreklit, H., "The American balanced scorecard versus the French *tableau de bord*: the ideological dimension", *Management Accounting Research*, 15, 2004, 107-134.
- Brown, R., Jones, M. and Steele, T., "Still flicking at the margins of existence? Publishing patterns and themes in accounting and finance research over the last two decades", *The British Accounting Review*, 39, 2007, 125-151.
- Budde, J., "Performance Measure Congruity and the Balanced Scorecard", *Journal of Accounting Research*, 45(3), June 2007, 515-539.
- Bushman, R., Indjejikian, R. and Smith, A., "CEO compensation: the role of individual performance evaluation", *Journal of Accounting and Economics*, 21, April 1996, 161-193.
- Carr, C. and Tomkins, C., "Context, culture and the role of the finance function in strategic decisions, a comparative analysis of Britain, Germany, the U.S.A. and Japan", *Management Accounting Research*, 9, 1998, 213- 239.
- Chapman, C., "Reflections on a contingent view of accounting", *Accounting, Organizations and Society*, 22(2), 1997, 189-205.
- Chenhall, R.H., "Management control systems design within its organizational context: findings from contingency-based research and directions for the future", *Accounting, Organizations and Society*, 28(2/3), 2003, 127-168.
- Chenhall, R.H., "Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study", *Accounting, Organizations and Society*, 30(5), 2005, 395-422.

Chenhall, R. H. and Langfield-Smith, K., "Factors influencing the role of management accounting in the development of performance measures within organisational change programs", *Management Accounting Research*, 9, 1998, 361–386.

Chhokar, J.S., Brodbeck, F.C., and House, R. J., *Culture and leadership across the world: the GLOBE book of in-depth studies of 25 societies*, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 2008.

Chintagunta, P.K. and Prasad, A.R., "An empirical investigation of the 'dynamic McFadden' model of purchase timing and brand choice: implications for market structure", *Journal of Business and Economic Statistics*, 16(1), January 1998, 2-12.

Chow, C.W., Harrison, G.L., McKinnon, J.L. and Wu, A., "The organizational culture of public accounting firms: evidence from Taiwanese local and US affiliated firms", *Accounting, Organizations and Society*, 27, 2002, 347–360.

Cone, J.D. and Foster, S., *Dissertation and theses: from start to finish*, American Psychological Association, Washington DC, 1993.

Conyon, M. and Murphy, K., "The prince and the pauper? CEO pay in the US and UK", *Economic Journal*, 110, 2000, F640-F671.

Core, J. and Guay, W., "Estimating the value of employee stock option portfolios and their sensitivities to price and volatility", *Journal of Accounting Research*, 40 (3), 2002, 613-630.

Cycyota, C.S. and Harrison, D.A., "What (not) to expect when surveying executives: a meta-analysis of top manager response rates and techniques over time", *Organizational Research Methods*, 9 (2), April 2006, 133-160.

Davis, S. and Albright, T., "An investigation of the effect of balanced scorecard implementation on financial performance", *Management Accounting Research*, 15(2), 2004, 135-153.

DeMaris, A., "A tutorial in logistic regression", *Journal of Marriage and the Family*, 57 (4), November 1995, 956-968.

Diekhoff, G., *Statistics for the social and behavioral sciences: univariate, bivariate, multivariate*, Wm. C. Brown Publishers, Dubuque IA, 1992.

DiMaggio, P.J. and Powell, W.W., "The iron cage revisited: institutional isomorphism and collective rationality in organizational fields", *American Sociology Review*, 48 (2), 1983, 147-160.

Donaldson, L., *American anti-management theories of organization: critique of paradigm proliferation*, Cambridge University Press, Cambridge, UK, 1995.

- Eccles, R.C., "The Performance Measurement Manifesto", *Harvard Business Review*, January-February, 1991, 131-137.
- Ely, K., "Interindustry differences in the relation between compensation and firm performance", *Journal of Accounting Research*, 29 (1), 1991, 37-58.
- Fama, E.F. and French, K.R., "Multifactor explanation of asset pricing anomalies," *Journal of Finance*, 51, 1996, 55-84.
- Feltham, G., and Xie, J., "Performance measure congruity and diversity in multi-task principal/agent relations", *The Accounting Review*, 69 (3), 1994, 429-453.
- Fisher, J.G., "Contingency theory, management control systems and firm outcomes: past results and future directions", *Behavioural Research in Accounting*, 10 (Supplement), 1998, 47-64.
- Gaver, J.J. and Gaver, M., "The association between performance plan adoption and corporate capital investment: a note", *Journal of Management Accounting Research*, Fall 1993, 145-158.
- Govindarajan, V. and Gupta, A.K., "Linking control systems to business unit strategy: impact on performance", *Accounting, Organizations and Society*, 10, 1985, 51-66.
- Graham, J. W., Cumsille, P. E. and Elek-Fisk, E. (eds.), "Methods for handling missing data". In J. A. Schinka & W. F. Velicer, "Research Methods in Psychology", *Handbook of Psychology*, 2, John Wiley & Sons, New York, 2003, 87-114.
- Hall, B. and Liebman, J., "Are CEOs really paid like bureaucrats?" *Quarterly Journal of Economics*, 103, 1998, 653-691.
- Hansen, D.R., Mowen, M.M., Senkow, D.W. and Pollanen, R.M, "Chapter 17 – strategic performance evaluation and management", in *Management Accounting*, Canadian 6th Edition, Nelson Thompson, 2004.
- Hartmann, F.G. H. and Moers, F., "Testing contingency hypotheses in budgetary research: an evaluation of the use of moderated regression analysis," *Accounting, Organizations and Society*, 24(4), 1999, 291-315.
- HassabElnaby, H.R., Said, A.A. and Wier, B., "The retention of non-financial performance measures in compensation contracts", *Journal of Management Accounting Research*, 7, 2005, pp. 23-42.
- Heiman, G.W., *Basic statistics for the behavioral sciences*, Houghton Mifflin Company, Boston MA, 1992.

Hoque, Z. and James, W., "Linking balanced scorecard measures to size and market factors: impact on organizational performance", *Journal of Management Accounting Research*, 12, 2000, 1-17.

Hussain, M. and Hoque, Z., "Understanding non-financial performance measurement practices in Japanese banks", *Accounting, Auditing and Accountability Journal*, 15 (2), 2002, 162-183.

Ittner, C.D. and Larcker, D.F., "Innovations in performance measurement: trends and research implications", *Journal of Management Accounting Research*, 10, 1998, 205-238.

Ittner, C.D. and Larcker, D.F., "Assessing empirical research in managerial accounting: a value-based management", *Journal of Accounting and Economics*, 32, 2001, 349-410.

Ittner, C., and Larcker, D., "Coming up short on non-financial performance measurement", *Harvard Business Review*, November 2003, 88-95.

Ittner, C., Larcker, D. and Meyer, "Subjectivity and the weighting of performance measures: evidence from a balanced scorecard", *The Accounting Review*, 78(3), 2003a, 725-758.

Ittner, C., Larcker, D. and Rajan, M., "The choice of performance measures in annual bonus contracts", *The Accounting Review*, 72(2), 1997, 231-255.

Ittner, C., Larcker, D. and Randall, T. "Performance implications of strategic performance measurement in financial services firms", *Accounting, Organizations and Society*, 28, 2003b, 715-741.

Jensen, M. and Meckling, W., "Theory of the firm: managerial behaviour, agency costs and ownership structure", *Journal of Financial Economics*, 3, 1976, 305-360.

Jensen, M. and Murphy, K., "CEO incentives – it's not how much you pay, but how", *Harvard Business Review*, May-June 1990, 138-149.

Jermias, J. and Gani, L., "Ownership structure, contingent-fit, and business-unit performance: a research model and empirical evidence", *The International Journal of Accounting*, 40, 2005, 65-85.

Kaplan, R.S. and Norton, D.P., "The balanced scorecard: measures that drive performance", *Harvard Business Review*, January-February 1992, 71-79.

Kaplan, R.S. and Norton, D.P., "Putting the balanced scorecard to work", *Harvard Business Review*, September-October 1993, 134-147.

Kaplan, R.S. and Norton, D.P., "Using the balanced scorecard as a strategic management system", *Harvard Business Review*, January-February 1996, 75-85.

Kaplan, R.S. and Norton, D.P., *The strategy-focused organization*, Harvard Business School Publishing, Boston, MA, 2001.

Kennerley, M. and Neely, A., "A framework of factors affecting the evolution of performance measurement systems", *International Journal of Operations and Production Management*, 22 (11), 2002, 1222-1245.

Khotari, S.P. and Warner, J.B., "Econometrics of event studies". In B. Espen Eckbo (ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Elsevier/North-Holland, Chapter 1, 2007.

Kim, K. and Nofsinger, J., *Corporate governance*, Pearson/Prentice Hall, 2nd edition, 2006.

Lambert, R.A., "Contracting theory and accounting", *Journal of Accounting and Economics*, 32, 2001, 3-87.

Langfield-Smith, K., "Management control systems and strategy: a critical review", *Accounting, Organizations and Society*, 22 (2), 1997, 207-232.

Larcker, D.F., "The association between performance plan adoption and corporate capital investment", *Journal of Accounting and Economics*, 5, 1983, 3-30.

Lee, T.W., Mitchell, T. R. and Sablinski, C. J., "Qualitative research in organizational and vocational psychology, 1979-1999", *Journal of Vocational Behaviour*, 55, 1999, 161-187.

Lipe, M. and Salterio, S., "The balanced scorecard: judgmental effects of common and unique performance measures", *The Accounting Review*, 75 (3), 2000, 283-298.

Little, R.J.A. and Rubin, D.B., *Statistical analysis with missing data*, John Wiley & Sons, New York, 1987.

Matyas, L. and Sevestre, P. (eds.), *The econometrics of panel data*, Springer-Verlag Berlin Heidelberg, 2008.

McClave, J.T., Benson, P.G. and Sincich, T., *Statistics for business and economics*, 8th edition, Prentice Hall, 2001.

McConnell, J. J. and Muscarella, C.J., "Corporate capital expenditure decisions and the market value of the firm", *Journal of Financial Economics*, 14, 1985, 399-422.

McDonald, R.P. and Moon-Ho, R.H., "Principles and practice in reporting structural equation analyses", *Psychological Methods*, 7(1), 2002, 64-82.

- Milgrom, P. and Roberts, J., "Complementarities and fit strategy, structure, and organizational change in manufacturing", *Journal of Accounting and Economics*, 19, (2-3), March-May 1995, 179-208.
- Morgan, G. and Smircich, L., "The case for qualitative research". *Academy of Management Review*, 5(4), 1980, 491-500.
- Motulsky, H., *Intuitive biostatistics: choosing a statistical test*, Chapter 37, Oxford University Press, 1995.
- Murphy, K., "Corporate performance and managerial remuneration: an empirical analysis", *Journal of Accounting and Economics*, 7, 1985, 11-42.
- Murphy, K.J., "Executive compensation". In O. A. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, 3, North Holland, Amsterdam, 1999.
- Neely, A., Gregory, M. and Platts, K., "Performance measurement systems design: a literature review and research agenda", *International Journal of Operations and Production Management*, 15 (4), 1995, 80-116.
- Norreklit, H., "The balance on the balanced scorecard: a critical analysis of some of its assumptions", *Management Accounting Research*, 11, 2000, 65-88.
- O'Brien, R.M., "A caution regarding rules of thumb for variance inflation factors," *Quality and Quantity*, 41(5), 2007, 673-690.
- Otley, D.T., "The contingent theory of management accounting: achievement and prognosis", *Accounting, Organizations and Society*, 5(4), 1980, 413-428.
- Otley, D.T., "Chapter 6 - performance management framework". In A. Berry, J. Broadbent and D. Otley (eds.), *Management control: theories, issues and performance*, 2nd edition, Palgrave Macmillan, New York, 2005.
- Park, H.M., "Univariate analysis and normality test using SAS, Stata, and SPSS", Indiana University, Bloomington, IN, 2008.
- Pollanen, R.M., "Management controls and their organizational and environmental context: a review with managerial emphasis", Paper Presented at the 5th International Management Control Systems Research Conference, London, United Kingdom, July 4 – 6, 2001 and at the American Accounting Association 2001 ABO Research Conference, St Louis, Missouri, October 12 – 13, 2001.
- Pollanen, R.M., "Performance measurement in municipalities: empirical evidence in Canadian context", *International Journal of Public Sector Management*, 18(1), 2005, 4-24.

Porter, M., *Competitive strategy*, New York: The Free Press, 1980.

Prendergast, C., "The tenuous tradeoff between risk and incentives", *Journal of Political Economy*, 39, 2002, 1071-1102.

Rigby, D., "Management tools and techniques: a survey", *California Manage Review*, 43 (2), 2001, 139-160.

Roberts, M.L., Albright, T.L. and Hibbets, A.R., "Debiasing balanced scorecard evaluations", *Behavioural Research in Accounting*, 16, 2004, 75-88.

Ronen, S. and Shenkar, O., "Clustering countries on attitudinal dimensions: a review and synthesis", *Academy of Management Review*, 10(3), 1985, 435-454.

Ryan, B., Scapens, R.W. and Theobald, M., *Research method and methodology in finance and accounting*, Academic Press, San Diego, 1992.

Said, A.A., HassabElnaby, H.R. and Wier, B., "An empirical investigation of the performance consequences of non-financial Measures", *Journal of Management Accounting Research*, 15, 2003, 193-223.

Scheytt, T. and Soin, K., "Ch11 – Culture". In A. Berry, J. Broadbent and D. Otley (eds.), *Management control: theories, issues and performance*, 2nd edition, Palgrave McMillan, New York, 2005.

Schuler, R.S. and Rogovsky, N., "Understanding compensation practice variations across firms: the impact of national culture", *Journal of International Business Studies*, 29 (1), 1st Qtr., 1998, 159-177.

Sila, I., "Examining the effects of contextual factors on TQM and performance through the lens of organizational theories: an empirical study", *Journal of Operations Management*, 25, 2007, 83-109.

Silk, S., "Automating the balanced scorecard", *Management Accounting*, May 1998, 38.

Simons, R., "The role of management control systems in creating competitive advantage: new perspectives", *Accounting, Organizations and Society*, 15, 1/2, 1990, 127-143.

Sloan, R., "Do stock prices fully reflect information earnings?" *The Accounting Review*, 71 (3), 1996, 289-315.

Speckbacher, G., Bischof, J. and Pfeiffer, T., "A descriptive analysis on the implementation of balanced scorecards in German-speaking countries", *Management Accounting Research*, 14, 2003, 361-387.

Statistics Canada, "Policy on informing users of data quality and methodology", March 2000.

Stevens, J., "Applied multivariate statistics for the social sciences", Hillsdale, NJ: Lawrence Erlbaum Associates, 1986.

Thomas, D.R., Lu, I. and Cedzynski, M., "A critique of partial least squares, and a preliminary assessment of an alternative estimation method", working paper, Carleton University, 2006.

Trochim, W.M.K., www.stat.ucla.edu/textbook, 2006.

Venkatraman, N., "The concept of fit in strategy research: toward verbal and statistical correspondence", *The Academy of Management Review*, 14(3), July 1989, 423-444.

Venkatraman, N. and Camillus, J.C., "Exploring the concept of "fit" in strategic management", *The Academy of Management Review*, 9(3), July 1984, 513-525.

Wagner, J.A, Rubin, P.A. and Callahan, T.J., "Incentive payment and nonmanagerial productivity: An interrupted time series analysis of magnitude and trend", *Organizational Behaviour and Human Decision Processes*, 42, 1988, 47-74.

Wallace, J.S., "Adopting residual income-based compensation plans: do you get what you pay for?" *Journal of Accounting and Economics*, 24, 1997, 275-300.

Waterhouse, J.H. and Tiessen, P., "A contingency framework for management accounting systems research", *Accounting, Organizations and Society*, 3, 1978, 65-76.

Wayman, J.C., "Multiple imputation for missing data: what is it and how can I use it?", Center for Social Organization of Schools, Johns Hopkins University, 2003.

Wong-On-Wong, B., Guo, L., Li, W. and Yang, D., "Reducing conflict in balanced scorecard evaluations", *Accounting, Organizations and Society*, 32, 2007, 363-377.

Appendices

Appendix 1: Excerpts from proxy statements

FORM TYPE: 6-K

DOCUMENT DATE: December 31, 2004

FILING DATE: April 11, 2005

ADDRESS: BURNABY, 00000

CIK: 0000868675

TICKER: TU

EXCHANGE: NYSE

SIC CODES: 4812 - Radiotelephone communications

INDUSTRY TYPE: Communications Services

SECTOR ID: Services

.....Specific targets were established using a balanced scorecard approach with 2004 corporate targets tied to profitable growth, customer metrics, business efficiency, and employee engagement. The profitable growth metrics (30%) were made up of TELUS Communications external revenue, TELUS Mobility network revenue, free cash flow, net debt/EBITDA (earnings before interest, taxes, depreciation and amortization) and earnings per share (from continuing operations). Customer metrics (40%) included TELUS Mobility churn rate, high speed subscriber churn (TELUS Communications), a customer loyalty index, and other indices that measure quality of performance and customer loyalty. Business efficiency metrics (20%) consisted of EBITDA margin percentage for TELUS Communications and EBITDA before cost of acquisition for TELUS Mobility. Lastly, employee engagement (10%) was measured, based on a survey of all employees.....

The peer group were:

BCE Inc.

Canadian Tire Corporation Limited

Celestica Inc.

EnCana Corporation

Hudson's Bay Company

Noranda Inc./Falconbridge

Petro-Canada

Quebecor Inc.

Rogers Communications Inc.

Sears Canada Inc.

Shaw Communications

Shell Canada Limited

Sobeys Inc.

Suncor Energy Inc.

The Thomson Corporation

Appendix 2: The list of sample firms

Adoption Year	Treatment Group	SIC	Control Group	SIC
2004	ADVANCED MICRO DEVICES INC	3674	SEMICONDUCTOR MFG INTL	3674
2000	ALLEGHENY TECHNOLOGIES INC	3312	CARPENTER TECHNOLOGY CORP	3312
2006	AMCOR LTD	2600	MERCER INTL INC	2600
2003	AMERICAN MANAGEMENT SYSTEMS INC	7370	INTERNET INITIATIVE	7370
2007	AMERICAN WATER CAPITAL CORP.	4941	CADIZ INC	4941
2003	ARACRUZ CELLULOSE S A	2611	BUCKEYE TECHNOLOGIES INC	2611
2007	ATHENAHEALTH INC	7370	CHINA FINANCE ONLINE CO	7370
1998	ATLANTIC ENERGY INC	4911	TRANSALTA CORP	4911
1998	ATLANTIC CITY ELECTRIC CO	4911	TRANSALTA CORP	4911
	AUSTRALIA & NEW ZEALAND BANKING GROUP LTD			
2005	LTD	6020	KB FINANCIAL GROUP INC	6020
2001	AXCELIS TECHNOLOGIES INC	3559	BROOKS AUTOMATION INC	3559
2008	BANK BRADESCO	6020	KB FINANCIAL GROUP INC	6020
2004	BANK OF NOVA SCOTIA	6020	KB FINANCIAL GROUP INC	6020
1999	BELK INC	5311	BON-TON STORES INC	5311
1996	BESTFOODS	2030	HEINZ (H J) CO	2030
2004	BEVERLY ENTERPRISES INC	8051	BROOKDALE SENIOR LIVING INC	8051
2007	BG GROUP PLC	4922	TRANSCANADA CORP	4922
	BRAZILIAN DISTRIBUTION CO COMPANHIA			
2007	BRASILEIRA DE DISTR CBD	5411	BLUE SQUARE-ISRAEL LTD	5411
2007	BUSINESS OBJECTS S.A.	7372	SOLERA HOLDINGS INC	7372
2002	CABOT CORP	2890	CYTEC INDUSTRIES INC	2890
2006	CAMPBELL SOUP CO	2030	HEINZ (H J) CO	2030
2001	CASTLE A M & CO	5051	OLYMPIC STEEL INC	5051
2001	CELANESE AG	2860	METHANEX CORP	2860
2007	CENTRAL VERMONT PUBLIC SERVICE CORP	4911	EMPRESA DISTRIB Y COMERC (EDENOR)	4911
2001	CHARTER FINANCIAL CORP	6035	PROVIDENT NEW YORK BANCORP	6035
1998	CHICAGO TITLE CORP	6361	FIRST AMERICAN CORP	6361
2001	CLOROX CO	2842	ECOLAB INC	2842

2004	COMMONWEALTH TELEPHONE ENTERPRISES INC	4813	ALASKA COMMUNICATIONS SYS GP	4813
2005	COMPANHIA DE SANEAMENTO BASICO DO	4941	AMERICAN STATES WATER CO	4941
2001	ESTADO DE SAO PAULO-SABESP	3571	SUN MICROSYSTEMS INC	3571
2000	COMPAQ COMPUTER CORP	4911	ENERSIS SA -ADR	4911
2002	GPU INC PA	4911	EMPRESA DISTRIB Y COMERC(EDENOR)	4911
2007	DUQUESNE LIGHT HOLDINGS INC	7370	INTERMOST CORP	7370
2008	EDGAR ONLINE INC	4924	SOUTHERN UNION CO	4924
	ENBRIDGE INC			
2007	COMPANHIA PARANAENSE DE ENERGIA -	4911	EMPRESA DISTRIB Y COMERC (EDENOR)	4911
	COPEL(ENERGY CO OF PARANA)			
	ENERGY CO OF MINAS GERAIS(CIA ENERGETICA			
2004	MINA GERA)	4911	EMPRESA DISTRIB Y COMERC (EDENOR)	4911
2008	FBL FINANCIAL GROUP INC	6311	SCOTTISH RE GROUP LTD	6311
1996	FMC CORP	2800	SINOPEC SHANGHAI PETRO	2800
2008	FROZEN FOOD EXPRESS INDUSTRIES INC	4213	TRAILER BRIDGE INC	4213
1992	GUARDSMAN PRODUCTS INC	2851	TOR MINERALS INTL INC	2810
1997	GENERAL COMMUNICATION INC	4813	PRIMUS TELECOMM GROUP INC	4813
2008	THE BANK OF IRELAND	6020	KB FINANCIAL GROUP INC	6020
2005	HANDLEMAN CO	5099	AUDIOVOX CORP	5065
2008	HARRY WINSTON DIAMOND CORP	1400	FRONTEER DEVELOPMENT GRP INC	1400
2002	HEWLETT PACKARD CO	3570	TOSHIBA CORP	3600
2006	HILTON HOTELS CORP	7011	STARWOOD HOTELS & RESORTS WRULD	7011
2006	HSBC HOLDINGS PLC	6020	MITSUBISHI UFJ FINL GP	6020
2006	HURON CONSULTING GROUP INC.	8742	PDI INC	8742
2005	INFOSYS TECHNOLOGIES LTD	7371	CACI INTL INC	7373
1997	IMPERIAL CHEMICAL INDUSTRIES PLC	2800	RHODIA	2800
2000	JDS UNIPHASE CORP	3663	SHARP CORP	3663
2004	JMAR TECHNOLOGIES INC	3559	AMTECH SYSTEMS INC	3559
2008	KENDLE INTERNATIONAL INC	8731	EXELIXIS INC	8731
1997	KERR MCGEE CORP	1311	EOG RESOURCES INC	1311
2007	KLA TENCOR CORP	3827	ADVANTEST CORP	3825
2000	KRISPY KREME DOUGHNUTS INC	5812	MORTONS RESTAURANT GROUP INC	5812
2006	LEGACY BANCORP, INC.	6036	FIRST BANCSHARES INC	6036
2008	LIHIR GOLD LTD	1040	IAMGOLD CORP	1040

2006	TELIASONERA AB	4812	CHINA UNICOM (HONG KONG)	4812
2005	TENET HEALTHCARE CORP	8062	HEALTHSOUTH CORP	8060
2008	TITAN MACHINERY INC.	5500	CLEAN ENERGY FUELS CORP	5500
1994	TRANSAMERICA CORP	6311	LINCOLN NATIONAL CORP	6311
2008	TRIANGLE PETROLEUM CORPORATION	1311	TOREADOR RESOURCES CORP	1311
2005	TRONOX INC	2810	MILLENNIUM CHEMICALS INC	2810
2004	TYCO INTERNATIONAL LTD	3678	mitsui & co ltd	3678
1999	TYSON FOODS INC	2011	SMITHFIELD FOODS INC	2011
2002	ULTRAPAR HOLDINGS INC	4924	TORM A/S	4412
2005	UMB FINANCIAL CORP	6020	SOUTH FINANCIAL GROUP INC	6020
2003	ITAU UNIBANCO HOLDINGS SA	6020	CREDICORP LTD	6020
2003	UNIBANCO UNION OF BRAZILIAN BANKS			
2003	SA(UNIBANCO UNIAO DE BANC)	6020	CREDICORP LTD	6020
1994	UNIVAR CORP	5160	AMPAL AMERICAN ISRAEL	5160
2008	VIVO PARTICIPACOES S.A.	4812	HUTCHISON TELECOMM INTL	4812
2001	WELLPOINT HEALTH NETWORKS INC	6324	HUMANA INC	6324
2004	WESTPAC BANKING CORP	6020	KB FINANCIAL GROUP INC	6020
1999	WHIRLPOOL CORP	3630	MAKITA CORP	3540
2008	XETA TECHNOLOGIES INC	3661	ECTEL LTD	3661
2004	ABB LTD*	3629	TOSHIBA CORP	3669
2008	AMERICAN ELECTRIC POWER CO*	4911	HUANENG POWER INTL INC	4911
2003	BHP BILLITON LTD*	1000	RIO TINTO GROUP (AUS)	1000
2003	BHP BILLITON PLC*	1000	TECK RESOURCES LTD	1000
2008	BRANDYWINE REALTY TRUST*	6798	COHEN & COMPANY INC	6798
2008	CAMECO CORP*	1090	DENISON MINES CORP	1090
2003	CANADIAN IMPERIAL BANK OF COMMERCE *	6020	DANSKE BANK AS	6020
2007	CASCADE FINANCIAL CORP*	6020	CENTERSTATE BANKS INC	6020
1997	CITIGROUP*	6199	ING GROEP NV	6311
2004	CLIFFS NATURAL RESOURCES INC*	1000	SOUTHERN COPPER CORP	1000
2002	COMMUNITY FIRST BANKSHARES INC*	6020	GRUPO FINANCIERO GALICIA	6020
2005	CRAY INC*	3571	GLOBAL EPOINT INC	3571
2005	DUN & BRADSTREET CORP*	7320	EQUIFAX INC	7320
2009	DYNASIL CORP OF AMERICA *	3827	ZYGO CORP	3827
2008	ENCANA CORP*	1311	CHESAPEAKE ENERGY CORP	1311

2008	EXELON CORP*	4911	EDP-ENERGIAS DE PORTUGAL	4911
2003	FANNIE MAE	6111	FEDERAL HOME LOAN MORTG CORP	6111
2008	FIRST CITIZENS BANCSHARES INC*	6022	GRUPO FINANCIERO GALICIA	6020
2002	FIRST COMMONWEALTH FINANCIAL CORP *	6020	GRUPO FINANCIERO GALICIA	6020
2005	FIRST MERCHANTS CORP*	6020	SUN BANCORP INC	6020
2008	FIRST NIAGARA FINANCIAL GROUP INC*	6036	GRUPO FINANCIERO GALICIA	6020
2007	FPL GROUP INC*	4911	ENERSIS SA	4911
2008	FULTON FINANCIAL CORP*	6020	WEBSTER FINANCIAL CORP	6020
2005	GALLAHER GROUP PLC*	2111	IMPERIAL TOBACCO GROUP	2111
2001	GERMAN AMERICAN BANCORP*	6020	PAB BANKSHARES INC	6020
2005	GREAT PLAINS ENERGY INC*	4911	PORTLAND GENERAL ELECTRIC CO	4911
1999	HARMON INDUSTRIES INC*	3669	CHECKPOINT SYSTEMS INC	3669
2005	HASBRO INC*	3944	JAKKS PACIFIC INC	3942
2007	HAWAIIAN ELECTRIC INDUSTRIES INC	4911	NV ENERGY INC	4911
2008	HAWAIIAN HOLDINGS INC*	4512	PINNACLE AIRLINES CORP	4512
2004	JP MORGAN CHASE & CO*	6020	SUMITOMO MITSUI FIN GP	6020
2000	KONINKLIJKE PHILIPS ELECTRONICS NV*	3600	TOSHIBA CORP	3600
	LLOYDS TSB GROUP PLC (LLOYDS BANKING GROUP) *			
2004	LUBRIZOL CORP*	6020	BANCO BILBAO VIZCAYA	6020
2007	MANPOWER INC*	2990	HEADWATERS INC	2990
2002	MARSH & MCLENNAN COMPANIES, INC. *	7363	KELLY SERVICES INC	7363
2008	MB FINANCIAL INC*	6411	AON CORP	6411
2009	MOLECULAR DEVICES CORP*	6020	NATIONAL PENN BANCSHARES INC	6020
2005	NABI BIOPHARMACEUTICALS	3826	CALIPER LIFE SCIENCES INC	3826
2003	NEXEN INC*	2836	INTERMUNE INC	2836
2007	SIERRA PACIFIC POWER CO*	1311	CHESAPEAKE ENERGY CORP	1311
2008	NYMAGIC INC*	4931	AVISTA CORP	4931
2007	OCWEN FINANCIAL CORP*	6351	RAM HOLDINGS LTD	6351
2009	PEOPLES BANCORP INC*	6035	OCEANFIRST FINANCIAL CORP	6035
2008	PROCTER & GAMBLE CO*	6020	AMERIS BANCORP	6020
2003	REUTERS GROUP PLC ADR*	2840	LOREAL CO	2844
1996	REYNOLDS METALS CO *	2731	WOLTERS KLUWER NV	2732
2006	ROYAL DUTCH SHELL PLC*	3334	WHX CORP	3330
		2911	ENI SPA	2911

2004	SCOTTISH POWER PLC*	4911	PROGRESS ENERGY INC	4911
2005	SEMCO ENERGY INC*	4924	METROGAS SA	4924
2006	SHIRE PLC*	2834	WATSON PHARMACEUTICALS INC	2834
2003	STANDARD REGISTER CO*	2761	IMPRESO INC	2761
2003	TELKOM SA LTD*	4813	PCCW LTD	4813
2003	TELUS CORP*	4813	QWEST COMMUNICATION INTL INC	4813
2006	TENNESSEE VALLEY AUTHORITY*	4911	EDP-ENERGIAS DE PORTUGAL	4911
1999	UIL HOLDINGS CORP*	4911	EL PASO ELECTRIC CO	4911

Note:

1. Asterisk * indicates firms that provide relative weights of financial and non-financial measures and are used in the sub-sample.
2. BSC Firms with Systematically Missing Values (not shown in the table above and excluded from the final sample):
 ANGLIAN WATER, COMPASS GROUP, CPC INTERNATIONAL INC, EPL INTERMEDIATE, INC. (EL POLLO LOCO, INC.), FIRST RELIANCE
 BANCSHARES INC, GCI INC, ICELAND GROUP, KANSAS CITY POWER & LIGHT CO, MONDAY LTD, MORRISON CONSTRUCTION GROUP,
 NATIONAL CONSUMER COOPERATIVE BANK, NOVA CORPORATION (NUVISTA ENERGY LTD), RAILTRACK GROUP, STRATEGIX
 SOLUTIONS INC, TESCO, APPALACHIAN POWER CO, OHIO POWER CO, APPALACHIA POWER, PUBLIC SERVICE CO OF OKLAHOMA,
 SOUTHWESTERN ELECTRIC POWER.

Appendix 3: K-density estimations: continuous data

EBIT margin **Net profit margin** **Dividend payout** **P/E ratio** **Price to cash flow** **Stock return**

Dividend yield

Revenue per employee

Labour cost per employee

Net income per employee

Sales growth

EPS growth

Appendix 4: Histograms: categorical data

Merger & acquisitions

New CEO

Product development cycle

Product life cycle

Appendix 5: Tests of significant control variables (full-sample)

Panel A: BSC and firm performance

		<u>Financial health</u>		<u>Industry volatility</u>	
		<u>Strong group</u>	<u>Weak group</u>	<u>High group</u>	<u>Low group</u>
Effectiveness (H2A)	coef.	-0.2220	0.2425	-0.0092	0.0174
	p-value	0.383	0.390	0.971	0.950
	F-test	0.3833	0.3904	0.9710	0.9498
	R	0.0047	0.0045	0.0000	0.0000
Efficiency (H2B)	Coef.	0.0737	-0.0279	-0.3364	0.0841
	p-value	0.504	0.9270	0.229	0.655
	F-test	0.5037	0.9271	0.2291	0.6553
	R	0.0027	0.0001	0.0089	0.0012

Panel B: BSC fit and firm performance

		<u>Financial health</u>		<u>Industry volatility</u>	
		<u>Strong group</u>	<u>Weak group</u>	<u>High group</u>	<u>Low group</u>
Effectiveness (H2A)	coef.	0.5317	1.1265	0.6863	0.7876
	p-value	0.216	0.038**	0.131	0.111
	F-test	0.2164	0.0383	0.1309	0.1106
	R	0.0103	0.0289	0.0158	0.0168
Efficiency (H2B)	Coef.	0.4819	0.1787	-0.4702	0.6036
	p-value	0.009***	0.7630	0.361	0.06*
	F-test	0.0090	0.7634	0.3611	0.0602
	R	0.0452	0.0006	0.0058	0.0232

Number of observations: 165 for all variables

***1% significance level

**5% significance level

*10% significance level

Panel B: Efficiency (H2B)

Independent variable			Model tests		
BSC dum	Intercept	F-test	R-squared	Adj R-squared	VIF
rev per employee	52.81 (0.641)	584.8 (0.000)***	(0.6411)	0.0007 -0.0024	1.00
laborCost per employee	-7.82 (0.089)*	22.9 (0.000)***	(0.0891)*	0.0088 0.0058	1.01
NI per employee('000 \$)	130.81 (0.074)*	-102.6 (0.048)**	(0.0739)*	0.0097 0.0067	1.01
sales growth rate	3.61 (0.388)	7.91 (0.008)***	(0.3885)	0.0023 -0.0008	1.00
EPS growth rate	142.1 (0.004)***	-125.3 (0.000)***	(0.0037)***	0.0254 0.0224	1.03
inventory turnover	520.6 (0.235)	538.9 (0.083)*	(0.2351)	0.0043 0.0013	1.00
receivables turnover	0.36 (0.781)	6.99 (0.000)***	(0.7812)	0.0002 -0.0028	1.00
total asset turnover	39.25 (0.163)	0.62 (0.975)	(0.1629)	0.0059 0.0029	1.01
operating cycle	-87.3 (0.646)	927.2 (0.000)***	(0.6461)	0.0006 -0.0024	1.00
quick ratio	0.36 (0.251)	0.82 (0.000)***	(0.2505)	0.0040 0.0010	1.00
interest coverage	6.64 (0.738)	11.37 (0.418)	(0.7378)	0.0003 -0.0027	1.00
D/E ratio	-167.8* (0.095)	294.9 (0.000)*	(0.0953)*	0.0085 0.0054	1.01

* 10% significance level

Appendix 7: Test of H2 – BSC and firm performance (with control variables)

Panel A: Effectiveness (H2A)

	Independent variables							Model tests					
	BSC dum	M&A	new CEO	Bankrupt Z-score	Industry ROA	industry volatility	prod dev cycle	Prod life cycle	Intercept	F-test	R-squared	Adj R-squared	VIF
ROA	3.52** (0.012)	1.26 (0.369)	-3.85 (0.024)**	0.45*** (0.002)	-0.14 (0.419)	-1.26*** (0.000)	-5.89** (0.012)	5.25** (0.024)	1.34 (0.402)	0.000***	0.1599	0.1389	1.19
ROE	37.09** (0.031)	12.10 (0.481)	-54.42 (0.01)***	1.50 (0.399)	0.96 (0.667)	-4.34 (0.14)	-44.74 (0.122)	60.06** (0.038)	-25.77 (0.189)	0.0053***	0.0655	0.042	1.07
ROI	10.52 (0.163)	5.46 (0.471)	-12.00 (0.191)	1.61** (0.039)	0.17 (0.862)	-4.81*** (0.000)	-22.34* (0.075)	10.44 (0.402)	4.80 (0.577)	0.0004***	0.0849	0.062	1.09
gross profit margin	7.45 (0.37)	17.15** (0.04)	1.42 (0.889)	1.28 (0.14)	-1.66 (0.118)	-1.84 (0.191)	-8.04 (0.562)	-8.39 (0.543)	44.65*** (0.000)	0.1417	0.037	0.013	1.04
EBIT margin	2.34 (0.905)	29.52 (0.134)	9.88 (0.68)	2.54 (0.212)	-3.30 (0.189)	-3.97 (0.231)	-31.10 (0.342)	-5.26 (0.872)	18.64 (0.406)	0.3013	0.0289	0.0047	1.03
net profit margin	4.13 (0.866)	36.79 (0.135)	9.06 (0.762)	3.12 (0.22)	-2.68 (0.391)	-3.38 (0.415)	-36.64 (0.369)	-0.21 (0.996)	-8.58 (0.759)	0.4949	0.0226	-0.0018	1.02
dividend payout	10.37 (0.718)	-32.03 (0.267)	-50.64 (0.15)	-0.12 (0.969)	-7.96** (0.031)	-7.78 (0.11)	-5.40 (0.91)	10.84 (0.82)	121.84*** (0.000)	0.2719	0.0301	0.0059	1.03
stock return	19.89*** (0.000)	5.93 (0.275)	-5.90 (0.371)	3.08*** (0.000)	-1.01 (0.142)	-2.58*** (0.005)	2.86 (0.751)	6.35 (0.479)	-10.48* (0.091)	0.000***	0.1703	0.1496	1.21
P/E ratio	19.91*** (0.000)	5.48 (0.206)	-13.06 (0.014)**	0.00 (0.999)	-0.08 (0.891)	-0.25 (0.734)	-3.86 (0.592)	11.16 (0.12)	-2.95 (0.55)	0.000***	0.1019	0.0795	1.11
dividend yield	1.77*** (0.000)	-0.32 (0.362)	-0.24 (0.575)	-0.05 (0.151)	-0.13*** (0.003)	-0.19*** (0.001)	-0.39 (0.502)	-0.13 (0.82)	2.87*** (0.000)	0.000***	0.1382	0.1167	1.16
price/cash flow	7.36*** (0.001)	0.50 (0.828)	-0.93 (0.739)	0.40* (0.089)	0.34 (0.24)	0.39 (0.311)	-11.99*** (0.002)	12.06*** (0.002)	-2.78 (0.289)	0.0002***	0.0899	0.0673	1.11

Panel B: Efficiency (H2B)

	Independent variables							Model tests					
	BSC dum	M&A	new CEO	Bankrupt Z-score	Industry ROA	industry volatility	prod dev cycle	Prod life cycle	Intercept	F-test	R-squared	Adj R-squared	VIF
rev per employee	20.84 (0.86)	-76.06 (0.522)	149.42 (0.302)	-14.64 (0.234)	3.68 (0.808)	12.62 (0.529)	105.87 (0.592)	-153.67 (0.435)	579.33*** (0)	0.8354	0.013	-0.0116	0.99
laborCost per employee	-7.82* (0.088)	0.61 (0.894)	7.03 (0.209)	-0.68 (0.151)	-2.21*** (0.000)	-1.47* (0.058)	-12.99* (0.089)	-0.49 (0.948)	41.04*** (0.000)	0.000***	0.1104	0.0882	1.10
NI per employee('000 \$)	103.14 (0.169)	96.56 (0.2)	-48.05 (0.599)	3.99 (0.608)	6.04 (0.528)	-41.42*** (0.001)	8.39 (0.946)	41.82 (0.737)	-52.21 (0.542)	0.0158**	0.0564	0.0329	1.03
sales growth rate	3.13 (0.464)	2.04 (0.635)	-2.15 (0.681)	-0.41 (0.355)	1.78*** (0.001)	-1.28* (0.079)	-2.43 (0.733)	-3.12 (0.661)	7.68 (0.117)	0.0131**	0.058	0.0345	1.04
EPS growth rate	151.36*** (0.003)	-43.24 (0.399)	-33.70 (0.588)	-0.03 (0.995)	-3.54 (0.587)	-4.43 (0.607)	-6.39 (0.94)	16.91 (0.842)	-83.78 (0.151)	0.268	0.0303	0.0061	1.01
inventory turnover	456.8 (0.319)	-289.3 (0.53)	-302.1 (0.589)	-1.6 (0.973)	-39.6 (0.499)	-47.3 (0.542)	-82.5 (0.914)	-365.6 (0.631)	1194.5** (0.023)	0.8367	0.0129	-0.0117	0.99
receivables turnover	-0.14 (0.915)	-1.19 (0.371)	1.45 (0.368)	0.01 (0.938)	0.61*** (0.000)	-0.01 (0.949)	-0.73 (0.742)	0.54 (0.806)	4.89*** (0.001)	0.0314**	0.0508	0.0271	1.03
total asset turnover	47.64 (0.105)	-39.27 (0.182)	22.11 (0.537)	-1.63 (0.592)	-2.57 (0.493)	-0.98 (0.843)	-10.91 (0.823)	-20.67 (0.671)	39.40 (0.24)	0.5246	0.0217	-0.0027	1.00
operating cycle	-107 (0.523)	343** (0.043)	-315 (0.126)	-34* (0.051)	-118*** (0.000)	-67** (0.019)	-262 (0.351)	-1122*** (0.000)	2211*** (0.000)	0.000***	0.2928	0.2752	1.38
quick ratio	0.34 (0.286)	-0.37 (0.249)	-0.27 (0.485)	0.07** (0.028)	0.05 (0.222)	0.06 (0.279)	0.29 (0.591)	-0.19 (0.723)	0.51 (0.169)	0.1622	0.0356	0.0116	1.01
interest coverage	14.50 (0.475)	-30.21 (0.139)	-39.32 (0.113)	-7.49*** (0.000)	3.76 (0.148)	1.06 (0.758)	-8.31 (0.806)	12.34 (0.714)	16.25 (0.484)	0.0192**	0.0549	0.0313	1.03
D/E ratio	-166.70 (0.109)	-73.26 (0.482)	264.71** (0.037)	-5.07 (0.638)	-15.38 (0.247)	-19.03 (0.279)	-120.30 (0.488)	-28.52 (0.869)	463.38*** (0.000)	0.0965*	0.0407	0.0168	1.02

p-value in bracket: *** 1% significance level ** 5% significance level * 10% significance level

Note: OLS regressions were run separately using each individual indicator as dependent variable.

Appendix 8: Tests of H3 – BSC fit and firm performance (no control variables)

Panel A: Effectiveness (H3A)

	Independent variable		Model tests			
	BSC fit	Intercept	F-test	R-squared	Adj R-squared	VIF
ROA	5.19 (0.028)**	-3.04 (0.038)**	0.0280**	0.0162	0.0128	1.02
ROE	64.31 (0.038)**	-43.95 (0.022)**	0.0377**	0.0146	0.0112	1.01
ROI	26.65 (0.047)**	-18.97 (0.023)**	0.0468**	0.0133	0.0099	1.01
gross profit margin	19.59 (0.182)	20.22 (0.027)**	0.1824	0.0060	0.0026	1.01
EBIT margin	23.36 (0.501)	-17.08 (0.427)	0.5007	0.0015	-0.0018	1.00
net profit margin	31.06 (0.471)	-36.83 (0.168)	0.4711	0.0018	-0.0016	1.00
dividend payout	15.62 (0.739)	43.79 (0.133)	0.7392	0.0004	-0.0030	1.00
stock return	28.54 (0.003)**	-15.58 (0.010)**	0.0034***	0.0284	0.0252	1.03
P/E ratio	11.96 (0.123)	3.57 (0.456)	0.1225	0.0080	0.0047	1.01
dividend yield	1.26 (0.049)**	1.63 (0.000)**	0.0492**	0.0130	0.0096	1.01
price/cash flow	7.72 (0.06)*	1.29 (0.611)	0.0598*	0.0119	0.0085	1.01

p-value in bracket: *** 1% significance level

** 5% significance level

* 10% significance level

Note: OLS regressions were run separately using each individual indicator as dependent variable.

Panel B: Efficiency (H3B)

	<u>Independent variable</u>		<u>Model tests</u>			
	BSC fit	Intercept	F-test	R-squared	Adj R-squared	VIF
rev per employee	77.3 (0.708)	623.4 (0.000)***	0.7085	0.0005	-0.0029	1.00
laborCost per employee	-21.9 (0.008)***	15.6 (0.002)***	0.0077***	0.0237	0.0204	1.02
NI per employee('000 \$)	334 (0.013)**	-221 (0.008)***	0.0125**	0.0208	0.0175	1.02
sales growth rate	6.08 (0.414)	10.29 (0.026)	0.4138	0.0023	-0.0011	1.00
EPS growth rate	185 (0.037)**	-160 (0.004)***	0.0374**	0.0145	0.0112	1.01
inventory turnover	14.4 (0.150)	29.5 (0.000)***	0.1503	0.0070	0.0036	1.01
receivables turnover	2.92 (0.139)	8.23 (0.000)***	0.1387	0.0074	0.0040	1.01
total asset turnover	52.8 (0.306)	-6.32 (0.843)	0.3064	0.0035	0.0002	1.00
operating cycle	41 (0.89)	787 (0.000)***	0.8900	0.0001	-0.0033	1.00
quick ratio	0.59 (0.291)	0.67 (0.057)*	0.2910	0.0038	0.0004	1.00
interest coverage	0.84 (0.975)	-2.21 (0.892)	0.9746	0.0000	-0.0034	1.00
D/E ratio	-64 (0.727)	256 (0.026)**	0.7271	0.0004	-0.0030	1.00

Appendix 9: Tests of H3 – BSC fit and firm performance (with control variables)

Panel A: Effectiveness (H3A)

Independent variable										Model tests			
	BSC fit	M&A	new CEO	Bankrupt Z-score	Industry ROA	industry volatility	prod dev cycle	Prod life cycle	Intercept	F-test	R-squared	Adj R-squared	VIF
ROA	4.17* (0.071)	1.45 (0.301)	-2.59 (0.122)	0.35* (0.08)	-0.01 (0.951)	-1.39*** (0)	-6.64*** (0.004)	4.72** (0.038)	1.32 (0.471)	0.0000***	0.1589	0.1357	1.16
ROE	57.84* (0.068)	12.73 (0.505)	-56.06** (0.015)	2.59 (0.338)	1.29 (0.616)	-5.58* (0.1)	-45.22 (0.152)	59.23* (0.061)	-38.31 (0.129)	0.0109**	0.066	0.04	1.04
ROI	20.87 (0.122)	4.09 (0.618)	-10.30 (0.292)	2.52** (0.03)	0.75 (0.485)	-5.72*** (0.000)	-24.21* (0.069)	6.66 (0.615)	0.35 (0.974)	0.0001***	0.1039	0.0791	1.09
gross profit margin	9.73 (0.522)	15.60* (0.092)	1.24 (0.911)	1.67 (0.202)	-1.79 (0.139)	-2.12 (0.186)	-8.00 (0.594)	-10.47 (0.485)	36.92*** (0.002)	0.1227	0.0424	0.016	1.02
EBIT margin	0.42 (0.991)	30.59 (0.163)	10.97 (0.675)	4.19 (0.177)	-3.44 (0.232)	-4.55 (0.231)	-31.45 (0.377)	-10.08 (0.777)	16.73 (0.56)	0.2743	0.0331	0.0065	1.01
net profit margin	2.35 (0.958)	39.12 (0.152)	12.29 (0.706)	5.29 (0.171)	-2.76 (0.441)	-4.15 (0.382)	-36.90 (0.405)	-7.93 (0.858)	-8.35 (0.815)	0.4275	0.0271	0.0003	1.00
dividend payout	31.52 (0.519)	-22.38 (0.451)	-45.61 (0.199)	-1.26 (0.764)	-6.87* (0.078)	-7.97 (0.122)	-15.93 (0.741)	2.76 (0.954)	110.9*** (0.005)	0.3754	0.029	0.0022	1.00
stock return	13.47 (0.134)	6.25 (0.253)	-8.50 (0.193)	6.49*** (0)	-1.32* (0.065)	-3.18*** (0.001)	5.38 (0.544)	1.04 (0.906)	-8.65 (0.227)	0.0000***	0.2513	0.2306	1.30
P/E ratio	11.05 (0.166)	10.44** (0.032)	-12.02** (0.039)	-0.49 (0.476)	-0.27 (0.674)	-0.54 (0.519)	-1.53 (0.845)	10.28 (0.192)	0.94 (0.882)	0.0379**	0.0542	0.0281	1.03
dividend yield	1.46** (0.025)	-0.16 (0.691)	-0.19 (0.694)	-0.06 (0.264)	-0.17*** (0.001)	-0.21*** (0.003)	-0.43 (0.502)	-0.24 (0.711)	3.26*** (0.000)	0.0002***	0.0974	0.0725	1.08
price/cash flow	5.47 (0.192)	1.02 (0.688)	-0.80 (0.791)	0.73** (0.042)	0.35 (0.291)	0.36 (0.412)	-12.50*** (0.003)	11.73*** (0.005)	-1.84 (0.581)	0.0049***	0.0724	0.0469	1.05

Panel B: Efficiency (H3B)

Independent variable										Model tests			
	BSC fit	M&A	new CEO	Bankrupt Z-score	Industry ROA	industry volatility	prod dev cycle	Prod life cycle	Intercept	F-test	R-squared	Adj R-squared	VIF
rev per employee	21.8 (0.92)	-93.02 (0.48)	173.84 (0.27)	-20.33 (0.276)	3.34 (0.846)	21.88 (0.339)	98.45 (0.645)	-137.13 (0.521)	585.33*** (0.001)	0.7954	0.0157	-0.0114	0.99
laborCost per employee	-23.3*** (0.004)	0.61 (0.901)	8.06 (0.172)	-1.58** (0.025)	-2.36*** (0.000)	-1.04 (0.226)	-14.23* (0.076)	2.07 (0.796)	32.80*** (0.000)	0.000***	0.1374	0.1136	1.13
NI per employee('000 \$)	304** (0.026)	99.16 (0.232)	-48.01 (0.628)	3.92 (0.738)	10.69 (0.325)	-52.01*** (0.000)	9.98 (0.941)	28.5 (0.832)	-157 (0.147)	0.0021***	0.0794	0.054	1.06
sales growth rate	21.80 (0.92)	-93.02 (0.48)	173.84 (0.27)	-20.33 (0.276)	3.34 (0.846)	21.88 (0.339)	98.45 (0.645)	-137 (0.521)	585*** (0.001)	0.7954	0.0157	-0.0114	0.99
EPS growth rate	192** (0.041)	-20.7 (0.715)	-31.6 (0.642)	2.51 (0.755)	-3.18 (0.67)	-7.88 (0.426)	14.07 (0.879)	-12.99 (0.888)	-117 (0.116)	0.7074	0.0185	-0.0086	0.99
inventory turnover	8.52 (0.408)	-5.32 (0.396)	-10.76 (0.151)	-1.04 (0.239)	-1.69** (0.04)	-1.82* (0.094)	9.10 (0.371)	-6.63 (0.514)	43.7*** (0.000)	0.055*	0.0506	0.0244	1.03
receivables turnover	2.56 (0.202)	-2.35* (0.053)	-0.55 (0.704)	0.26 (0.129)	0.59*** (0.000)	-0.03 (0.9)	-0.98 (0.62)	0.96 (0.625)	6.49*** (0.000)	0.0011***	0.0847	0.0595	1.06
total asset turnover	67.8 (0.21)	-40.93 (0.213)	23.29 (0.553)	-2.13 (0.646)	-3.29 (0.445)	-1.08 (0.849)	-10.26 (0.847)	-21.93 (0.68)	33.52 (0.436)	0.6391	0.0205	-0.0065	0.99
operating cycle	-75 (0.777)	224 (0.164)	-171 (0.373)	-56** (0.015)	-113*** (0.000)	-36 (0.198)	-236 (0.368)	-954*** (0.000)	1906*** (0.000)	0.000***	0.2997	0.2804	1.39
quick ratio	0.63 (0.285)	-0.37 (0.303)	-0.33 (0.438)	0.07 (0.185)	0.04 (0.41)	0.07 (0.271)	0.35 (0.541)	-0.25 (0.668)	0.39 (0.412)	0.577	0.0224	-0.0046	1.00
interest coverage	45.1* (0.063)	-0.19 (0.989)	-23.88 (0.174)	-19.00*** (0.000)	0.54 (0.779)	0.66 (0.795)	-33.85 (0.156)	27.05 (0.256)	9.70 (0.614)	0.000***	0.2437	0.2229	1.29
D/E ratio	-38 (0.842)	-111 (0.337)	278 (0.046)**	-13.5 (0.412)	-14.3 (0.349)	-17.1 (0.396)	-138 (0.464)	-6.87 (0.971)	425*** (0.006)	0.2273	0.0354	0.0088	1.01

(#) p-value : *** 1% significance level ** 5% significance level * 10% significance level
Note: OLS regressions were run separately using each individual indicator as dependent variable.

Appendix 10: Tests of significant control variables (sub-sample)

Panel A: non-financial weight and firm performance

		<u>Financial health</u>		<u>Industry volatility</u>	
		<u>Strong group</u>	<u>Weak group</u>	<u>High group</u>	<u>Low group</u>
Effectiveness (H2A)	coef.	-0.2960	-2.8034	-4.1205	-0.8792
	p-value	0.843	0.279	0.101	0.526
	F-test	0.8433	0.2790	0.0938	0.5260
	R	0.0015	0.0433	0.1005	0.0151
Efficiency (H2B)	Coef.	1.5773	-1.8240	1.3154	2.6528
	p-value	0.133	0.4820	0.300	0.236
	F-test	0.1327	0.4818	0.3001	0.2364
	R	0.0817	0.0185	0.0397	0.0515

Panel B: non-financial fit and firm performance

		<u>Financial health</u>		<u>Industry volatility</u>	
		<u>Strong group</u>	<u>Weak group</u>	<u>High group</u>	<u>Low group</u>
Effectiveness (H2A)	coef.	-5.2556	-7.4236	-8.3755	-5.9871
	p-value	0.017**	0.301	0.045**	0.003***
	F-test	0.0167	0.3015	0.0452	0.0030
	R	0.1941	0.0395	0.1404	0.2831
Efficiency (H2B)	Coef.	0.8942	-13.7107	0.2641	4.8613
	p-value	0.574	0.049**	0.905	0.165
	F-test	0.5736	0.0490	0.9047	0.1649
	R	0.0119	0.1360	0.0005	0.0702

Number of observations: 29 for all variables

***1% significance level

**5% significance level

*10% significance level

Appendix 11: Non-financial weights and firm performance (no control variables)

Panel A: Effectiveness

	Independent variable		Model tests			
	Nonfin Weight	Intercept	F-test	R-squared	Adj R-squared	VIF
ROA	13.64 (0.029)*	9.24 (0.005)***	0.0291**	0.0822	0.0658	1.09
ROE	65.39 (0.147)	51.12 (0.031)**	0.1472	0.0372	0.0200	1.04
ROI	43.75 (0.121)	31.05 (0.036)	0.1214	0.0423	0.0252	1.04
gross profit margin	1.73 (0.927)	38.89 (0.000)***	0.9269	0.0002	-0.0177	1.00
EBIT margin	11.44 (0.585)	19.82 (0.073)*	0.5855	0.0053	-0.0124	1.01
net profit margin	24.22 (0.185)	15.35 (0.107)	0.1850	0.0312	0.0139	1.03
dividend payout	188.8 (0.427)	-12.16 (0.921)	0.4270	0.0113	-0.0064	1.01
stock return	31.03 (0.371)	24.85 (0.170)	0.3711	0.0143	0.0033	1.01
P/E ratio	1.71 (0.925)	18.21 (0.060)*	0.9253	0.0002	-0.0177	1.00
dividend yield	1.39 (0.592)	2.52 (0.067)*	0.5921	0.0052	-0.0126	1.01
price/cash flow	13.02 (0.005)***	11.32 (0.000)***	0.0049	0.1327	0.1172	1.15

Panel B: Efficiency

Independent variable			Model tests			
	Nonfin Weight	Intercept	F-test	R-squared	Adj R-squared	VIF
rev per employee	896 (0.415)	375 (0.511)	0.4149	0.0119	-0.0057	1.01
laborCost per employee	-28.04 (0.391)	29.56 (0.085)*	0.3912	0.0132	-0.0045	1.01
NI per employee('000 \$)	102 (0.586)	24.7 (0.799)	0.5863	0.0053	-0.0124	1.01
sales growth rate	6.03 (0.361)	2.98 (0.385)	0.3612	0.0149	-0.0027	1.02
EPS growth rate	21.94 (0.802)	-4.18 (0.927)	0.8022	0.0011	-0.0167	1.00
inventory turnover	17.66 (0.204)	1.85 (0.796)	0.2039	0.0287	0.0113	1.03
receivables turnover	1.16 (0.772)	4.21 (0.046)**	0.7720	0.0015	-0.0163	1.00
total asset turnover	0.81 (0.015)*	0.89 (0.000)***	0.0150**	0.1010	0.0850	1.11
operating cycle	-1472 (0.225)	384 (0.541)	0.2253	0.0261	0.0087	1.03
quick ratio	0.669 (0.144)	0.981 (0.000)***	0.1438	0.0378	0.0206	1.04
interest coverage	34.4 (0.419)	30.2 (0.175)	0.4192	0.0117	-0.0060	1.01
D/E ratio	-39.5 (0.889)	153 (0.300)	0.8892	0.0003	-0.0175	1.00

p-value in bracket: *** 1% significance level ** 5% significance level * 10% significance level
 Note: OLS regressions were run separately using each individual indicator as dependent variable

Appendix 12: Non-financial weights and firm performance (with control variables)

	Independent variables								Model tests				
	nonfin weight	M&A	new CEO	bankruptcy Z-score	prod dev cycle	prod life cycle	industry ROA %	industry volatility	Intercept	F-test	R-squared	Adj R-squared	VIF
Panel A: Effectiveness													
ROA(%)	5.64 (0.329)	-0.57 (0.768)	-4.31 (0.044)**	2.67 (0.001)***	-6.89 (0.028)**	6.89 (0.03)**	-0.46 0.115	-1.28 (0)***	1.65 (0.614)	0.000***	0.5217	0.4436	2.09
ROE(%)	31.81 (0.493)	-7.18 (0.646)	1.62 (0.924)	0.46 (0.939)	-97.29 (0)***	93.21 (0)***	2.30 (0.32)	-6.70 (0.013)*	13.40 (0.611)	0.0012***	0.39	0.2904	1.64
ROI(%)	12.45 (0.687)	0.43 (0.967)	15.67 (0.168)	6.11 (0.13)	-43.51 (0.011)**	39.03 (0.022)**	-0.66 (0.666)	-3.54 (0.045)**	0.94 (0.957)	0.0132**	0.3111	0.1987	1.45
gross_profit_margin(%)	28.14 (0.194)	10.10 (0.167)	-12.13 (0.127)	0.57 (0.838)	-11.94 (0.300)	-2.38 (0.838)	0.63 (0.555)	-3.09 (0.013)**	36.45 (0.004)***	0.0824*	0.2361	0.1114	1.31
EBIT_margin(%)	22.86 (0.294)	4.12 (0.572)	-11.48 (0.151)	0.08 (0.978)	-14.09 (0.226)	8.63 (0.461)	-0.72 (0.502)	-5.95 (0)***	25.75 (0.041)**	0.0021***	0.3731	0.2708	1.60
net_profit_margin(%)	17.47 (0.338)	5.54 (0.367)	-9.96 (0.137)	3.05 (0.198)	-14.67 (0.134)	14.05 (0.155)	-0.93 (0.306)	-4.98 (0)***	6.62 (0.522)	0.0003***	0.4269	0.3333	1.74
dividend_payout(%)	365.66 (0.215)	-119.88 (0.228)	-96.92 (0.367)	3.44 (0.928)	-147.39 (0.348)	38.35 (0.808)	-4.57 (0.754)	-20.79 (0.211)	112.22 (0.501)	0.6544	0.108	-0.0376	1.12
stock_return(%)	15.03 (0.685)	2.18 (0.862)	-24.65 (0.073)*	14.44 (0.004)***	16.37 (0.41)	0.13 (0.995)	-6.94 (0)***	-1.50 (0.475)	2.15 (0.919)	0.0079***	0.3295	0.2201	1.49
P/E_ratio	20.47 (0.334)	8.35 (0.244)	-6.90 (0.372)	6.87 (0.015)**	12.03 (0.288)	-12.77 (0.264)	-0.40 (0.703)	-1.51 (0.207)	0.76 (0.949)	0.1298	0.2142	0.0859	1.27
dividend_yield (%)	3.66 (0.238)	-1.37 (0.191)	-0.51 (0.649)	-0.03 (0.945)	-2.25 (0.175)	0.82 (0.622)	-0.11 (0.464)	-0.36 (0.043)**	4.39 (0.015)**	0.2542	0.1777	0.0435	1.22
price/cash_flow	5.40 (0.185)	1.80 (0.189)	1.62 (0.274)	2.14 (0)***	-0.79 (0.715)	3.39 (0.123)	-0.05 (0.808)	0.03 (0.901)	1.31 (0.568)	0.000***	0.5835	0.5155	2.40

	<u>Independent variables</u>										<u>Model tests</u>		
	nonfin weight	M&A	new CEO	bankruptcy Z-score	prod dev cycle	prod life cycle	industry ROA %	industry volatility	Intercept	F-test	R- squared	Adj R- squared	VIF
Panel B: Efficiency													
rev_per_employee	997.92 (0.466)	-277.99 (0.547)	-522.25 (0.298)	-96.45 (0.586)	38.96 (0.957)	-44.71 (0.952)	104.45 (0.128)	-53.30 (0.489)	635.30 (0.415)	0.746	0.0937	-0.0543	1.10
laborCost_per_employee	-47.77 (0.224)	-1.62 (0.902)	-2.84 (0.842)	1.02 (0.841)	-18.20 (0.384)	-6.13 (0.771)	-2.63 (0.178)	-2.96 (0.181)	48.26 (0.034)**	0.3149	0.1646	0.0282	1.20
NI_per_employee	258.02 (0.249)	-25.53 (0.734)	-103.67 (0.205)	-3.55 (0.902)	-34.45 (0.772)	7.12 (0.953)	16.06 (0.151)	-28.15 (0.028)**	46.84 (0.711)	0.292	0.1693	0.0337	1.20
sales_growth_rate	2.64 (0.726)	1.97 (0.437)	0.01 (0.997)	0.56 (0.568)	3.83 (0.342)	-2.76 (0.497)	0.49 (0.192)	-1.04 (0.017)**	0.92 (0.829)	0.0761*	0.2398	0.1157	1.32
EPS_growth_rate	42.40 (0.661)	42.89 (0.192)	19.14 (0.588)	4.22 (0.736)	25.37 (0.623)	-9.74 (0.852)	3.01 (0.532)	-19.29 (0.001)***	-25.75 (0.64)	0.0284**	0.2818	0.1646	1.39
inventory_turnover	11.08 (0.514)	-4.51 (0.431)	4.30 (0.488)	-1.50 (0.494)	3.17 (0.726)	9.10 (0.321)	-0.25 (0.766)	-0.44 (0.647)	4.96 (0.607)	0.4803	0.1351	-0.0061	1.16
receivables_turnover	4.29 (0.316)	-2.71 (0.063)*	-0.75 (0.629)	0.52 (0.346)	-0.85 (0.707)	4.88 (0.038)**	0.26 (0.227)	0.00 (0.997)	0.62 (0.797)	0.0078***	0.3299	0.2205	1.49
total_asset_turnover	-0.55 (0.099)*	-0.02 (0.827)	0.05 (0.651)	0.07 (0.12)	-0.07 (0.69)	0.28 (0.121)	0.04 (0.021)**	0.02 (0.344)	0.36 (0.054)*	0.0001	0.4674	0.3805	1.88
operating_cycle(days)	1094 (0.279)	372 (0.275)	-1030 (0.007)***	-308 (0.021)**	-389 (0.469)	-1380 (0.014)**	-32 (0.517)	-55 (0.329)	2269 0	0.000***	0.6009	0.5358	2.51
quick_ratio	0.37 (0.366)	-0.11 (0.43)	0.32 (0.036)**	0.24 (0)**	0.40 (0.069)*	-0.29 (0.194)	0.01 (0.496)	0.01 (0.544)	0.33 (0.159)	0.000***	0.5451	0.4709	2.20
interest_coverage	66.53 (0.135)	-20.62 (0.169)	-21.39 (0.187)	28.14 (0)**	-6.20 (0.792)	-32.33 (0.177)	-2.86 (0.195)	-3.83 (0.127)	-6.29 (0.802)	0.0021***	0.3732	0.2709	1.60
D/E(%)	49.03 (0.886)	-107.60 (0.355)	-61.02 (0.627)	-81.86 (0.07)*	-249.09 (0.178)	234.31 (0.209)	20.78 (0.227)	-19.16 (0.324)	339.13 (0.087)*	0.4739	0.1361	-0.0049	1.16

p-value: *** 1% significance level

** 5% significance level

* 10% significance level

Note: OLS regressions were run separately using each individual indicator as dependent variable.

Appendix 13: Non-financial fit and firm performance (no control variables)

Panel A: Effectiveness

Independent variable			Model tests			
Nonfin fit	Intercept		F-test	R-squared	adj. R-squared	VIF
ROA	36.47 (0.000)***	20.34 (0.000)***	0.0002***	0.2180	0.2040	1.28
ROE	210 (0.004)***	121 (0.001)***	0.0035***	0.1424	0.1270	1.17
ROI	169 (0.000)***	91 (0.000)***	0.0001***	0.2340	0.2203	1.31
gross profit margin	11.57 (0.709)	45.36 (0.005)***	0.7090	0.0025	-0.0153	1.00
EBIT margin	21.53 (0.532)	24.73 (0.153)	0.5318	0.0070	-0.0107	1.01
net profit margin	51.95 (0.082)*	28.83 (0.054)*	0.0816	0.0532	0.0363	1.06
dividend payout	83 (0.831)	39 (0.841)	0.8312	0.0008	-0.0170	1.00
stock return	82.9 (0.143)	50.1 (0.078)*	0.1430	0.0379	0.0207	1.04
P/E ratio	14.3 (0.632)	24.4 (0.107)	0.6317	0.0041	-0.0137	1.00
dividend yield	1.62 (0.706)	2.41 (0.262)	0.7057	0.0026	-0.0152	1.00
price/cash flow	27.9 (0.000)***	18.5 (0.000)***	0.0002***	0.2256	0.2118	1.29

Panel B: Efficiency

	Independent variable		Model tests			
	Nonfin fit	Intercept	F-test	R-squared	adj. R-squared	VIF
rev per employee	491 (0.786)	572 (0.527)	0.7859	0.0013	-0.0165	1.00
laborCost per employee	-53.6 (0.317)	17.1 (0.522)	0.3171	0.0179	0.0003	1.02
NI per employee	83.5 (0.786)	114 (0.455)	0.7861	0.0013	-0.0165	1.00
sales growth rate	4.02 (0.711)	2.00 (0.712)	0.7114	0.0025	-0.0153	1.00
EPS growth rate	5.04 (0.972)	12.4 (0.863)	0.9721	0.0000	-0.0178	1.00
inventory turnover	31.8 (0.613)	-5.01 (0.657)	0.1632	0.0344	0.0172	1.04
receivables turnover	0.78 (0.905)	5.15 (0.120)	0.9050***	0.0003	-0.0176	1.00
total asset turnover	1.43 (0.009)***	1.19 (0.000)***	0.0090***	0.1156	0.0998	1.13
operating cycle	-2904 (0.144)	-311 (0.751)	0.1440	0.0377	0.0205	1.04
quick ratio	0.73 (0.335)	1.01 (0.009)***	0.3346	0.0166	-0.0009	1.02
interest coverage	106 (0.126)	65 (0.062)*	0.1259	0.0413	0.0242	1.04
D/E ratio	-292 (0.530)	31 (0.895)	0.5298	0.0071	-0.0106	1.01

p-value in bracket: *** 1% significance level ** 5% significance level * 10% significance level
 Note: OLS regressions were run separately using each individual indicator as dependent variable.

Appendix 14: Non-financial fit and firm performance (with control variables)

	<u>Independent variables</u>										<u>Model tests</u>		
	nonfin fit	M&A	new CEO	bankruptcy Z-score	prod dev cycle	prod life cycle	industry ROA %	industry volatility	Intercept	F-test	R- squared	Adj R- squared	VIF
Panel A: Effectiveness													
ROA(%)	14.96 (0.183)	-0.15 (0.937)	-3.77 (0.072)*	1.96 (0.013)**	-3.18 (0.366)	4.80 (0.144)	-0.46 (0.111)	-1.19 (0)***	11.64 (0.055)*	0.000***	0.5297	0.4529	2.13
ROE(%)	60.96 (0.502)	-5.15 (0.741)	4.31 (0.797)	-2.82 (0.648)	-80.61 (0.006)***	83.89 (0.002)***	2.34 (0.31)	-6.20 (0.017)**	57.99 (0.234)	0.0012***	0.3898	0.2901	1.64
ROI(%)	128.25 (0.03)**	2.76 (0.78)	18.28 (0.091)*	1.77 (0.653)	-18.07 (0.321)	24.42 (0.149)	-0.85 (0.559)	-3.23 (0.049)**	70.72 (0.025)**	0.0021***	0.3731	0.2707	1.60
gross_profit_margin(%)	34.82 (0.414)	10.58 (0.152)	-11.07 (0.163)	0.27 (0.926)	-13.25 (0.324)	-1.29 (0.917)	0.85 (0.431)	-2.75 (0.024)**	31.40 (0.17)	0.1158	0.2199	0.0926	1.28
EBIT_margin(%)	7.72 (0.857)	4.82 (0.513)	-10.31 (0.197)	-0.77 (0.793)	-11.44 (0.398)	7.35 (0.556)	-0.59 (0.588)	-5.65 (0)***	31.96 (0.166)	0.0033***	0.3591	0.2545	1.56
net_profit_margin(%)	2.75 (0.939)	6.20 (0.316)	-8.95 (0.181)	2.15 (0.382)	-11.07 (0.328)	12.16 (0.247)	-0.84 (0.357)	-4.73 (0)***	15.71 (0.413)	0.0005***	0.416	0.3206	1.71
dividend_payout(%)	112.13 (0.847)	-108.52 (0.281)	-78.12 (0.469)	-10.43 (0.793)	-102.86 (0.575)	16.66 (0.922)	-2.41 (0.87)	-15.95 (0.329)	217.25 (0.485)	0.8259	0.0801	-0.0701	1.09
stock_return(%)	83.17 (0.249)	3.94 (0.749)	-22.57 (0.094)*	11.29 (0.025)**	34.11 (0.135)	-9.98 (0.633)	-7.03 (0)***	-1.19 (0.552)	50.48 (0.191)	0.005***	0.3454	0.2385	1.53
P/E ratio	23.33 (0.575)	8.73 (0.225)	-6.10 (0.429)	6.60 (0.023)**	11.43 (0.383)	-12.20 (0.316)	-0.25 (0.816)	-1.26 (0.281)	-1.91 (0.931)	0.158	0.2042	0.0742	1.26
dividend_yield (%)	3.22 (0.597)	-1.29 (0.223)	-0.36 (0.753)	-0.10 (0.801)	-2.19 (0.257)	0.82 (0.643)	-0.09 (0.576)	-0.31 (0.072)*	4.38 (0.182)	0.3451	0.1586	0.0213	1.19
price/cash_flow	13.07 (0.099)*	1.80 (0.184)	1.51 (0.298)	2.01 (0)***	0.62 (0.799)	2.52 (0.272)	-0.10 (0.601)	-0.03 (0.89)	5.48 (0.192)	0.000***	0.5917	0.5251	2.45

	Independent variables										Model tests		
	nonfin fit	M&A	new CEO	bankruptcy Z-score	prod dev cycle	prod life cycle	industry ROA %	industry volatility	Intercept	F-test	R- squared	Adj R- squared	VIF
Panel B: Efficiency													
rev_per_employee(\$1000)	162.02 (0.952)	-240.09 (0.604)	-463.98 (0.353)	-148.01 (0.421)	245.30 (0.772)	-153.05 (0.845)	109.40 (0.113)	-39.54 (0.599)	1156.58 (0.421)	0.8051	0.0838	-0.0658	1.09
laborCost_per_employee	-88.09 (0.252)	-1.22 (0.926)	-1.48 (0.917)	1.36 (0.795)	-25.69 (0.288)	-1.25 (0.955)	-2.20 (0.258)	-2.41 (0.262)	25.16 (0.538)	0.3293	0.1617	0.0248	1.19
NL_per_employee(000 \$)	63.20 (0.886)	-15.42 (0.839)	-88.28 (0.282)	-17.50 (0.562)	22.76 (0.87)	-23.14 (0.857)	17.29 (0.126)	-24.57 (0.051)*	192.30 (0.416)	0.4111	0.1466	0.0073	1.17
sales_growth_rate	3.49 (0.812)	1.83 (0.471)	-0.19 (0.944)	0.78 (0.438)	2.73 (0.556)	-2.15 (0.617)	0.49 (0.197)	-1.08 (0.011)**	-1.99 (0.8)	0.0779*	0.2387	0.1144	1.31
EPS_growth_rate	80.95 (0.669)	43.21 (0.188)	20.31 (0.562)	4.60 (0.721)	18.23 (0.759)	-5.12 (0.926)	3.39 (0.48)	-18.80 (0.001)***	-47.64 (0.637)	0.0285**	0.2817	0.1644	1.39
inventory_turnover	12.67 (0.703)	-4.30 (0.453)	4.73 (0.443)	-1.65 (0.468)	2.84 (0.786)	9.42 (0.333)	-0.17 (0.844)	-0.30 (0.745)	3.49 (0.844)	0.5117	0.13	-0.012	1.15
receivables_turnover	6.86 (0.414)	-2.66 (0.069)*	-0.61 (0.692)	0.52 (0.363)	-1.34 (0.612)	5.21 (0.037)**	0.29 (0.169)	0.05 (0.832)	-0.92 (0.836)	0.0089***	0.3253	0.2151	1.48
total_asset_turnover	0.47 (0.477)	-0.04 (0.745)	0.03 (0.803)	0.08 (0.083)*	-0.08 (0.69)	0.28 (0.149)	0.03 (0.041)**	0.01 (0.563)	0.36 (0.308)	0.0002***	0.4425	0.3515	1.79
operating_cycle(days)	-1624 (0.413)	387 (0.258)	-993 (0.009)***	-312 (0.024)**	-489 (0.432)	-1309 (0.026)**	-23 (0.641)	-42 (0.443)	1938 (0.071)*	0.000***	0.5968	0.531	2.48
quick_ratio	0.69 (0.388)	-0.13 (0.338)	0.29 (0.055)*	0.28 (0)***	0.21 (0.402)	-0.18 (0.435)	0.01 (0.513)	0.01 (0.716)	-0.18 (0.673)	0.000***	0.5445	0.4701	2.20
interest_coverage	60.32 (0.493)	-19.14 (0.208)	-18.56 (0.256)	26.79 (0)***	-5.32 (0.847)	-32.08 (0.213)	-2.39 (0.285)	-2.99 (0.226)	-7.19 (0.878)	0.0043***	0.35	0.2439	1.54
D/E(%)	-489.68 (0.465)	-113.07 (0.328)	-65.57 (0.596)	-69.81 (0.13)	-329.11 (0.122)	281.26 (0.153)	22.02 (0.197)	-19.06 (0.31)	115.25 (0.747)	0.4194	0.1452	0.0057	1.17

p-value: *** 1% significance level

** 5% significance level

* 10% significance level

Note: OLS regressions were run separately using each individual indicator as dependent variable.