The Consistent Estimation of Future Cash Flow and Future Earnings:

An Accounting Prediction Model with Double Entry Constraint

A THESIS

Submitted by

Ehsan Khansalar

For the award of the Degree of

DOCTOR OF PHILOSOPHY

SCHOOL OF BUSINESS, MANAGEMENT AND ECONOMICS

MANTEL BUILDING, UNIVERSITY OF SUSSEX, BRIGHTON, BN1 9RF

August 2011

ACKNOWLEDGEMENTS

This thesis would not have been possible without the kind support, the trenchant critiques, the probing questions, and the remarkable patience of my thesis advisor - Prof. Stuart McLeay. I am heartily thankful to him.

He has taught me, both consciously and unconsciously, how good empirical accounting work is done. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for his research was contagious and motivational for me, even during tough times in the Ph.D. pursuit. I am also thankful for the excellent example he has provided as a successful accountant, auditor and professor.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Azad University in Iran- Shiraz Branch. My work was also supported by School of Business, Management and Economics in the University of Sussex, and the Discipline of Accounting in the University of Sydney. My time there was made enjoyable in large part due to the friends that became a part of my life. I am grateful for time spent with my colleagues from Iran who are experienced lecturers and auditors and accountants, Dr. Javad Izadi and Dr. Mahmoud Lari .

In the end, I would like to thank my family for all their love and encouragement. For, my parents who raised me with a love of science and supported me in all my pursuits - Thank you.

Ehsan Khansalar
August 2011

DECLARATION

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature: Name:

The Consistent Estimation of Future Cash Flow and Future Earnings: A Predictive Model with Accounting Double Entry Constraint

Abstract

In empirical financial accounting research, there continues to be a debate as to what the best predictors of future earnings and future cash flows might be. Past accruals, earnings and cash flows are the most common predictors, but there is no consensus over their relative contributions, and little attention to the underlying accounting identities that link the components of these three prominent variables. The aim of this thesis is to investigate this controversy further, and to apply an innovative method which yields consistent estimations of future earnings and cash flows, with higher precision and greater efficiency than is the case in published results to date. The estimation imposes constraints based on financial statement articulation, using a system of structural regressions and a framework of simultaneous linear equations, which allows for the most basic property of accounting - double entry book-keeping - to be incorporated as a set of constraints within the model. In predicting future cash flows, the results imply that the constrained model which observes the double entry condition is superior to the models that are not constrained in this way, producing (a) rational signs consistent with expectations, not only in the entire sample but also in each industry, (b) evidence that double entry holds, based on the Wald test that the estimated marginal responses sum to zero, and (c) confirmation of model improvement by way of a higher likelihood and greater precision attached to predictor variables. Furthermore, by then using an appropriately specified model that observes the double entry constraint in order to predict earnings, the thesis reports statistically significant results, across all industries, that cash flows are superior to accruals in explaining future earnings, indicating also that accruals with a lower level of reliability tend to be more relevant in this respect.

Keywords: Double entry system; Cash flow prediction; Earnings prediction; Reliability of accruals.

CONTENTS

LIST OF ABBREVIATIONS

In Alphabetic Order

ACP	Accrued payroll
APA	Accounts payable
ARE	Accounts receivable
CED	Capital expenditure (net of asset disposals)
CGS	Cost of goods sold
CSI	Cash and short-term investments
DDA	Depreciation, depletion and amortization
DIN	Deferred income
DIR	Debt issues (net of debt repayments)
DOP	Discontinued operations
DPA	Dividend payable
DTU	Deferred tax in untaxed reserves
DTX	Deferred tax
EBIT	Earnings before interest & taxes
ECH	Extraordinary charge
ECR	Extraordinary credit
EGL	Extraordinary items and gain(loss) on sale of assets
EIE	Equity interest earnings
FCF	Financial cash flows
GRI	Gross income
ICF	Investing cash flows
ICP	Interest capitalized
IED	Interest expense on debt
IIN	Interest income
INV	Total inventory
ITP	Income tax payable
ITX	Income taxes
IUS	Investments in unconsolidated subsidiaries
LTD	Long-term debt
LTR	Long-term receivables
MIE	Minority interest in earnings
MIN	Minority interest in net assets
NCD	New capital (net of dividends)
NER	on-equity reserves

By Financial Statement Item

CSI	Cash and short-term investments	Current assets
ARE	Accounts receivable	
INV	Total inventory	
PRE	Prepaid expenses	
OCA	Other current assets	
TCA	Total current assets	
PPE	Property plant and equipment	Long-term assets
LTR	Long-term receivables	
IUS	Investments in unconsolidated subsidiaries	
OIN	Other investments	
OLA	Other long-term assets	
TLA	Total long-term assets	
TA	Total assets	*TCA+TLA*
APA	Accounts payable	Current liabilities
STD	Short-term debt	
ACP	Accrued payroll	
ITP	Income tax payable	
DPA	Dividend payable	
OCL	Other current liabilities	
TCL	Total current liabilities	
LTD	Long-term debt	Long-term liabilities
PRC	Provision for risks and charges	
DIN	Deferred income	
DTX	Deferred tax	
DTU	Deferred tax in untaxed reserves	
OLL	Other long-term liabilities	
TLL	Total long-term liabilities	
TL	Total liabilities	
SEQ	Shareholders' equity	Equities
PST	Preferred stock	
NER	Non-equity reserves	
MIN	Minority interest in net assets	
TEQ	Total equity	
TLE	Total liabilities and equity	TCL+ TL+TEQ

Balance sheet Change

NI	Net income (before preferred dividends)	*SAL*	Sales	Operating items (OPI)	Income statement
NIE	Net income before extraordinary items	*CGS*	Cost of goods sold		
OCA	Other current assets	*DDA*	Depreciation, depletion and amortization		
OCF	Operating cash flows	*GRI*	Gross income		
OCL	Other current liabilities	*SAE*	Selling and administrative expenses		
OIA	Other income (expense) - after tax	*OOE*	Other operating expense		
OIE	Other income (expense) – net	*TOE*	Total operating expenses		
OIN	Other investments	*OPI*	Operating income		
OLA	Other long-term assets	*ECR*	Extraordinary credit	Non-operating items (NOI)	
OLL	Other long-term liabilities	*ECH*	Extraordinary charge		
OOC	Other operating cash flow	*IIN*	Interest income		
OOE	Other operating expense	*PEI*	Pre-tax equity interest earnings		
OPI	Operating income	*OIE*	Other income (expense) – net		
ECF	Exceptional cash flow	*EBIT*	Earnings before interest & taxes		
PAY	Supplier payments	*IED*	Interest expense on debt		
PEI	Pre-tax equity interest earnings	*ICP*	Interest capitalized		
PPE	Property plant and equipment	*PTI*	Pre-tax income		
PRC	Provision for risks and charges	*ITX*	Income taxes		
PRE	Prepaid expenses	*MIE*	Minority interest in earnings		
PST	Preferred stock	*EIE*	Equity interest earnings		
PTI	Pre-tax income	*OIA*	Other income (expense) - after tax		
REC	Customer receipts	*DOP*	Discontinued operations		
SAE	Selling and administrative expenses	*NIE*	Net income before extraordinary items		
SAL	Sales	*EGL*	Extraordinary items and gain(loss) on sale of assets		
SEQ	Shareholders' equity	*NI*	Net income (before preferred dividends)	OPI + NOI	

STD	Short-term debt	*REC*	Customer receipts	Operating	Cash flow statement
TA	Total assets	*PAY*	Supplier payments		
TCA	Total current assets	*OOC*	Other operating cash flow		
TCL	Total current liabilities	*ECF*	Exceptional cash flow		
TEQ	Total equity	*TXP*	Tax payments and provision settlements		
TL	Total liabilities	*OCF*	Operating cash flows		
TLA	Total long-term assets	*CED*	Capital expenditure (net of asset disposals)	Investing	
TLE	Total liabilities and equity	*ICF*	Investing cash flows		
TLL	Total long-term liabilities	*DIR*	Debt issues (net of debt repayments)	Financing	
TOE	Total operating expenses	*NCD*	New capital (net of dividends)		
TXP	Tax payments and provision settlements	*FCF*	Financial cash flows		

TABLES

FIGURES

CHAPTER 1- INTRODUCTION

CHAPTER 1- Introduction

The 'articulation' of financial statements is an inherent outcome of the double entry system, because every transaction that is recorded in an account will always be mirrored in one or more other accounts, as debit and credit entries. Thus, when double entry is applied consistently to all transactions, and the financial statements are fully articulated, it will be theoretically possible to calculate an item such as operating cash flow not only directly by deducting the cash outflows reported in the cash flow statement from the cash inflows, but also by differencing the revenues and expenses reported in the income statement and adjusting this for changes in the relevant line-items in the balance sheet.

The most basic acknowledgement of the articulation of financial statements in the context of cash flow prediction research is already implicit in studies such as Barth et al. (2001) and Arthur et al. (2010), who suppose that the current cash flow from operations that is reported in the cash flow statement is equal to the difference between the earnings reported in the income statement and the accruals reflected in balance sheet changes. Along similar lines, Basu (1999), Jordan and Waldron (2001) and Fischer et al. (2008), who present other versions of disaggregated cash flow, all imply articulation as they also draw on accounting line-items that are reported in the different financial statements.

With regard to predictive ability, it should be noted that each of the above studies finds the same result with respect to future operating cash flow - the disaggregated models which imply articulation are superior to the more

parsimonious aggregated models in terms of their explanatory power. The aim of the present study is to further develop research design in this area, using an advanced level of financial statement articulation based on all line-items in the income statement, the cash flow statement and the opening and closing balance sheets, and by making explicit the most basic property of double entry book-keeping in the disaggregation of financial statements items in order to predict components of future cash flow and components of future earnings. A more comprehensive definition and categorization of accruals is used in this thesis, as the complete set of financial statement items is employed both in the prediction of future cash flows and the prediction of future earnings, and in evaluating in the latter case the reliability of accruals in such prediction. It is expected *a priori* that the models presented, which by design draw on all variables disclosed in the accounts together with their double entry identity, will be superior (in terms of the explanatory power of the predictor variables) to prior models that do not consider the double entry equilibrium condition under articulation.

More specifically, the first part of this thesis considers the prediction of three components of cash flow (operating, investing and financing cash flows) based on the accounting identities governing each of the three components involved. The predictor variables in the cash flow forecasts are the revenue and cost items in the income statement, the balance sheet changes that make up the total accrual, and the changes in the components of shareholders' equity. In this respect, three different statistical models are employed: (a) multiple *OLS* regressions, (b) seemingly unrelated regressions, or *SUR*, and (c) constrained *SUR*. With the first model, the least squares regression, it is assumed that each of the linear equations

describing the above-mentioned cash flow components (operating, investing and financing) is unrelated to the other two. In the second model, these three linear cash flow equations are re-expressed as a set of simultaneous equations, related only through their correlated residuals. In the third model, the accounting relation among the predictor variables is imposed as a set of constraints, based on the double entry system.

In summary, the aim of the first part of the study is thus to consider whether imposing the double entry constraint on the simultaneous linear equations provides a superior model with which to estimate future cash flows. To this end, in addition to assessing whether the standard errors on predictor variables are lower in the case of a more powerful model (and, consequently, whether the t-statistics are higher), the following criteria are also evaluated:

- whether the model results in *reasonable signs* for coefficients, e.g. that there is a positive sign for current receivables in the prediction of one year ahead operating cash flow, given that an increase in customer receivables in the present year would be consistent with an increase in cash from customers in the following year;

- whether financial statement items that incorporate both *cash* and *accrual* information make a higher contribution to prediction (i.e. they are more significant and larger) than those which are either just cash amounts or accrual amounts, e.g. as the sales figure contains both cash and accrual information, sales may be expected to contribute more to cash flow prediction than receivables, which is based just on accrual information.

As a second step, if a model based on double entry constrained *SUR* can be identified as a superior cash flow prediction model in line with the above, then it is argued here that its demonstrated accounting validity should support the use of such a model in the prediction of future earnings. In this respect, therefore, the aim of the second part of this thesis is to compare the explanatory power of cash flows, accruals and changes in equities in the prediction of future income components, which in this case are simply the future operating income and the future non-operating income.

The motivation for the second part of the thesis is to attempt to unravel the confusion relating to the relative power of accrual versus cash flow predictors, where the emphasis appears either to be on the relative strength of accrual information (e.g.Chan et al. (2004)) or on the greater predictive power of cash flow information (Arthur et al. (2010) and Sloan (1996)). The results reported in this thesis suggest that, using two simultaneous equations based on the accounting identities for operating and non-operating income, together with double entry constraints on the predictor variables, cash flow items prove to be significantly more powerful in predicting future earnings than accruals and changes in equities.

Finally, a qualitative assessment is carried out of low and high reliability accruals in order to assess their predictive ability with respect to future earnings. Prior work by Lev and Sougiannis (1996) and Richardson et al. (2005) suggests that accruals which are *less reliable* in terms of the greater managerial discretion that they afford are expected to be *more relevant* and similarly that accruals which are *more reliable*, and thus offering less managerial discretion, will be *less relevant*.

However, contrary to the work of Richardson et al. (2005), the results presented here show that, by using the modelling framework that is built up in the first and second parts of this thesis, lower (higher) reliability accruals make a stronger (weaker) contribution to explaining future income, a finding which offers further insights into the trade-off between the reliability and relevance of accruals.

These issues can be summarized in the following three research questions that guide this thesis:

1. Does imposing the double entry constraint on accounting variables improve the usefulness of the model by comparison with other models that are unconstrained?

And, with respect to one of the key unresolved issues in accounting research:

2. Do cash flow variables estimate future income better than accruals?

And, from a qualitative viewpoint:

3. Are accruals with a high propensity to manipulation (low reliability accruals) associated with a higher degree of explanatory power and predictive ability in the prediction of future earnings by comparison with more reliable accruals?

With this in mind, and before moving to the second chapter, it is worth emphasising that the approach taken in this thesis attempts to draw together into the prediction model all accounting variables that are reported as financial statement line items, specifically in the opening and closing balance sheets, the income statement and, by a process of deduction, the reconciling cash flow

statement. The comprehensive financial package is presented for the average company - i.e. averaged over 21,698 firm-year observations relating to US companies - in Tables 1-1, 1-2 and 1-3.

It should be noted that accounting data downloaded for this purpose are always reported without sign. This implies a positive sign for assets as well as for liabilities, for instance, with a negative sign being given in the source database when a balance falls for some reason below zero, whether it is an asset line item that is reported as a net liability or a liability line item that is reported as a net asset. Therefore, in order to operationalise the double entry system among the financial statement variables, appropriate signs have to be imposed on the variables based on whether they arise from debit or credit entries in accounting journals. In the above case, for example, it might be natural for assets (debit balances) to be presented as positives, and for liabilities (credit balances) to be presented as negatives. Therefore an increase in debit items (increasing assets) takes the positive sign, and an increase in credit items (increasing liabilities or equities) takes the negative. Similarly, a decrease in debit items (decreasing assets) takes the negative sign, and a decrease in credit items (decreasing liabilities or equities) takes the positive. The result of imposing the signs on the complete set of accounting variables in this way is that the summation of all accounting variables in each financial statement is equal to zero, and it follows that the summation of the full set of accounting variables is equal to zero. In the following chapters, this issue will be explained in greater detail, and the following tables provide some initial intuition regarding the structural approach that is adopted for this research study.

Table 1-1 Balance sheet changes for the average company

This table shows the arithmetic average of the change in each balance sheet line item across 21,698 firm-year observations, for all US firms covered by the WorldScope database. The table includes all balance sheet line items reported by Worldscope, with two items in the standard WorldScope balance sheet (Deferred tax in untaxed reserves and Non-equity reserves) not being applicable in the USA, and therefore shown as zero below. The variable abbreviations used in this thesis are given in the third column. The sum of changes in all balance sheet is shown to be exactly zero at the foot of this table.

Balance sheet	Current assets	Δ*CSI*	Δ Cash and short-term investments	6,305,885.96		
		Δ*ARE*	Δ Accounts receivable	6,705,832.30		
		Δ*INV*	Δ Total inventory	4,688,963.84		
		Δ*PRE*	Δ Prepaid expenses	250,598.04		
		Δ*OCA*	Δ Other current assets	1,666,743.98		
		Δ*TCA*	Δ Total current assets		19,618,024.12	
	Long-term assets	Δ*LTR*	Δ Long-term receivables	737,382.53		
		Δ*IUS*	Δ Investment in unconsolidated subsidiaries	623,114.72		
		Δ*OIN*	Δ Other investments	812,611.89		
		Δ*PPE*	Δ Property plant and equipment	7,235,843.45		
		Δ*OLA*	Δ Other long-term assets	9,523,992.69		
		Δ*TLA*	Δ Total long term assets		18,932,945.28	
		Δ*TA*	Δ Total assets			38,550,969.40
	Current liabilities	Δ*APA*	Δ Accounts payable	-3,691,604.05		
		Δ*STD*	Δ Short-term debt	-2,841,324.27		
		Δ*ACP*	Δ Accrued payroll	-1,214,710.19		
		Δ*ITP*	Δ Income tax payable	-8,043.11		
		Δ*DPA*	Δ Dividend payable	-60,446.76		
		Δ*OCL*	Δ Other current liabilities	-4,439,444.67		
		Δ*TCL*	Δ Total current liabilities		-12,255,573.05	
	Long-term liabilities	Δ*LTD*	Δ Long-term debt	-7,240,398.80		
		Δ*PRC*	Δ Provision for risks and charges	-1,859,375.58		
		Δ*DIN*	Δ Deferred income	-778,768.71		
		Δ*DTX*	Δ Deferred tax	514,122.94		
		Δ*DTU*	Δ Deferred tax in untaxed reserves	0.00		
		Δ*OLL*	Δ Other long-term liabilities	-1,962,689.07		
		Δ*TLL*	Δ Total long-term liabilities		-11,327,109.22	
		Δ *TL*	Δ Total liabilities			-23,582,682.27
	Equities	Δ*SEQ*	Δ Shareholders' equity	-9,685,539.31		
		Δ*PST*	Δ Preferred stock	-4,928,679.83		
		Δ*NER*	Δ Non-equity reserves	0.00		
		Δ*MIN*	Δ Minority interest in net assets	-354,067.99		
		Δ*TEQ*	Δ Total equity		-14,968,287.13	
		Δ*TLE*	Δ Total liabilities and equity			-38,550,969.40
			Sum of balance sheet items			0

Table 1-2 Income statement for the average company

This table shows the arithmetic average of each income statement line item across 21,698 firm-year observations, for all US firms covered by the WorldScope database. The table includes all income statement line items reported by Worldscope, with five lines in the standard WorldScope income statement (Allocation from (to) reserves, Pre-tax equity interest earnings, Interest capitalized, Other after-tax income (expense) and Extraordinary items & gain (loss) on sale of assets) not being applicable in the USA, and therefore shown as zero below. The variable abbreviations used in this thesis are given in the third column. The sum of the income statement variables including Net income is shown to be exactly zero at the foot of this table.

Income statement	Operating income	*SAL*	Sales	-440,982,275.27		
		CGS	Cost of goods sold	283,079,122.56		
		DDA	Depreciation, depletion and amortization	16,203,502.65		
		GRI	Gross income		-141,699,650.06	
		SAE	Selling and administrative expenses#	94,054,625.03		
		OOE	Other operating expense#	5,052,213.55		
		TOE	Total operating expenses#		99,106,838.58	
		OPI	Operating income			-42,592,811.48
	Non-operating income	*ECR*	Extraordinary credits	-532,418.19		
		ECH	Extraordinary charges	7,289,417.84		
		IIN	Interest income	-2,112,771.27		
		RES	Allocation from (to) reserves	0.00		
		PEI	Pre-tax equity interest earnings	0.00		
		OIE	Other (income) expense – net	-1,150,500.56		
		EBIT	Earnings before interest & taxes		-39,099,083.66	
		IED	Interest expense on debt	4,987,181.04		
		ICP	Interest capitalized	0.00		
		PTI	Pre-tax income		-34,111,902.62	
		ITX	Income taxes	13,040,134.49		
		MIE	Minority interest in earnings	327,962.93		
		EIE	Equity interest earnings	-797,967.59		
		OIA	Other income (expense)-after tax	0.00		
		DOP	Discontinued operations	62,296.40		
		NIE	Net income before extraordinary items		-21,479,476.39	
		EGL	Extraordinary items & gain (loss) on sale of assets	0.00		
		NI	Net income (before preferred dividends)			21,479,476.40
			Sum of income statement items			0

Table 1-3 Cash flow statement for the average company

This table shows the arithmetic average of each of the cash flow statement variables used in this thesis, across the 21,698 firm-year observations for US firms that are the source for Tables 1-1 and 1-2 above. The variables included below are computed by deduction from the related income statement items and balance sheet changes, as described in greater detail in the thesis, and therefore do not take account of the amounts attributable to factors such as structural change (e.g. mergers, discontinued operations, etc.) that are normally offset in reported cash flow statements. The variable abbreviations used in this thesis are given in the third column. The sum of the cash flow statement variables (and hence the sum of the cash flows: Operating cash flow OCF, Investing cash flow ICF and Financing cash flow FCF) is shown to be exactly zero at the foot of this table.

Cash flow statement	Operating	*REC*	Customer receipts	-434,276,442.97	
		PAY	Supplier payments	376,916,397.19	
		OOC	Other operating cash flow	17,954,844.84	
		TXP	Tax payments and provision settlements	11,686,838.74	
		ECF	Exceptional cash flow	12,432,131.52	
		OCF	Operating cash flow		-15,286,230.68
	Investing	*CED*	Capital expenditure (net of asset disposals)	9,408,952.59	
		ICF	Investing cash flow		9,408,952.59
	Financing	*DIR*	Debt issues (net of debt repayments)	-901,427.34	
		NCD	New capital (net of dividends)	6,778,705.44	
		FCF	Financing cash flow		5,877,278.10
			Sum of cash flow statement items		0

CHAPTER 2- FINANCIAL STATEMENT ARTICULATION AND THE PREDICTION OF CASH FLOWS AND EARNINGS

CHAPTER 2: Financial statement articulation and the prediction of cash flows and earnings

This chapter sets out the background to the present thesis, and considers the following: (i) the concept of financial statement articulation, and its development in prior studies; (ii) the nature of accounting-based models that are used in the prediction of future cash flows and future earnings, and the role of articulated accruals in these prediction models; and (iii) recent work on the reliability of accruals and its consequences for accounting-based predictions when financial statement articulation is taken into account.

2-1 Financial statement articulation

The notion of 'articulation' in financial statements as employed in this thesis is explained clearly in the work of Mann (1984). In that study, which is normative, a worksheet is presented that demonstrates the articulating identities governing the basic financial statements, mainly with a view to bringing about a better understanding of the linkages between the beginning and ending balance sheets, the income statement and the cash flow statement. Indeed, the specific concern of Mann's study is to facilitate our appreciation of the effects of transactions on the financial statement identities of an entity, and this is done by way of a simple case study.

The worksheet prepared by Mann (1984) relates to an example company ABC, and is presented below in Table 2-1 This provides a useful starting point for

Table 2-1 Articulation among the financial statements – Mann (1984)

ABC Co. Financial Statement Data

Financial Position at December 31, 19x1				Income Statement for the year 19x2		Cash Flow Statement for the year 19x2		Financial Position at December 31, 19x2			
	Current	Long-term	Equity						Current	Long-term	Equity
Accounts receivable	200			Sales	2560	Collections from sales	2500	Accounts Receivable	260		
Accounts payable	-110			Purchases	-555	Disbursements for merchandise	-540	Accounts payable	-125		
Accrued wages	-8			Wages	-1030	Paid for wages	-1028	Accrued wages	-10		
Accrued liabilities	-11			Services	-420	Paid for services	-422	Accrued liabilities	-9		
Inventory	170			Decrease in inventory	-15			Inventory	155		
				Operating profit	540	Cash inflow from operations	510				
				Interest expense	-17	Interest paid	-17				
					523		493				
Deferred income tax		-84		Income tax expense	-130	Income tax paid	-125	Deferred income tax		-89	
					393	Net operating cash inflow	368				
Fixed assets(net)		1100		Depreciation expense	-232			Fixed assets (net)		868	
					161						
				Loss on sale of assets	-15	Cash from sale of assets	20	assets sold	-35		
					146						
						cash spent for new assets	-290	Assets acquired	290	1123	
Bonds payable		-875		Gain on retirement of bonds	21	Cash outlay to redeem bonds	-154	Bonds payable after retired	-700		
						Cash inflow from bond issue	193	Plus sold	-193	-893	
						Net capital cash outflow	-231				
						Net operating cash inflow	368				
						Net cash inflow	137				
Dividends payable	-40					Dividends paid	-150	Dividends payable	-20		
Cash in bank	76							Cash in bank	63		
Working capital	227							Working capital	314		
Retained earnings			-198					Retained earnings			-235
Capital stock			-220					Capital stock			-220
		141	418	Net income for the year	167	Decrease in cash	-13			141	455

this thesis, as it shows the rudimentary structural framework that will be developed in later chapters in order to impose articulation among the basic financial statements on the type of data that is available in large scale accounting datasets. In fact, it will be seen that this thesis develops such a structural framework on a larger scale than the outline presented in Mann (1984), incorporating every available accounting line item in one of the commercial financial datasets that is regularly employed in academic research (i.e. WorldScope), using observations for all of the US corporations that are covered by that database.

Each accounting line item in Table 2.1 is presented as a clean surplus articulation, whereby the change in the relevant balance sheet lines plus the related income statement entry reconciles to the cash flow statement entry. Nevertheless, it is evident that the signs (+ or -) on these variables are implicit and the informed reader will only become aware of these through an accounting understanding. A further point to consider is that, for each financial statement in this matrix-based worksheet, the sum of all variables is equal to zero, but this again is implicit in the presentation and requires an accounting understanding in order to fully appreciate that, appropriately signed, the sum of the entire matrix is itself zero.

In Table 2.2, which is based on Table 2.1 (with the opening and closing balance sheets rearranged to show the change during the year, and with subtotals deleted), Mann's illustrative example is reworked in order to demonstrate how such a matrix may be appropriately signed in a way that is consistent with the

underlying accounting identities. For instance, an increase in an item in the balance sheet such as Accounts receivable is presented here with a positive sign (+), and Sales in the income statement with negative sign (-), and the implied cash receipt from customers equal to the sum of these two being given a negative sign (-). By the same logic, an increase in Accounts payable is presented here with a negative sign (-) and Purchases in the income statement with positive sign (+), again with the sum of these two equalling the implied cash payment to suppliers, which thus takes the positive sign (+). Following Christodoulou and McLeay (2010), this thesis refers to the imposition of sign in this way (i.e. so that the overall accounting identity is a 'clearing identity' that sums to zero) as the 'double entry condition', as it is the fact that every transaction has a debit and a credit entry that leads to the zero sum.

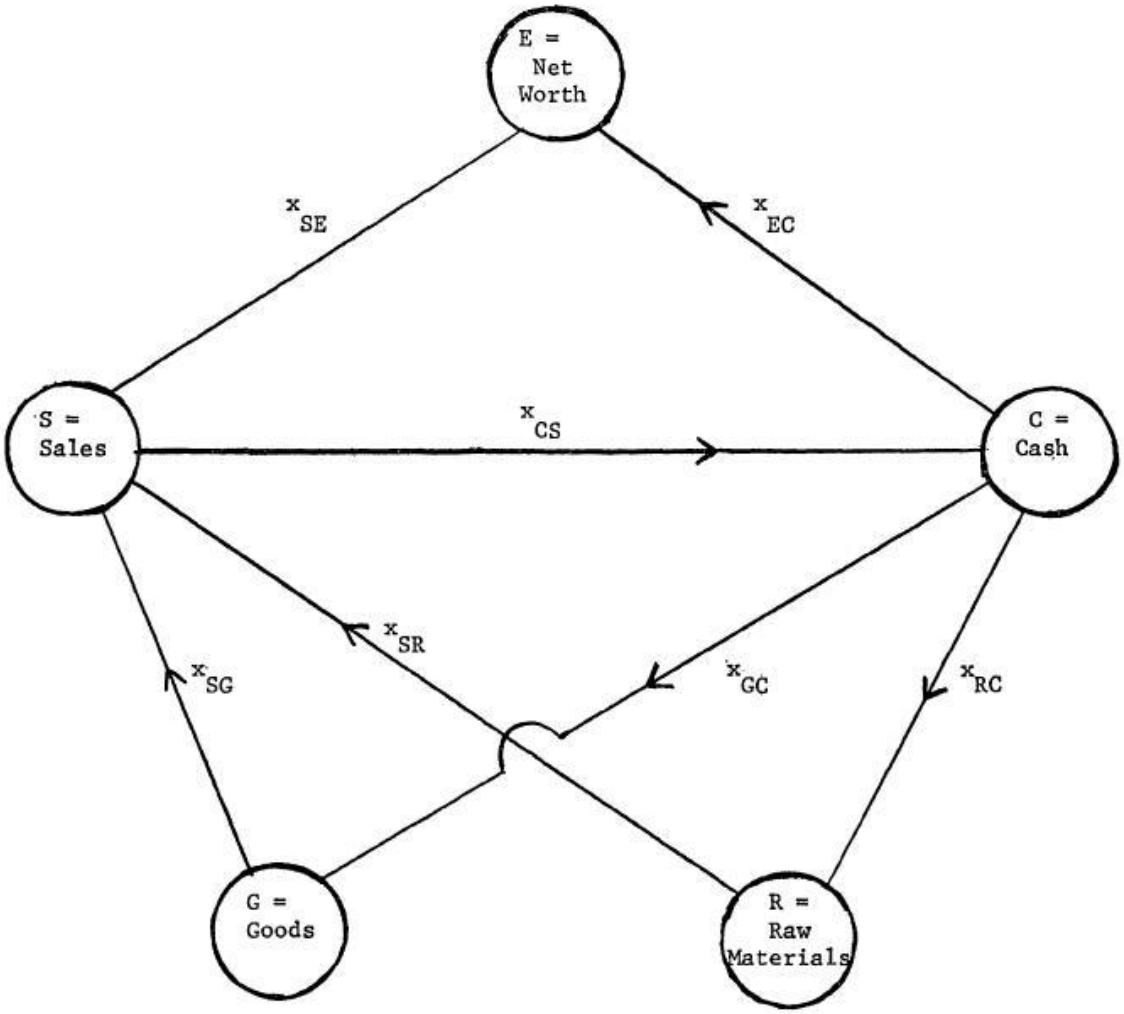
Table 2-2 Articulation among the financial statements - appropriately signed

Balance Sheet change + Income Statement = Cash Flow Statement *Assets (+) Liabilities (-) Expenses (+) Revenues (-) Profit (+) Loss (-)*								
	Balance Sheet			**Income Statement**			**Cash Flow Statement**	
Variables	**Opening**	**Closing**	**Change**					
Accounts receivable	200	260	60	Sales	-2560	=	Collections from sales	-2500
Accounts payable	-110	-125	-15	Purchases	555	=	Disbursements	540
Inventory	170	155	-15	Decrease in inventory	15			
Accrued wages	-8	-10	-2	Wages	1030	=	Paid for wages	1028
Accrued liabilities	-11	-9	2	Services	420	=	Paid for services	422
Deferred income tax	-84	-89	-5	Income tax expense	130	=	Income tax paid	125
Fixed assets(net)	1100	1123	23	Depreciation expense	232	=	Cash from sale of assets	-20
				Loss on sale of assets	15		Cash spent for new assets	290
Bonds payable	-875	-893	-18	Gain on retirement of bonds	-21	=	Cash outlay to redeem bonds	154
				Interest expense	17		Cash inflow from bond issue	-193
							Interest paid	17
Dividends payable	-40	-20	20	Net income (loss) for the year	-167	=	Dividends paid	150
Retained earnings	-198	-235	-37					
Capital stock	-220	-220	0					
Cash in bank	76	63	-13				Decrease in cash	13
Summation	**0**	**0**	**0**	**Summation**	**0**		**Summation**	**0**

Although Mann's case study illustration in *The Accounting Review* is primarily a teaching note, it nevertheless reflects some of the key ideas in accounting to be found in the influential theoretical works of the time, such as Charnes, Cooper and Ijiri (1963), and which have continued to be debated in subsequent discussion papers, e.g. Arya, Fellingham, Glover, Schroeder and Strang (2000) and Demski, Fitzgerald, Ijiri and Lin (2006).

Indeed an early attempt to explore financial statement articulation was through a linear programming application, in the Charnes et al. (1963) paper cited above. A set of constraints is employed in that study in order to show the articulation of accounting variables using the duality specification of debit and credit. The figure below demonstrates the view of these authors regarding the way in which transactions are recorded in two (or more) accounts and how this underlies their articulation application:

Figure 2-1 Association between financial variables



Source - Charnes et al. (1963)

The lines connecting the above variables indicate the accounting entries, where the two-letter suffix to each double entry 'x' represents the account debited (first letter of the suffix) and then the account credited (second letter). For instance, the line between Cash (C) and Raw Material (R) is specified as X_{RC}, which signifies purchasing materials with cash, thus increasing Raw material (R) and decreasing Cash (C). The other double entries in this figure are given below:

- X_{GC} – Purchasing other goods with cash: debit Goods (G), credit Cash (C)
- X_{CS} – Selling the company's products for cash: debit Cash (C), credit Sales (S)
- X_{SR} – Using raw materials increases the cost of goods sold: debit Sales (S), credit Raw material (R)
- X_{SG} - Using other goods also increases the cost of goods sold: debit Sales (S), credit Goods (G)
- X_{EC} – Withdrawals by owner – debit Net worth (E), credit Cash (C)
- X_{SE} – Adding the year's net profit to equity – debit sales (S), credit Net worth (E)

Structurally the above can also be expressed using a number of constraint equations in the form of zero sum 'clearing identities', as follows

$$\mathbf{0} = \boldsymbol{X_{SG} - X_{GC}}$$

$$\mathbf{0} = \boldsymbol{X_{SR} - X_{RC}}$$

$$\mathbf{0} = \boldsymbol{-X_{CS} + X_{GC} + X_{RC} + X_{SE}}$$

It can be seen, in this extract from their work, how Charnes et al. (1963) employ the structural articulation framework to describe a small set of conventional accounting variables, specifying their constraint equations on the basis of the double entry condition. Later in this thesis, it will be shown how these

'clearing identities' set to zero may be developed to describe each financial statement, and indeed the entire set of financial statements.

The above study only briefly examines some of the possible relations between mathematics and double entry. In another paper, Ellerman (1985) extends further the mathematics of the conventional accounting equation with respect to the duality principle of the double entry system. He notes that double entry bookkeeping is based on the construction of integers (positive and negative) as equivalence classes of ordered pairs of natural numbers. The ordered pairs of this construction correspond to the *T*-accounts of double entry bookkeeping. The left-hand entry in the ordered pair corresponds to the debit side of the *T*-account, and the right-hand entry to the credit side. Following Ellerman's specifications, the notation *[x//y]* will be used for a debit entry of *x* and a credit entry of *y*, as follows

$$[\boldsymbol{x//y}] = \begin{array}{c|c} \boldsymbol{Debit} & \boldsymbol{Credit} \\ \hline \boldsymbol{x} & \boldsymbol{y} \end{array}$$

Since the label '*T*-account' will be used later in a specific accounting context, the general ordered pairs *[x//y]* will be called *T*-terms; also, a *T*-term equal to the zero *T*-term *[0//0]* will be called a zero-term.

For example, *[x//x]* is a zero-term even when *x* is nonzero. That is, for any nonnegative numbers *w* and *y*,

$$\boldsymbol{w = y}$$

if and only if

$$[\boldsymbol{w//O}] + [\boldsymbol{O//y}] = [\boldsymbol{w//y}] = [\boldsymbol{0//0}]$$

In more general terms, given any equation where all numbers are nonnegative such as $w1 + \cdots + wn = y1 + \cdots + yn$, which represents n transactions, we encode each left-hand-side item as a *debit T*-term, such as *[w//O],* and we encode each right-hand-side item as a *credit T*-term, such as *[O//y]*. Then the original equation holds if and only if the sum of the encoded *T*-terms is a zero-term, i.e.

$$\boldsymbol{w1 + \cdots + wn = y1 + \cdots + yn}$$

if and only if

$$[\boldsymbol{w1\ //O}] + \cdots + [\boldsymbol{wn\ //O}] + [\boldsymbol{0//\ y1}] + \cdots + [\boldsymbol{0//\ yn}]$$

is a zero-term.

This translation or encoding of equations into zero-terms works even if the original equation contains negative numbers since any equation can be transformed to one in which all of the terms are positive by transferring the negative numbers to the other side.

Another relevant aspect of double entry bookkeeping is that transactions must be recorded in a way that maintains the integrity of the governing accounting identities that describe the financial statements, such as, for example, the balance sheet equation

$$\boldsymbol{Assets} \equiv \boldsymbol{Liabilities} + \boldsymbol{Net\,Worth}$$

That is, transactions must be recorded as valid algebraic operations which maintain this identity. As mentioned before, such an identity can be rearranged into a zero-term as:

$$\mathbf{0} \equiv \boldsymbol{Liabilities} + \boldsymbol{Net\,Worth} - \boldsymbol{Assets}$$

A further issue, which concerns the way in which accounting numbers are appropriately signed, is explored in Arya et al. (2000). The main aim of their study is to embed the double entry accounting structure in a matrix representing all of the firm's transactions. In a first step, a linear algebraic representation of the accounting process is constructed in order to specify a double entry matrix that transforms the *large* number of transactions into a *small* number of account balances. An example may be given supposing that a firm undertakes the following seven transactions:

A. Purchase inventory for cash

B. Plant acquisition for cash

C. Cash expenses

D. Cash sales

E. Cost of goods sold

F. Depreciation (product cost)

G. Depreciation (period cost)

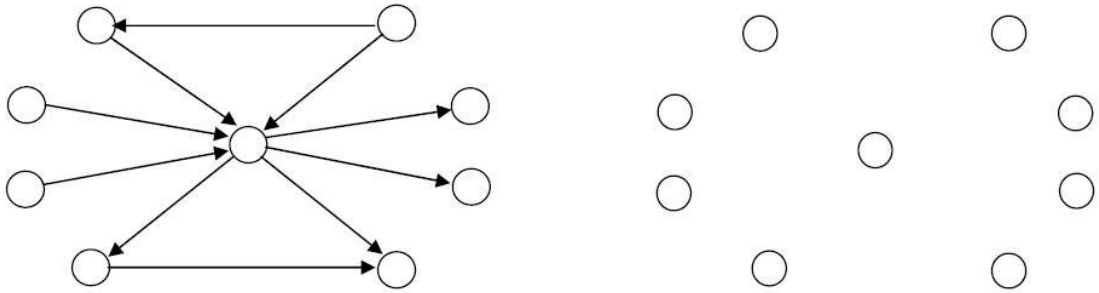
The core of the linear representation of this double entry system is a transformation matrix with *m* rows and *n* columns, where *m* is the number of accounts and *n* is the number of transactions. Similar accounting transactions aggregate so that *m*<*n*. Two non-zero entries are made in each column, which denote the accounts that are 'connected' by the transaction journal entry, using the following sign convention: debits to an account are denoted by +1, and credits are denoted by −1. For the above example, we have

	A	***B***	***C***	***D***	***E***	***F***	***G***
Cash	−1	−1	−1	1	0	0	0
Inventory	1	0	0	0	−1	1	0
PPE	0	1	0	0	0	−1	−1
Sales	0	0	0	−1	0	0	0
Cost of goods sold	0	0	0	0	1	0	0
Other expenses	0	0	1	0	0	0	1

For example, the cash sales entry in column D is shown above with a negative sign for sales (-1), as it is credited, and a positive sign for cash (+1), as it is debited. Later in this thesis, consistent with Arya et al (2000), all of the debit items are denoted by +1 and all of the credit items are denoted by -1, whether their aggregations remain in the balance sheet or are taken to the income statement.

Another issue concerning the articulation of financial statements has been raised by Demski et al. (2006) . Their study reinforces the importance of the double entry condition and, hence, the duality of accounting variables by comparing 'double entry' information with 'single-entry' information. They suggest that the difference lies in the 'connectivity' of information. For example, imagine that a firm finds an extra $10,000 of cash in the safe. As the authors comment in an aside, for anyone else this is good news, but not for accountants! Accountants cannot record the event until they find a reason why the cash increases. Indeed, 'single entry' merely records what happened but not why, and this is the importance of the 'double entry' system. In the latter case, the record of what happened in the firm (*Dr/Cr*) always has to be the reflection of why these things happened (*Cr / Dr*).

Figure 2-2 Double Entry Information (Left) and Single Entry Information (Right)



Source - Demski et al. (2006)

Connectivity is based on the accountant's judgment regarding causality. As shown in the left hand graph of Figure 2-2, if the circle in the centre represents cash, and an arrow into the circle represents an increase in cash and an arrow from

the circle represents a decrease in cash, the cash circle at the centre is always connected for any transaction to one of the four circles with arrows pointing inwards. This is not the case for the single entry diagram on the right.

In the second part of their study, Demski et al. (2006) introduce two important rules about the articulation of financial statements:

Rule 1: Beginning balances + Increases = Decreases + Ending balances

Rule 2: A change in an account cannot occur without causing a corresponding change in another account, thereby allowing the inference to jump from one account to another in the network of accounts

They point out in this context that any missing accounting number can be found by deduction using Rule 1 if three of the four numbers are already known. Note that the inference here need not stop at the first account. Having found the missing fourth number, this may make it possible for the inferential process to continue with another account, and a domino-like effect will result from combining Rules 1 and 2. For example, if say an inventory account has only two known numbers, a beginning balance and an ending balance, then as soon as an entry for the transfer out to cost of goods is recorded (i.e. a decrease), the fourth number, which in this case would be the purchases of inventory (an increase), can be deduced using Rule 1. When the purchase is entered on the debit side of the inventory account, the same amount must be entered on the credit side of accounts payable, thereby creating a chain reaction. Again, this principle, expounded in the context of double entry transaction accounting, is taken up later in this thesis with

respect to the periodic aggregates when we consider the full articulation of each financial statement.

In addition to the above issues concerning articulation, a relevant question is asked by Bahnson et al. (1996) as to whether or not financial statements actually articulate in practice. They believe that many balance sheets and cash flow statements are not articulated, in the sense that the change in current assets and liabilities in the opening and closing balance sheets seems often to be reported in cash flow statements at substantially different amounts. The findings, based on nearly 10,000 financial statements, reveal many unexplained differences between the calculated operating cash flow based on the articulation of balance sheets and income statements and the amounts actually disclosed in cash flow statements. They find that about 25 percent of the data points show small differences between the reported cash flow and the articulated cash flow, within the range from -3% to +3% of the reported cash flow. The more material differences are symmetrically distributed above or below the reported cash flow, with about 17 per cent recording a difference whose absolute value exceeds reported cash flow by 100% or more. The rest of the data points, about 58 percent of observations, lie between ±3% and ± 100% difference between reported and articulated cash flow. Their examination of company reports in these cases reveals that the companies generally fail to provide sufficient information in the annual report to fully explain the causes of all the non-articulated differences. Nevertheless, the research shows that, in some situations, reasons for significant non-articulated differences can be identified. For example, one annual report describes the management's decision to reclassify certain plant and equipment as a current asset in anticipation of disposal

in the next year. Other examples include disclosure of the management's decision to reclassify a part of the company's inventory as a long-term asset because it is being used to demonstrate the company's products, or the unusual issuance of equity shares to settle a significant portion of accounts payable. In this thesis, the large scale empirical work does not permit examination of reports in such detail, and the articulations that are imposed on the data in constructing the models described later will necessarily be measured with error in this respect, as illustrated here. However, an important insight in Bahnson et al. (1996) is that such articulation differences are distributed symmetrically, and hence are unbiased.

The final study in this area that informs the approach to financial statement articulation in this thesis is the work of Christodoulou and McLeay (2010). The key proposition of this paper lies in the treatment of financial statements as a matrix of codetermined information bounded by the double entry constraint. This again raises the issue of appropriate signing of accounting variables. In the study by Christodoulou and McLeay (2010), items that add to the accrual are deemed to take a positive sign and those that reduce the accrual a negative sign. Therefore, purchases and other expenses are positive under this system, and cash payments to suppliers are negative. In bookkeeping terms, the source entries in this case are debits [+] to the purchases account and credits [–] to the cash account respectively. It follows that sales revenues take a negative sign whilst cash receipts from customers are positive. For those accounts that are not totals of similar transactions, but instead offset debit and credit items, a net balance is reported. For example, debt issued is positive and debt repaid negative, and therefore the

net financial flow is positive if debt issues exceed debt repayments and negative if more debt is retired than raised. Similarly, when new contributions of equity capital exceed distributions, the net dividend takes a positive sign, and vice versa when there is a net return to the shareholders. Any addition to equity capital by way of profit also takes a positive sign, whilst a loss is negative.

This signing of accounting variables must be made explicit in any accounting identity, such as the cash flow equation that describes the cash flow statement, and the earnings equation that describes the income statement. Each of these two accounting identities is represented by seven variables in Christodoulou and McLeay (2010). As mentioned, the sign either is known or, as in the case of net balances, there is no fixed sign. The accounting identities are expressed as clearing identities that sum to zero in a similar way to Charnes et al. (1963), leading to the following representation where the implicit sign can be seen below each equation in the accounting logic that follows the *Cr* and *Dr* entries that make up each accounting variable:

Figure 2-3 Double entry equilibrium

Cash Flow Identity

	REC +	PAY +	SET +	TXP +	NCE +	NFF +	NDV ≡ 0
Cr [–]		Payments to Suppliers	Settlement of Provisions	Tax Payment	New Investment	Debt Repayment	Dividend Distribution
Dr [+]	Receipts from Customers		Reimbursementt of Settlements	Tax Refund	Asset Disposal	New Debt Issues	Capital Contribution

Earnings Identity

	SAL +	PUR +	PRO +	TXC +	DEP +	INT +	EAR ≡ 0
Dr [+]		Purchases	Provision Recognition	Tax Charge	Depreciation	Interest Expense	Net Profit
Cr [–]	Sales		Provision Cancellation	Tax Rebate			Net Loss

Source- Christodoulou and McLeay (2010)

Accruals may be defined on the basis of the operating components of the above accounting identities. That is to say, accruals will include the increase in trade creditor accounts when purchases *PUR* exceed supplier payments *PAY* and the increase in other operating liabilities when provision recognitions net of cancellations *PRO* exceed provision settlements net of reimbursements *SET*, less any increase in the trade debtor accounts when the sales revenues *SAL* exceed the receipts from customers *REC*, together with the depreciation charge *DEP*. Christodoulou and McLeay (2010) refer to the aggregation of these variables as the 'operating function of accruals', *OPAC*:

Operating function of accruals

$$OPAC \equiv SAL + REC + PUR + PAY + PRO + SET + DEP$$

The remaining components of the cash flow and earnings identities include the changes to the equity holders' position in the form of earnings *EAR* and net dividends *NDV*, the changes to the debt holders' position arising from net financing flows *NFF* and interest charges *INT*, and changes to the fiscal position attributable to tax charges *TXC* and tax payments *TXP*. Together with net capital expenditure *NCE*, Christodoulou and McLeay (2010) refer to this identity that brings together all of the earnings and cash flow variables that are not contributing to the operating function of accruals as the 'financing and investment function of accruals', *FIAC*:

Financing-investment function of accruals

$$FIAC \equiv NCE + TXC + TXP + INT + NFF + EAR + NDV$$

It follows therefore, for articulated financial statements, where the cash flow and earnings identities must sum to zero, that double entry bookkeeping ensures that the operating accruals OPAC is identical to the financing-investment accruals FIAC, but with opposite sign, i.e. the equilibrium condition $\text{OPAC} + \text{FIAC} \equiv 0$ must hold. Using a similarly cohesive structural framework, the present thesis builds on this approach by rearranging the accounting identity functions in order to predict future cash flows and future earnings, rather than to obtain estimates of accruals.

In fact, a principal difference between this thesis and the work of Christodoulou and McLeay (2010) concerns the variables that are estimated. In the

work of Christodoulou and McLeay (2010), a structural model is applied to the estimation of the equilibrium level of 'accruals' that is achieved through the imposition of the double entry constraint over estimation. Their study uses the reconciliation of financial statements to demonstrate how accruals are codetermined by two basic intra-firm functions, namely the operating function and the investment financing-function, which are defined by the cash flow identity and the earnings identity. The equilibrium level can be represented through the equation below:

$$\text{OPAC} + \text{FIAC} \equiv 0$$

However the first part of this thesis considers the prediction of three components of cash flow (*operating, investing and financing cash flows*) based on the accounting identities governing each of the three components involved. The predictor variables in the cash flow forecast are the revenue and cost items in the income statement, the balance sheet changes that make up the total accrual, and the changes in the components of shareholders' equity. The equilibrium level can be presented as below:

$$OPC + ICF + FCF = 0$$

Moreover, in addition to above difference, by using the above constraint, the aim of the second part of this thesis is to compare the explanatory power of cash flows, accruals and changes in equities in the prediction of future income components, which in this case are simply the future *operating income* and the future *non-operating income*. It should be noted that this issue has not been investigated in the work of Christodoulou and McLeay (2010).

As another less important difference, the number of the variables that has been used in the work of Christodoulou and McLeay (2010) is 14 variables for both operating and financing accruals. They worked mainly with aggregated variables; for instance, they have one variable representing Current assets (*CA*). However, in this thesis, consistent with the works of Basu (1999), Jordan and Waldron (2001) and Fischer et al. (2008), which all find the same result with respect to disaggregated models (suggesting therefore that articulation through disaggregation is superior to the use of more parsimonious aggregated models in terms of explanatory power), the models will be presented at the most disaggregated level possible; for instance, Current assets (*CA*) is divided here into five components, namely Cash and short-term investments (*CSI*), Accounts receivable (*ARE*), Total inventory (*INV*), Prepaid expenses (*PRE*), and Other current assets (*OCA*).

It follows that the present study aims to further develop research design in this area, using a more advanced level of financial statement articulation than Christodoulou and McLeay (2010), based on all line-items in the income statement, the cash flow statement and the opening and closing balance sheets available in the data source that is used (Worldscope). By implication, a more comprehensive definition and categorization of accruals is used in this thesis, as the *complete* set of financial statement items is employed both in the prediction of future cash flows and the prediction of future earnings, and, in the latter case, in evaluating the reliability of accruals in such prediction. Indeed, the third part of this thesis approaches this last issue from a qualitative viewpoint, attempting to answer the question as to whether accruals with a high propensity to manipulation (low

reliability accruals) are associated with a higher degree of explanatory power and predictive ability in the forecasting of future earnings by comparison with more reliable accruals. It should be noted that this issue also has not been evaluated in Christodoulou and McLeay (2010).

A final difference between this thesis and the work of Christodoulou and McLeay (2010) is the issue of appropriately signing the variables. As mentioned in their paper, items that add to the accrual are deemed to take a positive sign and those that reduce the accrual a negative sign and consequently assets are shown with the negative sign and liabilities are shown with the positive sign. However, in this study, an increase in an item in the balance sheet such as Accounts receivable (*ARE*) is presented here with a positive sign (+), and Sales (*SAL*) in the income statement with negative sign (-), and the implied Cash receipt from customers (*REC*) equal to the sum of these two being given a negative sign (-). By the same logic, an increase in Accounts payable (*APA*) is presented here with a negative sign (-) and Purchases (*PUR*) in the income statement with positive sign (+), again with the sum of these two equalling the implied Cash payment to suppliers (*PAY*), which thus takes positive sign (+). Therefore in this thesis, for variables in both balance sheet and income statement, any debit variable or increase in debit variable will take the positive sign (+) and consequently any credit variable or increase in credit variable will take a negative sign (-). Thus, the sign of the variables in the cash flow statement depends on the relative magnitude of the debit and credit variables in the other two statements (balance sheet and income statement) from which the cash flow statement variables are derived. This issue will be explained in greater detail in Tables 3-6 and 4-1.

Policy issues and implications for research design

The issue of financial statement articulation, which is central to this thesis, has been addressed as a policy issue by the main body representing financial analysts, the CFA, in pursuing their view that the current conceptual framework of accounting needs to be refined CFA (2005). An update to the conceptual framework is already under way on an international scale, in the form of a joint project by the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB). In the view of the financial analysts, one of the most controversial issues concerns the inconsistent pattern of aggregation and netting of items in financial statements, along with their obscured, even opaque, articulation, which analysts consider as a barrier to efficient analysis in forming what is described as nothing more than "*a best guess*" (CFA, 2005, p5) . The same discussion document offers twelve suggestions in order to improve the quality of financial statements, which are summarised below in Table 2-3.

Table 2-3 CFA suggestions to improve financial statement quality

1. The primary financial statements must provide the information needed by suppliers of risk capital.
2. In financial reporting, the company must be viewed from the perspective of an equity investor in the company.
3. Fair value information is the only information relevant for financial decision making.
4. Recognition and disclosure must be determined by the relevance of the information to investment decision making and not based upon measurement reliability alone.
5. All transactions and events must be recognized in the financial statements as they occur.
6. Investors' information requirements must determine the materiality threshold.
7. Financial reporting must be neutral.
8. All changes in net assets, including fair values, must appear in a single financial statement, the Statement of Change in Net Assets Available to Common Shareowners (SCNA).
9. Operating cash flows must be reported by the direct method.
10. Changes affecting each financial statement should be reported and explained on a disaggregated basis.
11. Income statement line-items must be reported by their nature or object classification, such as labor cost and utilities expense, rather than their functional classification, such as cost of goods sold (COGS) or selling, general, and administrative (SG&A) expenses.
12. Disclosures must provide added information to understand financial statement items, their measurement properties, and their risk exposures.

With regard to financial statement articulation, the eighth suggestion in Table 2-3 is particularly relevant, with the general discussion in CFA (2005) attempting to resolve the problems surrounding weak articulation. The discussion document notes:

"...although financial statements must of necessity articulate (that is, how individual items flow from one to another of the statements; for example, for credit sales, receivables in the balance sheet, cash collections in the cash flow statement, and revenues in the income statement), that articulation is opaque to investors. The reason is that the various statements do not follow a consistent structure or the same pattern of aggregation. Only when all of the changes to individual items are made clear and the measurement characteristics of items are fully disclosed will investors be able to understand a company's process of wealth generation and the prospects for their investments." (p,12)

The CFA view is that, for investors to understand (1) the economic activities reported in the financial statements, they also need to understand (2) the processes managers have used to produce the financial statements and any additional information that is disclosed. The ability of investors to distinguish between these two depends directly on the business reporting model and the quality of the measurements involved, as well as the articulation of the financial statement package as a whole, given that the financial statements are all interrelated. To understand one statement, investors must have the means to understand them all.

Currently, accounting standards permit assets and related liabilities, revenues, and expenses, as well as investing and financing cash inflows and outflows, to be reported with a high degree of aggregation, or on a netted basis, which causes important information to be obscured or lost altogether. Consistent with the work of Fischer et al. (2008), the tenth of the CFA suggestions in Table 2-3 addresses this last issue, i.e. that any factor that affects any of the financial statements should be reported on with sufficient disaggregation to allow an understanding of the linkage with the other financial statements. An example given by Fischer et al. (2008) is that

> *"In certain aggregations of working capital accounts such as changes in other current liabilities it is difficult to tell whether some of the amount might be related to investing or financing activities rather than operations."* (p,152)

In addition to Fischer et al. (2008), a paper by Largay and Weaver (2009) also addresses the non-articulation problem of financial statements, noting that further research would be required to confirm the significance of this important issue. Their point is that, even when financial statements articulate, the articulation is not sufficiently transparent to permit easy replication of the underlying computations.

In this context, further pertinent discussion concerning the articulation of financial statements can be found the work of Hribar and Collins (2002). Their work presumes that articulation breaks down when non-operating events such as reclassifications, acquisitions, divestitures, accounting changes and foreign currency translations occur. In fact, the aim of their paper is to examine the impact of measuring accruals as the change in successive balance sheet accounts, as opposed to measuring accruals directly from the statement of cash flows. Their finding is that studies using a balance sheet approach to test for earnings management are potentially contaminated by measurement error in accruals estimates as the result of the above-mentioned non-operating events. Using the balance sheet approach, accruals in their study are calculated as below:

$$ACCRUALS = (\Delta TCA - \Delta TCL - \Delta CSI + \Delta STD - DDA)$$

They note that mergers and acquisitions induce a positive bias to estimated accruals based on the the above equation, as net current assets (i.e., current assets minus current liabilities) tend to increase when one firm acquires another (e.g., accounts receivable and inventory increase to reflect the combined operations of the merged entity). However, this particular issue is not expected to place a serious

limitation on the present thesis, as the main focus here concerns the modelling of the complete set of financial statement items through the articulation process; for instance, from the balance sheet we take all long-term assets and liabilities alongside short-tem, in addition to equity items, and the related flows in the income statement; however, the above equation from Hribar and Collins (2002) just includes the change in working capital in the balance sheet. From an accounting policy perspective, with regard to the effect of acquisitions and mergers, it is also useful to make reference to *IFRS 3*, the *International Financial Reporting Standard* with regard to business combinations. Such business combination bring together separate entities or businesses into one reporting entity, with the acquirer obtaining control of one or more other businesses, the acquirees. Using the purchase method given in *IFRS* 3, the acquirer needs to transfer the assets or accept the acquiree's liabilities or issue equity shares at the acquisition date. Each one of these situations can affect other parts of the balance sheet rather than just changes in working capital; however, the equation presented by Hribar and Collins (2002) is limited to a few numbers of the variables involved. The paper by Hribar and Collins (2002) infers that events like mergers and acquisitions do not have a similar impact on the income statement; however, *IFRS* 3 clearly requires the acquirer to reassess the identification and measurement of the acquiree's identifiable assets, liabilities and contingent liabilities and the measurement of the cost of the business combination if the acquirer's interest in the net fair value of the items exceeds the cost of the combination. Any excess remaining after that reassessment must be recognised by the acquirer immediately in profit or loss. Again, for this thesis, this particular issue is not expected to create

a serious limitation as the articulation of the complete set of financial statement items allows for all movements to assets and liabilities to be accounted for.

Finally, on this issue, it is worth noting that Dechow et al. (2010) raise the issue that mergers and acquisitions result in a difference between the number reported in the statement of cash flows and the related changes in consecutive balance sheet accounts, and they suggest that these differences may be expected to have a significant effect on forecasting future cash flows or future earnings. For example, increases in inventory that are the result of an acquisition will not be reflected in changes in inventory reported in the statement of cash flows. The present thesis, however, cannot proceed by extracting merger and acquisition effects, as the necessary information is not available in the data set. Whilst recognising that future elaboration of the model proposed here might attempt to incorporate such accounting entries, the present thesis concentrates on developing the articulation framework in the context of the large-scale financial data currently available.

It is worth pointing out in this context that, with regard to the basic principles of double entry book keeping, given that a change in one variable is required to be offset by an equal change in one or more other variables, then when a complete set of financial statement items is employed, all potential accounting distortions are systematically redistributed in the form of both positive and negative entries. This issue is not recognised in published research prior to this thesis on cash flow prediction and the paper by Christodolou and McLeay (2010) on accruals estimation. In summary, the present study attempts to describe the

articulation of financial statements in order to see whether drawing such links among the different financial statements could be useful or not. Interestingly, the results reported in this thesis will show the superiority of the model that integrates such articulation of financial statements, presenting accurate signs for estimators, in addition to greater statistical efficiency.

These thesis findings, with respect to the prediction of both future cash flows and future earnings, are important not only for standard setters but also for investors and business managers. As set out in *SFAC No. 1*, a primary objective of financial reporting is to provide information useful to decision makers. Estimates and projections are potentially helpful to investors because they are the primary means for managers to convey credibly forward-looking proprietary information to them. Indeed, predicting future cash flows and future earnings are complex tasks confronting not only equity investors but also creditors and others, and accrual and cash flow accounting information are well known to be relevant to such predictions.

With respect to standard setters, the impact of the estimates underlying reporting procedures on the usefulness of financial information is an open question. The relevance of this cannot be overstated. Accounting estimates and projections underlie many Generally Accepted Accounting Principles (GAAP) and consume much of the standard-setters' time and efforts. As mentioned above, the predictive use of financial information is central to financial statement analysis and valuation, and as such is a fundamental premise of the FASB's Conceptual Framework. With respect to many of the major issues addressed by the FASB and

IASB, the 'non-articulation' problem concerning financial statements clearly occupies their ongoing attention. If this issue were not to contribute significantly to the usefulness of financial information, the efforts of accounting regulators and practitioners are misdirected, and perhaps more importantly, the resources society devotes to the generation of estimates in the process of financial statement preparation, their auditing and eventual analysis may be equally misplaced.

2-2 The prediction of future cash flows and earnings

There is considerable interest by accounting researchers in the implications of current and past earnings for the prediction of future cash flow, and it should be noted that there exists also a related but to some extent separate stream of research that asks about the implications of current and past cash flow for future earnings. Clearly, as already shown in Section 2.1 above, with an earnings identity that reflects the income statement, and a cash flow identity that reflects the cash flow statement, the issue of articulation between these financial statements must be considered in any research design that attempts an integrated approach to such predictions. In this section, the development of research with regard to both cash flow and earnings prediction is considered, followed by a discussion regarding the relevance in this respect of the way in which this thesis imposes articulation on financial statement data sets.

2-2-1 The prediction of future cash flows

One of the seminal empirical studies regarding the prediction of future cash flows is the work of Dechow et al. (1998). Their argument is that earnings, which occupies a central position in accounting as a measure of firm performance, and is widely used in share valuation, in debt contracts and to measure management performance, should be better able to explain future cash flow than is current cash flow or accruals. This inference follows from their demonstration that both cash information and accruals information are embedded in earnings, thus implying accounting identities similar to those mentioned above. They further propose that income statement items from which earnings are calculated, such as sales and cost of sales, are also likely to be good predictors of future cash flows, for two reasons. First, they are expected to be superior to cash flows as they suffer less from timing and matching problems, and superior to accruals as they also incorporate the relevant cash inflows and cash outflows. This thesis will go on to provide evidence in support of a more comprehensive theoretical position than that advanced by Dechow et al. (1998), by showing that *all* income statement items including earnings are superior to accruals.

We begin by giving precedence to an assumption about the overall revenue generating process rather than the operating cash flow generating process, because revenue contracts determine both the timing and amount of the cash inflows (and often related cash outflows) and the recognition of earnings. That is, the relevant cash inflow is informed by the revenue contract with respect to when and under what conditions the customer has to pay. These conditions determine

the pattern of the related cash receipts and so it may be said that the revenue contract is more primitive than the cash receipts. The revenue conditions also determine when a relevant future cash inflow is verifiable and thus included in earnings (along with associated relevant cash outflows).

Consistent with Ball and Watts (1972), let us assume that sales follow a random walk process, where sales for period *t*, SAL_t is given by

$$SAL_t = SAL_{t-1} + \varepsilon_t$$

where ε_t is a random variable with variance σ^2 and $COV\,(\varepsilon_t, \varepsilon_{t-\tau}) = 0$ for $|\tau| > 0$. The relation between sales and the cash flow arising from those sales is not one-to-one because some sales are made on credit. Specifically, it is assumed that proportion α of the firm's sales remains uncollected at the end of the period such that accounts receivable for period *t*, ARE_t, are

$$ARE_t = \alpha SAL_t$$

It is evident that this accounts receivable accrual incorporates future cash flow forecasts (collection of accounts receivable) into earnings.

We assume (for this discussion) that all expenses in the income statement, including purchases (*PUR*), vary with sales so that the total expense for period *t* is $(1 - \pi)SAL_t$ where π is the net profit margin on sales. It follows that Net income NI_t is πSAL_t. Therefore, if it is assumed that proportion β of the firm's purchases remains unpaid at the end of the period, accounts payable for period *t*, APA_t, is

$$APA_t = \beta[(1 - \pi)SAL_t]$$

In this way, all of the income statement variables can be written as a proportion of sales. Also all relevant cash inflows and outflows can be displayed as the proportion $(1-\alpha)$ of the firm's sales and, respectively, the proportion $(1-\beta)$ of purchases and other expenses. As income statement variables are associated with both the proportion of accruals $(\alpha, \beta, ..)$ and the proportion of cash $[(1-\alpha), (1-\beta), \dots]$, this means that both cash information and accruals information are embedded in income statement variables, and therefore these variables are more able in explaining the future performance of the company as they are associated with all cash and accrual information in compare to each one of cash or accrual information, individually.

Following on from this portrayal of the role of accruals as shifting or adjusting cash flow recognitions over time, and the arguments put forward in Dechow et al (1998) regarding the superiority of such adjusted earnings numbers over cash flow in anticipating future cash flows, Dechow and Dichev (2002) expand on these initial assertions by focusing on the *quality* of accruals. They suggest that this must also be factored into the issue of predictability, as low quality accrual information is likely to be less reliable than high quality cash flow information. For example, if the net proceeds from a receivable are less than the original estimate, then the subsequent entry records both the cash collected and the correction of the estimation error, and such estimation errors will detract from the forecasting accuracy. Thus, according to Dechow and Dichev (2002), the superiority of accruals and earnings to cash does not always hold, and a trade-off is inherent that should be recognised in building cash flow prediction models.

In another influential paper, which further develops the modelling framework first introduced in the work of Dechow et al. (1998), Barth et al. (2001) provide evidence that is consistent with the FASB's assertion that knowledge of the *components* of earnings is important for predicting future cash flows, extending in this way the understanding of the temporal relationship among accruals, cash flows and earnings. Following on from this initial work, in a study which compares the ability of current operating cash flow and earnings to predict future operating cash flow, Barth et al. (2002) show again that earnings is more able than cash flow to explain future cash flow only when the cash and accrual components of earnings are included separately in the prediction equation (i.e. the coefficients are allowed to vary across the various accruals).

Jordan, Jordan and Waldron (2001) also indicate that individual components of accrual-based earnings and cash flow provide unique information not captured in models examining aggregated accrual earnings and cash flows from operations. In related empirical work by Yanagawa (2003), on firms in the Japanese financial market, which is based on Barth et al. (2001), the main concern is again whether the various accrual components of earnings may be associated differently with future operating cash flow. Whilst this paper, and many others that follow, make no advances in theory or in research design, the results confirm in a setting other than the US that the current cash flow and accrual components that comprise the earnings identity have more predictive ability for future cash flows than aggregate earnings.

Further confirmation is also given by Cheng and Hollie (2008), who again build their study on the work of Barth et al. (2001), again examining the role of cash flow components beyond that of accrual components in predicting future cash flows. Their paper takes the research one step further however, by using a model that decomposes cash flows from operations into core and non-core cash flow components, to parallel the presentation of operating income as core and non-core in the income statement, a rudimentary form of articulation. Indeed, this approach is shown to be consistent with the view of the Special Committee on Financial Reporting of the American Institute of Certified Public Accountants (see Nusbaum and Weiss (1993)), who note that core and non-core cash flows are often separated in practice as they are considered by practitioners to be differentially persistent, which implies in turn that they should be uncoupled in predicting future cash flows, as supported by the empirical results reported by Cheng and Hollie (2008).

A further finding in this line of analysis is the work of Habib (2010) on Australian companies, which again examines empirically the relative abilities of current operating cash flow and earnings in predicting future operating cash flow. In this case, the results differ from previously, suggesting that the cash flow-based models are more accurate in predicting future operating cash flows than the earnings-based models. Although the paper has nothing to say about disaggregation, this result is worth noting because the research design introduces moderation of firm-specific cash-flow series, i.e. negative versus positive cash flow patterns and the degree of cash flow variability, together with other contextual covariates in the form of the firm size and the firm operating cycle.

Another paper to note in this context, by Waldron and Jordan (2010), investigates the comparative predictive abilities of past cash flow from operations and accrual-based net income during the economic boom leading up to the IT Bubble and the period of economic duress following the burst of that Bubble. Generally, their results indicate that cash flow outperforms net income in predicting future cash flows during these periods of economic turbulence. Additionally, the evidence reveals great variability in the predictive ability of accrual-based earnings during the time period studied, suggesting that accrual-based accounting estimates lose some of their precision during periods of extreme economic fluctuation.

Developing the modelling framework first introduced in the work of Sloan (1996), an interesting perspective is offered by Dechow et al. (2008). They focus on the fact that Sloan (1996) shows that the cash component of earnings is more persistent than the accrual component, following which they decompose the cash component into: (1) the change in the cash balance, (2) issuances/distributions to debt, and (3) issuances/distributions to equity. They find that the higher persistence of the cash component is entirely due to the subcomponent related to equity and that the other subcomponents have persistence levels almost identical to accruals. In this thesis, building on the above, and as previous work by Fischer et al. (2008) and a recent study by Lev et al. (2010) confirms the superiority of the disaggregated model to the aggregated model, the cash flow is decomposed into eight components, as listed previously in Table 1-3.

Future cash flows are at the core of asset and liability valuation rules. Thus, for example, asset impairment (*SFAS* 144) is determined by expected cash flows, and the useful lives of acquired intangibles (*SFAS* 142) are a function of future cash flows. More fundamentally, enterprise cash flows are postulated by economic theory as the major determinants of asset values. Given a certain ambiguity about the specific definition of cash flows used by investors, one of the recent studies by Lev et al. (2010) performs their test by regressing accruals and cash flows on both cash flows and future earnings. With respect to estimation of cash flows they run their tests with two widely used and frequently prescribed cash flow constructs: cash from operations (*i.e. OCF*) and free cash flow; however, this thesis decomposes free cash flow into investing cash flow (*ICF*) and financing cash flow (*FCF*). As mentioned earlier, this thesis considers the three cash flow estimating equations for *OCF, ICF* and *FCF* as a set of simultaneous equations in the prediction of future cash flows, whereby the *SUR* system assumes that all predictor variables, across all equations, are interrelated. This sort of estimation, which is attributable to Christodoulou and McLeay (2010), has not been applied in Lev et al. (2010), nor to our knowledge in any of the published studies investigating the estimation of cash flows.

A further matter to note in the work of Lev et al. (2010) is that accruals are disaggregated into just changes in working capital (ΔWC), depreciation, depletion and amortization (DDA), change in deferred tax (ΔDTX) together with other aggregated accruals. We believe this is a simplified definition of accruals, which is also confirmed by them as follows:

"... our prediction tests are based on fairly simple models. Users may be using different, more sophisticated models where estimates could prove to be useful" (p.783)

In this thesis however, consistent with the work of Fischer et al. (2008), a more comprehensive definition and categorization of accruals is used, as a *complete* set of financial statement items is employed both in the prediction of future cash flows (and in the prediction of future earnings, as discussed in the next section in Chapter 2; in the following chapter, Table 3-11 provides further details on all applied accrual variables that are used in both predictions).

The prediction tests in the work of Lev et al. (2010) indicate that accounting variables beyond those in working capital (excluding inventory) do not explain the prediction of cash flows. It will be seen later in this thesis that, when using *SUR,* all accounting variables contribute to explaining future cash flows.

The definition of accruals is still evolving. Accruals as a component of earnings are the most studied determinant of persistence. One confusing aspect of this area of research is that the definition of 'accruals' has changed over time. In early research done prior to mandatory reporting of the statement of cash flows, accruals were frequently defined as non-cash working capital and depreciation. These numbers were backed out of the statement of working capital or the balance sheet. Sloan (1996), Jones (1991), and Healy (1985) use this type of definition of accruals. Since the introduction of the statement of cash flows, accruals are more often defined as the difference between earnings and cash flows where cash flows are obtained from the statement of cash flows.

Realizing that all balance sheet accounts (except cash) are the result of the accrual accounting system, Richardson et al. (2005) provide a more comprehensive measure of accruals (intuitively, the change in net operating assets other than cash) with the change in the cash balance reflecting 'cash earnings'. This definition is more consistent with research that assumes clean surplus (often necessary for research focusing on valuation) since it is necessary to reconcile the change in equity from the balance sheet with earnings and dividends.

The motivation for the use of more exacting measures of accruals stems from the research by Hribar and Collins (2002) mentioned earlier, who suggest that their definition mitigates error induced by mergers and acquisitions. Dechow et al. (2010) also mention that mergers and acquisition and other non-cash transactions result in a difference between the numbers reported in the statement of cash flows and those implied by changes in consecutive balance sheet accounts Although, for example, increases in inventory that are the result of an acquisition will not be reflected in changes in inventory reported in the statement of cash flows, such increases could be relevant nevertheless for predicting future write-offs, which would also be reflected in the change in inventory in consecutive balance sheets.

Finally, returning to a point raised in the discussion of each of the seminal papers discussed above, by Dechow et al. (1998), Barth et al. (2001) and Dechow and Dichev (2002), it would seem reasonable to infer that income statement items from which earnings are calculated, such as sales and the cost of sales, would also be likely to be good predictors of future cash flows. The reasoning behind this assertion is that such variables are likely to be (i) superior to cash flows as they

will suffer less from timing and matching problems, and (ii) superior to accruals as they already incorporate relevant cash inflows and cash outflows. However, it is important to emphasise in concluding this review that this more general implication is not tested, neither in the papers cited here nor, to our knowledge, elsewhere. Consequently, consistent with this viewpoint, comprehensive evidence is provided in this thesis with the aim of showing that *all* income statement items including earnings are superior to accruals in cash flow forecasting.

2-2-2 The prediction of future earnings

As noted in the introduction to this section of Chapter 2, whilst the implications of current and past earnings for future cash flow have been addressed directly, by the authors cited above and many others, the reversal of this is equally well researched in a number of papers that ask about the implications of current and past cash flow for future earnings. Indeed, one of the seminal empirical studies regarding the prediction of future earnings is the work of Sloan (1996) who finds that cash is superior to accruals items in the prediction of future earnings. Sloan (1996) mentions that the extent to which current earnings performance persists into the future depends on the *relative magnitudes* of the cash and accrual components of current earnings. He proposes that accruals are less likely to persist into future earnings because of their transitory nature, and hence earnings with a larger portion of cash signify higher quality income as these earnings are more likely to persist into future periods.

Developing the modelling framework first introduced in the work of Sloan (1996), both Pfeiffer and Elgers (1999) and Barth et al. (1999) provide evidence that cash is an important indicator of future earnings, having a stronger relationship with future earnings (i.e. higher persistence) than the accrual component, and with different valuation effects as a result. A subsequent study by Chan et al. (2004) reports that aggregate future earnings will decrease by $0.046 in the next year for a $1 increase of current accruals. They note that this negative accrual effect is more significant for firms with high accruals, high price-earnings ratios and high market-to-book ratios, where earnings management is more likely to occur, and they show that incorporating the accrual effect is useful in improving the accuracy of earnings forecasts for these firms, consistent with the notion that earnings management may be the cause of such a negative relationship between current accruals and future earnings.

In another approach to the comparison between the accrual variables and cash flow variables, the work of Hirshleifer et al. (2009) shows that, in market-based earnings persistence regressions, the accrual component of earnings is less persistent than the cash flow component, with a difference in coefficients that is much larger than that in the regressions of Sloan (1996). These findings suggest that accruals and cash flows contain information about changes in the cost of capital, where accruals and cash flows are correlated with shifts in discount rates, inferring that firms manage earnings in response to market-wide undervaluation. The initial assumption in their work is based on the earnings fixation hypothesis, where naive investors fixate on earnings and fail to attend separately to the cash flow and accrual components of earnings. Consistent with the work of Sloan

(1996), where the cash flow component of earnings is a stronger predictor of future earnings than the accrual component of earnings, investors who ignore this become overly optimistic about the future performance of firms with high accruals but low cash flows, and overly pessimistic about the future prospects of firms with low accruals but high cash flows.

Further insight from market-based research is evident in the work of Barone and Magilke (2009) who qualify prior results, such as those of Collins et al. (2003) who focused exclusively on the returns to an accruals-based trading strategy. They note that, because the prior literature did not control for cash flow mispricing, all of the returns were attributed solely to accruals mispricing. Accordingly, they explicitly identify the specific source of the returns that allows us to more appropriately identify the underlying mispricing that generated those returns. The motivation for including cash flows alongside accruals in their study dates back to the work of Desai et al. (2004), which documents empirically the effect that negative correlation between accruals and cash flows has on the composition of firms and on the returns generated from independent sorts on accruals and cash flows. They note that finding a different composition of firms, along with differing return characteristics of those firms, introduces the possibility that the accrual and cash flow mispricing uncovered in Sloan (1996) derives from different firms, although it may be the case that different investor classes price these firms differently. Nevertheless, it is evident that the market-based evidence also places considerable emphasis on the separation of cash flow and accrual components of earnings.

One of the more recent studies by Dechow et al. (2008) also compared the predictive ability of current cash flows and current accruals in forecasting future earnings. Their results provide an alternative interpretation to the results in Sloan (1996), who had argued that the lower persistence of accruals arises because accruals are more subjective and hence more susceptible to measurement error. The alternative interpretation is that accruals measure changes in invested capital and that changes in invested capital are associated with diminishing marginal returns to new investment (and related overinvestment). They note that these alternative interpretations are not mutually exclusive and probably coexist. Indeed, since generally accepted accounting principles (GAAP) require the immediate impairment of unprofitable investments, one could even argue that the two interpretations are indistinguishable.

With regard to the comparison between the different components of accruals in the prediction of future earnings, one explanation for the lower persistence of the accruals component as documented by Sloan (1996) is that it is the result of measurement problems with the accounting system either because of how it reflects fundamental performance or because of the discretion allowed in the accounting system. However, other research like Dechow et al. (2010) has argued instead that the lower persistence of accruals is related to the effect of fundamental performance on persistence and in particular growth in fundamental performance.

Alongside the work on cash flow prediction by Lev et al. (2010), they also introduced into their study an examination of the usefulness of accounting

estimates for the prediction of earnings, in which context they address the issue of 'funadmental performance' by considering operating income (*OPI*) as well as net income (*NI*). These variables, *OPI* and *NOI*, are also employed in the present thesis, and are introduced as dependent variables in the *SUR* estimating procedure in order to differentiate between the company's earning power from ongoing operations and the portion of income that is derived from activities not related to core operations, given that net income reflects a combination of both a company's operating and non-operating activities.

Finally, an interesting empirical result concerning earning persistence is in the work of Ebaid (2011), on firms in the Egyptian financial markets, again comparing the cash flow component and the accrual component in earnings prediction. Although the prediction of earnings and earnings components has spawned a growing body of research in developed markets, little is known about this fundamental issue in emerging markets. The purpose of the study is primarily to examine whether earnings persistence is more attributable to cash flow or accrual component of earnings, and the conclusion is reached that the earnings of Egyptian firms are persistent, although the persistence parameters are less than those documented in developed markets. Consistent with research in developed markets, however, the findings suggest that the persistence of earnings performance is attributed more to the cash flow component than to the accrual component of earnings. Of particular interest, however, is that the findings reveal that the persistence of accruals is affected by their reliability, where less reliable accrual components have lower persistence than more reliable accrual components. It will be seen later that, consistent with Ebaid (2011), the findings

reported in this thesis also reveal that the persistence of earnings performance is attributed more to the cash flow component than to the accrual component of earnings, but our results show that less reliable accrual components have higher persistence than more reliable accrual components. This is consistent with the Statement of Financial Accounting Concepts 2 issued by the FASB, where it is assumed that less reliable accruals are more relevant and will have higher predictive ability. This issue is taken further below in section 2-3.

2-3 Reliability of accruals in the prediction of future earnings

So far three different issues have been discussed: the articulation of financial statements, the prediction of future cash flows and the prediction of future earnings. The aim of this thesis is to further develop research design in these three areas, using an advanced level of financial statement articulation and by making explicit the most basic property of double entry book-keeping in the disaggregation of financial statements items in order to predict components of future cash flow and components of future earnings. Indeed, the discussion above led to the first question this that thesis attempts to answer - whether imposing the double entry constraint on accounting variables improves the usefulness of the model in prediction of future cash flows by comparison with other models that are unconstrained? Further, with respect to one of the key unresolved issues in accounting research, a second question has also been posed - whether cash flow variables estimate future income better than accruals?

In providing the motivation for the thesis, the previous section has attempted to establish what prior research tells us might be the best accounting-based predictors of future cash flows and future earnings. Although prior accruals, earnings and cash flows are the most common predictors, there is no consensus over their relative contributions, and little attention to the underlying accounting identities that link the components of these three prominent variables. The aim then is to investigate this controversy further, and to propose and apply the innovative method mentioned above which yields consistent estimations of future earnings and cash flows, with higher precision and greater efficiency than is the case in the published results to date. As also mentioned, these improvements in terms both of explanatory power and predictive ability will be achieved by imposing constraints based on financial statement articulation and by assuming that all accounting items in the model are contemporaneously codetermined in drawing up the financial statements. The next chapters of the thesis will outline how the estimation is achieved with a system of structural regressions, and a framework of simultaneous linear equations, which will permit related stochastic accounting variables to be included on either sides of the regression, also allowing for the most basic property of accounting - double entry book-keeping - to be incorporated within the model using a set of well-specified constraints.

Beforehand, however, it is worthwhile to consider the trade-off between the reliability and relevance of accruals in such prediction, as it may be expected that a low level of accrual 'reliability' will be associated with a high level of 'relevance', as these accruals are associated with the greater managerial discretion. Consistent with the work of Sloan (1996), we may also refer to the degree of subjectivity in

accounting measurement, or its 'verifiability', when introducing the notion of 'reliability'. For instance, the Financial Accounting Standard Board's Statement of Financial Accounting Concepts describes 'verifiability' as:

"*The ability through consensus among measurers to ensure that information represents what it purports to represent or that the chosen method of measurement has been used without error or bias.* (SFAC 2, Glossary of Terms)

Likewise, in a similar vein, 'reliability' is defined as:

"*The quality of information that assures that information is reasonably free from error and bias and faithfully represents what it purports to represent."* (SFAC 2, Glossary of Terms)

Dechow and Dichev (2002) think along the same lines with regard to error and bias when they argue that the quality of accruals decreases as the magnitude of estimation error in accruals increases. That is, one role of accruals is to shift the recognition of cash flows over time so that the adjusted numbers (earnings) better measure firm performance. Accruals must therefore require assumptions and estimates of future cash flows, and hence accrual quality is positively related to earnings persistence. However, the work of Richardson et al. (2005) is already the most influential on this issue, and plays a significant role in this thesis, motivating two particular arguments. The first of these, instead of providing a traditional definition of accruals in the form of working capital plus depreciation, offers an understanding of accruals based on the non-equity and non-cash balance sheet changes, and the second builds on this clarification of the scope of accruals by examining their reliability. Their paper makes a strong link between accrual reliability and earnings persistence, believing that accruals with a lower degree of reliability are associated with the lower level of earning persistence. A

comprehensive categorization of accruals based on the balance sheet lines is presented and each category rated according to the reliability of the underlying accruals. Their empirical findings confirm that less reliable accruals lead to lower earnings persistence and that investors do not fully anticipate the lower earnings persistence, leading in their view to significant security mispricing. These results firmly suggest that there are significant costs associated with incorporating less reliable accrual information in financial statements.

Richardson et al (2005) builds nevertheless on the work of Sloan (1996), drawing a natural link between Sloan's notion of subjectivity and the well-known accounting concept of reliability. Indeed, although much prior research is about the decision-relevance of accounting numbers, these papers help to switch the attention to reliability, quantifiable in terms of error and bias, as described above. Yet it is Lev and Sougiannis (1996) who first emphasise the link between reliability and relevance, noting that high relevance is associated with low reliability, and Watts (2003) also makes this connection, noting that less verifiable and hence less reliable estimates in accounting can seriously compromise their usefulness. But it is Richardson et al (2005) who highlight the crucial trade-off between relevance and reliability, suggesting, as noted above, that the recognition of less reliable accrual estimates introduces measurement error that reduces earnings persistence leading to significant security mispricing.

FASB, in *Statement of Financial Accounting Concepts 2*, demonstrates the qualitative characteristics of accounting information such as relevance, along with reliability. Though, ideally, the choice of an accounting alternative should produce

information that is both more reliable and more relevant, it may be necessary to sacrifice some of one quality for a gain in another.

The statement notes that, to be relevant, information must be timely and it must have predictive value or feedback value or both and, as mentioned previously, to be reliable, information must have representational faithfulness and it must be verifiable and neutral.

Relevant accounting information is capable of making a difference in a decision by helping users to form predictions about the outcomes of past, present, and future events or to confirm or correct prior expectations. Information can make a difference to decisions by improving decision makers' capacities to predict or by providing feedback on earlier expectations. Usually, information does both at once, because knowledge about the outcomes of actions already taken will generally improve decision makers' abilities to predict the results of similar future actions. Without knowledge of the past, the basis for a prediction will usually be lacking. Without an interest in the future, knowledge of the past is sterile.

In a more explicit way, relevance for financial reporting can be defined as information's capacity to 'make a difference' that identifies it as relevant to a decision. Moreover, paragraph 47 of *FASB* 2, provides a definition of relevant information as pertaining to any 'event' that helps users to form predictions. Using their words:

"....'Event' is a happening of consequence to an enterprise, and in this context can mean, for example, the receipt of a sales order or a price change in something the enterprise buys or sells. 'Outcome' is the effect or result of an event or series of events

and in this context can mean, for example, that last year's profit was $X or the expectation that this year's profit will be $Y. The event in question may be a past event the outcome of which is not already known, or it may be a future event the outcome of which can only be predicted." (p.16)

Paragraph 55 highlights this issue again, directly relating 'predictive ability' to 'relevance' as follows:

"... predictive ability is an important consideration in distinguishing relevant from irrelevant accounting information." (p.17)

In this respect, this thesis is consistent with the work of Richardson et al. (2005). First, the research takes on board their assumption with respect to the trade-off between reliability and relevance. Second, the study formally models the implications of reliability for earnings persistence in the same way as these authors. However, this thesis does not follow them in their attempt to corroborate and extend the work of Sloan (1996) regarding market efficiency with respect to the information in accruals. In this respect, it can be noted, Sloan shows that investors act as if they do not anticipate the lower persistence of the accrual component of earnings, resulting in significant security mispricing, but this does not fall within the scope of the present thesis.

Richardson et al. (2005) argue that increased relevance could be at the expense of reliability and consequently less reliable accrual estimates introduce measurement error that reduces earnings persistence. Instead of 'measurement error', this thesis, consistent with FASB's *Statement of Financial Accounting Concepts 2*, assumes that an accrual with less reliability is associated with high relevance, i.e. that the more relevant the variable, the more capable it is of making

a difference to decision making and consequently the more it provides help to users to form predictions. This is precisely as discussed in paragraph 55 of the FASB's statement, that information with high relevance is associated with higher predictive ability. Such higher predictive ability for less reliable accruals will be reflected in larger and more significant coefficients by comparison with the more reliable accruals, and the difference will be confirmed statistically by a Wald linear coefficient test (this will be explained in greater detail in section 5-3 of this thesis).

With respect to the components of the net accrual, Richardson et al. (2005)[1] attribute a low level of reliability to those accruals and deferrals arising from changes in Accounts receivables, Inventories, Property, plant and equipment and Provision for risks and charges. The reasoning, and related empirical evidence, appears to be as follows:

Accounts receivable - These balances require the estimation of uncollectable items. Indeed, Dechow et al. (1998) note that these are the most common items used by management for managing earnings, for instance by exploiting trade-loading, premature revenue recognition and channel stuffing.

Inventories - This is also an item strongly associated with earnings management - see Bauwhede and Willekens (2003), for example. Earnings can be managed by applying different subjective cost flow assumptions; for instance, when prices increase during a specific period of time, changing the cost flow from FIFO to LIFO could lead to a decrease in earnings. Other accounting methods can

[1] Some further analysis in this respect has also been presented in the following conference paper: KHANSALAR, E. 2010. The reliability of accruals and the prediction of future cash flow. *European Accounting Association* (Aston)and *British Accounting Association* (Istanbul).

also cause inventories to be an unreliable item, for instance applying the lower of cost or market method, i.e. inventories should be recorded based on their historical value but if the market value is lower than the historical value, then the former is chosen. But market values are not transparent, nor accurate in most cases, as they are associated with a high level of uncertainty and consequently will provide room for management to exercise their own discretion to manipulate earnings.

Property, plant and equipment - This is another item with a low level of reliability. Holthausen (1981) mentions this as possibly the most unreliable item in the balance sheet, allowing management to exercise considerable discretion. Depreciation and amortization methods, useful life and salvage value are considered as the most unreliable factors. Furthermore, property, plant and equipment should be write-down when impaired, but the estimation of the amount of these impairments involves considerable unreliability. Asset disposal gains also provide opportunities – for example, Shah et al. (2009) note the propensity for firms to include such non-cash revenues in earnings. Capitalizing expenses also entails uncertainty – for instance, assigning software development cost to assets mainly depends on manager decisions instead of a set of transparent rules (the *Securities and Exchange Commission*, in *Accounting and Auditing Enforcement Release* 2393, point to a case of a company that significantly overstated its software assets by more than 400% by improperly capitalizing operating expenses).

Provision for risks and charges - Such provisions are set aside to cover losses or liabilities that are certain or probable, but for which the amount or timing

cannot be determined at the end of the period. For example, the WorldScope definition tells us that Provision for risks and charges includes items such as provisions for pensions, repairs and maintenance, and litigation claims. The allowances made for these provisions is supposed to reflect the best possible estimate based on the facts available at the date of preparing the financial statements. Other risks, for which a liability is only possible but not expected, are to be mentioned in the notes, but no provision is to be stated for them in the accounts. Therefore, this item too is based on a best guess, suggesting a lower level of reliability.

In the context of the modeling approach to accrual estimation that is used in this thesis, where all non-equity and non-cash balance sheet changes are involved in financial statement articulation (and not simply the above four that have been discussed at length in prior published studies), it is evident that we should consider the trade-off between reliability (estimation error and bias) and relevance (significance of the coefficient estimate) for the entire set of variables involved. Hence, in contrast to the low reliability variables identified above, certain other balance sheet items are deemed here to have a high level of reliability. These are: Short-term debt and Long-term debt, Accounts payable, Accrued payroll, Income tax payable and Other current liabilities. Such financial liabilities are valued at the present value of the future cash obligations owed by the firm, with the discount rate fixed at the issue date. Firms are not permitted to book an allowance for the anticipated non-payment of their own financial obligations, so there is relatively little subjectivity involved in the measurement of these liabilities. We therefore expect the reliability of these variables to be relatively

high. With regard to payables, Richardson et al. (2005) in their work note that, in contrast to receivables and inventory, payables can generally be measured with a high degree of reliability. Accounts payable are financial obligations to suppliers that are recorded at their face value. If a firm is to continue as a going concern, then it will typically have to pay its suppliers in full. The main subjectivity that arises in the case of payables is in estimating discounts that may be offered by suppliers. But since the amount of any discounts can typically be verified with the suppliers, there is relatively little room for error. Accrued payroll and Other current liabilities may similarly be categorized as accruals with a relatively high level of reliability.

Given the above, it follows that the remaining items on the balance sheet may be presumed to have a medium level of reliability, i.e. Deferred income, Deferred tax, Prepaid expenses, Investments in unconsolidated subsidiaries, Other investments, Other long-term assets, Other current assets, and Other long-term liabilities. In some ways, this third group arises primarily by default, as it is difficult to assign such variables either with a low level of reliability or a high level reliability.

Table 2-4 provides a summary of accruals based on the above categorisation into low, medium and high levels of reliability, based on the balance sheet changes that are key to the articulation of financial statements.

Table2-4 Accruals reliability

Variables		Reliability level
Δ Accounts receivable	Δ*ARE*	LOW
Δ Accounts payable	Δ*APA*	HIGH
Δ Total inventory	Δ*INV*	LOW
Δ Accrued payroll	Δ*ACP*	HIGH
Δ Other current assets	Δ*OCA*	MEDIUM
Δ Other current liabilities	Δ*OCL*	HIGH
Δ Deferred income	Δ*DIN*	MEDIUM
Δ Prepaid expenses	Δ*PRE*	MEDIUM
Δ Deferred tax	Δ*DTX*	MEDIUM
Δ Income tax payable	Δ*ITP*	HIGH
Δ Provision for risks and charges	Δ*PRC*	LOW
Δ Other long-term assets	Δ*OLA*	MEDIUM
Δ Other long-term liabilities	Δ*OLL*	MEDIUM
Δ Property plant and equipment	Δ*PPE*	LOW
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	MEDIUM
Δ Other investments	Δ*OIN*	MEDIUM
Δ Short-term debt	Δ*STD*	HIGH
Δ Long-term debt	Δ*LTD*	HIGH

Following this detailed discussion of the prior research that motivates this thesis and the research questions that it leads to, the next two chapters (Chapter 3 and Chapter 4) describe the data structure used in this research study and the methodology that is followed.

CHAPTER 3 - DATA STRUCTURE AND METHODOLOGY

CHAPTER 3- Data structure and methodology

This chapter is divided into four sections. First, the process of imposing signs onto accounting variables based on the double entry system is explained. Second, all equations related to cash flow prediction, including the constraint equation, are presented. Third, all equations related to earnings prediction are presented, again including the constraint equation; and, finally, a comparison is drawn between the *OLS* and *SUR* systems in the prediction of both cash flows and earnings.

3-1 Signs for variables under financial statement articulation

When downloading data from corporate datasets, all debit items as well as all credit items are presented with the same sign, which is positive (+) by default; however, in this study, following the work of Arya et al. (2000), the appropriate sign is to be allocated to the set of variables under investigation. As mentioned in the previous chapter, in the work of these authors all of the debit items are denoted by +1 and all of the credit items are denoted by -1, whether the underlying transactions remain as aggregations in the balance sheet or are taken to the income statement.

In this thesis, the main categories of variable in the balance sheet are treated as follows, on the basis of their book-keeping treatment: a positive sign is imposed on all assets, on the grounds that they are recorded on the debit side of the journal in the company's books, and consequently liabilities and equities take the negative sign, as they are recorded as credits in the journal. It follows that, as a result of the

addition of all balance sheet variables with appropriate sign, they will sum to zero overall. It is this characteristic of the underlying book-keeping that will be defined as a constraint arising from the double entry system, as the total of all debit items (+) is equal to the total of all credit items (-), and consequently they sum to zero.

The same situation can also be seen in the income statement. All revenues will be given a negative sign (as they are credit items in the journal, based on the double entry system) and all expenses a positive sign (as they are on the debit side of the journal), so that together with the net profit/loss, the income statement also sums to zero. Furthermore, in cash flow statements, the sum of all payments and receipts plus the change in cash will also be equal to zero.

This convention was adopted at the outset of this thesis, for the illustrative financial statements presented in Chapter 1 that introduced the full set of variables employed in this study in financial statement form, illustrated for the average US company, and again in the case study presented in Table 2-2 in Chapter 2, which provided an initial example of the appropriately signed articulation among the financial statements.

In the following sections, imposing the signs onto the variables in balance sheets, income statements and cash flow statements is explained in greater detail.

3-1-1 Signs for balance sheet variables

The logic of the double entry system is simple, as, for each accounting transaction, it is expected that for each entry on the debit side the same amount will also appear on the credit side, and vice-versa. In balance sheets, as stated

above, it is evident that the summation of all items on the asset side equals the summation of all liabilities and equities. Therefore, when we impose a positive sign on all assets in the balance sheet, and a negative sign on all liabilities and equities, we expect these to sum to zero. This was illustrated in Table 1-1 in Chapter 1, where the value of zero is shown at the end of the balance sheet. Below, in Tables 3-1 and 3-2, the same balance sheet items are presented again, grouped according to the requisite sign.

Table 3-1 provides a list of the accounting variables for which the positive sign needs to be imposed:

Table 3-1 Balance sheet items requiring an imposed positive sign

Variable	Abbreviation	Imposed sign
Short-term assets		
Cash and short-term investments	*CSI*	+
Accounts receivable	*ARE*	+
Total inventory	*INV*	+
Prepaid expenses	*PRE*	+
Other current assets	*OCA*	+
Long-term assets		
Long-term receivables	*LTR*	+
Investment in unconsolidated subsidiaries	*IUS*	+
Other investments	*OIN*	+
Property plant and equipment	*PPE*	+
Other long-term assets	*OLA*	+

Table 3-1 is divided into two parts. The first section lists all of the variables used in this thesis that may be described as short-term assets, and with long-term assets listed below. For all of these items, increases from one balance sheet date to

the next, i.e. debit entries in the accounting journal, will also show a positive sign. It is the imposition of sign on such balance sheet changes that comprises the double entry constraint operationalised in this thesis.

To complete the constraint, it is evident that the opposite and hence negative sign must be imposed on credit entries when they are accumulated in the balance sheet, i.e. the liabilities and equities. Table 3-2 summarises the allocation of the negative sign to the set of these variables used in this thesis.

Table 3-2 Balance sheet items requiring an imposed negative sign

Variable	Abbreviation	Imposed sign
Short-term liabilities		
Accounts payable	*APA*	-
Short-term debt	*STD*	-
Accrued payroll	*ACP*	-
Income tax payable	*ITP*	-
Dividend payable	*DPA*	-
Other current liabilities	*OCL*	-
Long-term liabilities		
Long-term debt	*LTD*	-
Provision for risks and charges	*PRC*	-
Deferred income	*DIN*	-
Deferred tax	*DTX*	-
Other long-term liabilities	*OLL*	-
Equities		
Shareholders' equity	*SEQ*	-
Preferred stock	*PST*	-
Non-equity reserves	*NER*	-
Minority interest in net assets	*MIN*	-

In Table 3-2, the credit items in the balance sheet have been divided into three groups: short-term and long-term liabilities, and equities. Based on the same

logic as before, these items are accumulated from initial transactions on the credit side of the accounting journal, and hence the negative sign which is imposed on them in formulating the double entry constraint.

As demonstrated, allocating signs to the balance sheet items is not that difficult as the statement is divided into two, where the debit and credit sides are self-evident, and balance sheets conventionally group items into assets and liabilities. However, the next statement for which signs are to be imposed is the income statement, where we cannot follow the simple procedure that was applied to the balance sheet, as the income statement is a mixture of debit and credit items. For instance, sales (SAL) may be categorized as a credit item, but the following item – the cost of goods sold (CGS) – is the opposite and thus in total would be categorized as a debit item. Nevertheless, the logic here is the same as in the work of Arya et al. (2000), whereby a negative sign will be imposed onto all credit items in the income statement, and a positive sign on all debit items.

3-1-2 Signs for income statement variables

In this case, we look first at revenue items, which are credit entries in the accounting journal. For the income statement items listed in Table 3-3, a *negative* sign will be imposed upon them, given that they are credits.

Table 3-3 Income statement items requiring an imposed negative sign

Variable	Abbreviation	Imposed sign
Sales	*SAL*	**-**
Extraordinary credits	*ECR*	**-**
Interest income	*IIN*	**-**
Pre-tax equity interest earnings#	*PEI*	**-**
Other income (expense) – net#	*OIE*	**-**
Equity interest earnings	*EIE*	**-**
Other income (expense)-after tax#	*OIA*	**-**
Discontinued operations#	*DOP*	**-**
Extraordinary items & gain (loss) on sale of assets#	*EGL*	**-**

***#** Items that are presented under both negative and positive signs in the WorldScope database*

Note that, while all of these items are deemed to be negative for this purpose, at least by default, five of these variables may in fact be disclosed in the source data with either sign, because they are in fact accumulated as the net of specific revenue items and expense items. These variables are:

Pre-tax equity interest earnings (*PEI)*

Other income (expense) – net (*OIE)*

Other income (expense) – after tax *(OIA)*

Discontinued operations (*DOP*)

Extraordinary items and gain (loss) on sale of assets (*EGL*)

For such variables, the negative sign is imposed where net revenue is recorded (which reflects the default interpretation in accordance with the logic applied in

the reported financial statements and the WorldScope summaries), and a positive sign if the reported amount is a net expense.

In Table 3-4, the debit items in the income statement are listed, showing in each case that a positive sign is to be imposed upon them.

Table 3-4 Income statement items requiring an imposed positive sign

Variable	Abbreviation	Imposed sign
Cost of goods sold	*CGS*	+
Depreciation, depletion and amortisation	*DDA*	+
Selling and administrative expenses	*SAE*	+
Other operating expenses	*OOE*	+
Extraordinary charges	*ECH*	+
Interest expenses on debt	*IED*	+
Interest capitalised	*ICP*	+
Minority interest on earnings	*MIE*	+
Income taxes	*ITX*	+
Net income (before preferred dividends) #	*NI*	+

 # Items that are presented with both negative and positive signs in the WorldScope database

Here, it should be noted that, whereas accounting duality leads readily to the full set of items reported in a published balance sheet summing together to zero when appropriately signed, this is not how a published income statement is usually envisaged, given the presentation of net income (*NI*) as a bottom line difference. What we need to do is to overlook the conventional way in which we see the income statement, and to include the value of net income (*NI*) in the summation process. Therefore, in Table 3-4, where the last row is allocated to the net income (*NI*), the imposed sign for net income (*NI*) is again changeable,

whereby a net loss bears a negative sign and net income a positive sign. Using this approach, the overall result for the summation of all income statement items will be equal to zero.

Up to this point, by imposing positive and negative signs onto debit and credit items respectively, the summation of each balance sheet and each income statement will equal zero, leading to the double entry constraint on financial statements, as mentioned above. Later, the implications of this constraint will be discussed in greater detail.

3-1-3 Signs for cash flow statement variables

In this thesis, *seven* different variables are created in order to prepare a cash flow statement. The method used to compute the values for these seven variables is similar to using the direct method of cash flow calculation, with the sign of each variable in the cash flow statement depending on its covariates in other financial statements. Following on with the earlier logic, cash receipts are given a negative sign and payments a positive sign. For instance, the sign of customer receipts is negative, resulting from of its articulation with sales (*Cr*) and the change in receivables (*Dr*). In the following sections this issue will be explained in greater detail. The cash flow statement is presented in Table 3-5, completing the articulation process,

Table 3-5 Cash flow statement item signs

Cash flow statement	Operating	*REC*	Customer receipts	-469,519,576	
		PAY	Supplier payments	404,086,054	
		OOC	Other operating cash flows	35,979,100	
		TXP	Tax payments and provision	11,500,591	
		OCF	**Operating cash flow**		**-17,953,830**
	Investing	*CED*	Capital expenditure (net of asset disposals)	8,275,224	
		ICF	**Investing cash flow**		**8,275,224**
	Financing	*DIR*	Debt issues (net of debt	4,662,065	
		NCD	New capital (net of dividends)	5,016,541	
		FCF	**Financing cash flow**		**9,678,606**
			Sum of cash flow statement		**0**

Later in this thesis, a system of structural regressions is introduced for estimating future values of the operating cash flow *(OCF)*, investing cash flow *(ICF)* and financing cash flow *(FCF)* variables in the above statement, an estimation that is based on the treatment of the entire financial statements as a single matrix of endogenous information. In order to achieve consistent estimation of future cash flows, in addition to the three simultaneous equations for operating, investing and financing cash flows, a fourth equation is added to the set of equations governing the matrix of endogenous variables, which is the constraint equation. This reflects the requirement that the summation of all predictor variables based on the full set of line-items in the opening and closing balance sheets, the income statement and the cash flow statement has to be set to zero, making explicit the double entry book-keeping condition discussed above, requiring an advanced level of financial statement articulation.

At present, however, financial statements in corporate databases are more likely to face the problem of *non-articulation*. For instance, CFA (2005, p 12) note that although financial statements must of necessity articulate in a manner that reveals how individual items flow from one to another of the statements, that articulation is argued to be relatively opaque to investors, often due to structural changes such as mergers and acquisitions, and this detail is not generally available in large corporate databases. Note that therefore, in order to resolve this problem for the purposes of the modeling in this thesis, and recognising the limitations that arise, the cash flow statement in this study is created in a way that at least makes us able to see how individual items flow from one to another of the statements, which could not be observed just by using reported cash flows in databases. This issue will be discussed again later in chapter 3.

3-1-4 A summary of imposed signs

The final table gives a summary of all signs for all financial statement variables, for reference specifying also whether they are debit (*Dr*) or credit (*Cr*). As can be seen, the first two main columns concern the balance sheet and the income statement, and the signs imposed onto their variables are based on whether the underlying transaction accounting is debit (+) or credit (-). The last column contains the cash flow statement, which is computed from the other two statements, and, as can be seen, all receipts are shown as negative, while all payments are shown as positive.

Table 3-6 A summary of sign allocation

Balance sheet				***Income statement***				***Cash flow statement***		
Variable		***Dr/Cr***	**Sign**	**Variable**		***Dr/Cr***	**Sign**	**Variables**		**Sign**
Cash and short-term investments	*CSI*	*Dr*	**+**	*Sales*	*SAL*	*Cr*	**-**	Customer receipts	*REC*	Receipt (**-**)
Accounts receivable	*ARE*	*Dr*	**+**	*Cost of goods sold*	*CGS*	*Dr*	**+**	Supplier payments	*PAY*	Payment (**+**)
Total inventory	*INV*	*Dr*	**+**	*Depreciation, depletion and amortisation*	*DDA*	*Dr*	**+**	Other operating cash flows	*OOC*	Receipt (**-**) Payment (**+**)
Prepaid expenses	*PRE*	*Dr*	**+**	*Selling and administrative expenses*	*SAE*	*Dr*	**+**	Tax payments and provision settlements (net of refunds and cancellations)	*TXP*	Receipt (**-**) Payment (**+**)
Other current assets	*OCA*	*Dr*	**+**	*Other operating expenses*	*OOE*	*Dr*	**+**	Capital expenditure (net of asset disposals)	*CED*	Receipt (**-**) Payment (**+**)
Long-term receivables	*LTR*	*Dr*	**+**	*Extraordinary credits*	*ECR*	*Cr*	**-**	Debt issues (net of debt repayments)	*DIR*	Receipt (**-**) Payment (**+**)
Investment in unconsolidated subsidiaries	*IUS*	*Dr*	**+**	*Extraordinary charges*	*ECH*	*Dr*	**+**	New capital (net of dividends)	*NCD*	Receipt (**-**) Payment (**+**)
Other investments	*OIN*	*Dr*	**+**	*Interest income*	*IIN*	*Cr*	**-**			
Property plant and equipment	*PPE*	*Dr*	**+**	*Pre-tax equity interest earnings #*	*PEI*	*Cr*	**-**			
Other long-term assets	*OLA*	*Dr*	**+**	*Other income (expenses) – net #*	*OIE*	*Dr/Cr*	**-**			
Accounts payable	*APA*	*Cr*	**-**	*Interest expense on debt*	*IED*	*Dr*	**+**			
Short-term debt	*STD*	*Cr*	**-**	*Interest capitalised*	*ICP*	*Dr*	**+**			
Accrued payroll	*ACP*	*Cr*	**-**	*Income taxes*	*ITX*	*Dr*	**+**			
Income tax payable	*ITP*	*Cr*	**-**	*Minority interest in earnings*	*MIE*	*Dr*	**+**			
Dividend payable	*DPA*	*Cr*	**-**	*Equity interest earnings*	*EIE*	*Cr*	**-**			
Other current liabilities	*OCL*	*Cr*	**-**	*Other income (expenses) after tax #*	*OIA*	*Dr/Cr*	**-**			
Long-term debt	*LTD*	*Cr*	**-**	*Discontinued operations #*	*DOP*	*Dr/Cr*	**-**			
Provision for risks and charges	*PRC*	*Cr*	**-**	*Extraordinary items & gain (loss) on sale of assets #*	*EGL*	*Dr/Cr*	**-**			
Deferred income	*DIN*	*Cr*	**-**	*Net income (before preferred dividends) #*	*NI*	*Dr/Cr*	**+**			
Deferred tax	*DTX*	*Cr*	**-**							
Other long-term liabilities	*OLL*	*Cr*	**-**							
Shareholders' equity	*SEQ*	*Cr*	**-**							
Preferred stock	*PST*	*Cr*	**-**							
Non-equity reserves	*NER*	*Cr*	**-**							
Minority interest in net assets	*MIN*	*Cr*	**-**							

 # Items that are presented with both negative and positive signs in the WorldScope database

Building on Table 3-6 above, we will be able to draw up a more comprehensive table that articulates all variables from all financial statements, including their imposed signs, consequently showing the double entry constraint, either for each financial statement (vertically, as previously in Table 3.5, for instance, where the sum of the cash flow statement is zero), or for every single line item (i.e., horizontally, relating each cash flow statement variable to the corresponding balance sheet and income statement line items). This development is discussed again later in the thesis, after first considering an example of the basic approach to sign imposition outlined above.

3-1-5 An example: Petmed Express Company

Following the discussion in the previous section, demonstrating the procedure of imposing signs onto the variables, this section looks at the example of *Petmed Express Company*. This company has been chosen as an example because their total assets (*TA*) are close to the average of total assets for the US market. Table 3-7 gives a brief introduction to the company.

Table 3-7 Case study - Petmed Express Company

Name	***Year***	***Industry***	***Supersector***	***Sector***	***Subsector***
Petmed Express Inc.	2008	Consumer Services	Retail	Food & Drug Retailers	Drug Retailers

In Tables 3-7 and 3-8 below, the balance sheet and income statement for this company are presented in order to show with a specific case study how signs are imposed on the debit (*Dr*) and credit (*Cr*) variables. As can be seen, the structure of the balance sheet facilitates this exercise. The first part of the balance sheet gives the assets (debit), upon which a positive sign (+) is to be imposed, and the second part the liabilities and equities (credit), on which a negative sign (-) will be imposed. In the income statement, on the other hand, as mentioned earlier, the debit and credit items are not separated, and signs are imposed based on whether the underlying accounting flow is expense (debit) or revenue (credit). The cash flow statement is not presented here, as the articulation of financial statements that shows how we can get cash flow statement variables will be presented in more detail later in the thesis.

Table 3-8 Balance sheet for Petmed Express Company

				Dr/Cr	Imposed sign			
Balance sheet	Current assets	*CSI*	Cash and short-term investments	**Dr**	+	12,602,681		
		ARE	Accounts receivable	**Dr**	+	792,585		
		INV	Total inventory	**Dr**	+	9,011,094		
		PRE	Prepaid expenses	**Dr**	+	0		
		OCA	Other current assets	**Dr**	+	348,105		
		TCA	**Total current assets**				**22,754,465**	
	Long term assets	*LTR*	Long-term receivables	**Dr**	+	0		
		IUS	Investment in unconsolidated subsidiaries	**Dr**	+	0		
		OIN	Other investments	**Dr**	+	12,447,799		
		PPE	Property plant and equipment	**Dr**	+	957,632		
		OLA	Other long-term assets	**Dr**	+	183,648		
		TLA	**Total long term assets**				**13,589,079**	
		TA	**Total assets**					**36,343,544**
	Current liabilities	*APA*	Accounts payable	**Cr**	-	-2,193,094		
		STD	Short-term debt	**Cr**	-	0		
		ACP	Accrued payroll	**Cr**	-	-112,834		
		ITP	Income tax payable	**Cr**	-	-93,204		
		DPA	Dividend payable	**Cr**	-	0		
		OCL	Other current liabilities	**Cr**	-	-831,395		
		TCL	**Total current liabilities**				**-3,230,527**	
	Long term liabilities	*LTD*	Long-term debt	**Cr**	-	0		
		PRC	Provision for risks and charges	**Cr**	-	0		
		DIN	Deferred income	**Cr**	-	0		
		DTX	Deferred tax	**Cr**	-	614,769		
		DTU	Deferred tax in untaxed reserves	**Cr**	-	0		
		OLL	Other long-term liabilities	**Cr**	-	0		
		TLL	**Total long-term liabilities**				**614,769**	
		TL	**Total liabilities**					**-2,615,758**
	Equities	*SEQ*	Shareholders' equity	**Cr**	-	-33,722,755		
		PST	Preferred stock	**Cr**	-	-5,031		
		NER	Non-equity reserves	**Cr**	-	0		
		MIN	Minority interest in net assets	**Cr**	-	0		
		TEQ	**Total equity**				**-33,727,786**	
		TLE	**Total liabilities and equity**					**-36,343,544**
			Sum of balance sheet items					***0***

Table 3-9 Income statement for Petmed Express Company

					Dr/Cr	*Imposed sign*			
Income statement	Operating income	*SAL*	Sales		***Cr***	**-**	-94,760,488		
		CGS	Cost of goods sold		***Dr***	**+**	57,420,092		
		DDA	Depreciation, depletion and amortisation		***Dr***	**+**	296,850		
		GRI	**Gross income**					**-37,043,546**	
		SAE	Selling and administrative expenses		***Dr***	**+**	22,763,489		
		OOE	Other operating expenses		***Dr***	**+**	1		
		TOE	**Total operating expenses**					**22,763,490**	
		OPI	**Operating income**						**-14,280,056**
	Non-operating income	*ECR*	Extraordinary credits		***Cr***	**-**	0		
		ECH	Extraordinary charges		***Dr***	**+**	194,214		
		IIN	Interest income		***Cr***	**-**	-891,398		
		PEI	Pre-tax equity interest earnings				0		
		OIE	Other (income) expenses – net		***Dr/Cr***	**-**	-323,965		
		EBIT	**Earnings before interest & taxes**					**-15,301,205**	
		IED	Interest expenses on debt		***Dr***	**+**	0		
		ICP	Interest capitalised		***Dr***	**+**	0		
		PTI	**Pre-tax income**					**-15,301,205**	
		ITX	Income taxes		***Dr***	**+**	5,227,127		
		MIE	Minority interest in earnings		***Dr***	**+**	0		
		EIE	Equity interest earnings		***Cr***	**-**	0		
		OIA	Other income (expense)-after tax		***Dr/Cr***	**-**	0		
		DOP	Discontinued operations		***Dr/Cr***	**-**	0		
		NIE	**Net income before extraordinary items**					**-10,074,078**	
		EGL	Extraordinary items & gain (loss) on sale of assets		***Dr/Cr***	**-**	0		
		NI	**Net income (before preferred dividends)**		***Dr/Cr***	**+**			**10,074,078**
			Sum of income statement items						***0***

3-2 The double entry constraint in cash flow prediction

In the previous section, the imposition of signs onto variables was demonstrated, on the basis of whether they are debit or credit. An example was also given of a specific company in order to show the procedure of allocating signs to each financial statement variable. Following the works of Charnes et al. (1963), Ellerman (1985), Demski et al. (2006) and Christodoulou and McLeay (2010), as explained in Chapter 2, the double entry constraint has been defined for each statement by showing the zero summation at end of that statement. The aim of this section is to demonstrate the double entry constraint, not just by illustrating it in the context of a financial statement formats, but also by writing an an accounting identity for each financial statement and consequently showing the zero sum at the end of such equations. The balance sheet is the first statement that will be demonstrated in this section, while the income statement and cash flow statement will be discussed later.

3-2-1 Articulation

The balance sheet

In this section, the aim is to set down an equation showing the double entry constraint in the balance sheet. To begin with, it is worth considering the most obvious accounting identity governing the duality of accounts, where assets is equal to the sum of liabilities and equities. This can be written as $TA \equiv TL + TEQ$, which can be rearranged as the following 'clearing identity':

$$TA - TL - TEQ \equiv 0 \qquad (1)$$

It is obvious now, from the above, how the convention adopted in this thesis is reflected in the governing identities, with assets being positively signed whereas liabilities and equities are negatively signed.

In addition, it is evident that, by allowing for short-term and long-term components of both assets and liabilities in accordance with the groupings presented in the previous section of this chapter, the fundamental accounting identity in (1) may then be extended as

$$(TCA + TLA) - (TCL + TLL) - TEQ \equiv 0 \qquad (2)$$

By extending the equation further to incorporate all of the line items in the balance sheet, as presented in the tables above, the accounting identity can be fully specified as

$$\big((CSI + ARE + INV + PRE + OCA) + (PPE + LTR + IUS + OIN + OLA)\big) - ((APA + STD + ACP + ITP + DPA + OCL) + (LTD + PRC + DIN + DTX + DTU + OLL) - (SEQ + PST + NER + MIN) \equiv 0 \qquad (3)$$

The double entry constraint for the full set of variables is implicit in the above equation, given that the sum of these variables must equal zero.

The income statement

Disregarding the imposition of signs in the first instance, the equation below shows a simple decomposition of the income statement, where operating income plus non-operating income must add up to net income, i.e. $OPI + NOI \equiv NI$. This in turn may be rearranged to provide the clearing identity in the form that shall be adopted to describe the income statement, i.e.

$$NI - OPI - NOI \equiv 0 \qquad (4)$$

This equation can be extended to involve all of the income statement variables used in this research study, as presented in the tables above, where $OPI = SAL - CGS - DDA - SAE - OOE$, and $NOI = ECR - ECH + INN + PEI + OIE - IED - ICP - ITX - MIE + EIE + OIA + DOP + EGL$. By now imposing appropriate signs onto the variables in the way that was described in the previous section of this chapter, on the basis of debit (+) and credit (-), and recalling that net income (*NI*) takes a positive sign when a profit or a negative sign when a loss, and must therefore be interpreted with care, the income statement clearing identity can be more appropriately written as

$$((-SAL + CGS + DDA + SAE + OOE) + (-ECR + ECH - INN - PEI - OIE + IED + ICP + ITX + MIE - EIE - OIA - DOP - EGL) + NI \equiv 0. \qquad (5)$$

Again, for the income statement this time, the double entry constraint is evident in the signs used in the clearing identity, where that the sum of all variables is equal to zero.

The cash flow statement

As mentioned earlier, instead of using the *reported* cash flows from the WorldScope database in the estimation process, it is possible to determine the articulated values of all of the line-items in the cash flow statement using the information in the other two financial statements. To do this, the articulation is assembled from the income statement and the *change in* balance sheet items, thus deriving the values for the variables in the cash flow statement. In predicting future cash flows in this thesis, seven cash flow variables are used: four associated with operating activities, one with investing activities and two with financing activities. The aim of this section, in addition to providing a formal demonstration of the double entry constraint for the cash flow statement, is to also place the constraint on each of these seven line-items. In the following section, the accounting identity for each line-item will be explained.

- ***Sales and receipts from customers***

Following the first rule presented by Demski et al. (2006) in Chapter 2, which indicates that the sum of *beginning balances* plus *increases* is equal to

decreases plus *ending balances*, any missing accounting number can be found by deduction if three of the four numbers are already known.

Therefore, following this logic, the first variable in the articulated cash flow statement, customer receipts (REC), can be calculated as the change in accounts receivable (ΔARE) between the opening and closing balance sheets plus the sales (SAL) from the income statement. Therefore, this accounting identity is written as:

$$\Delta ARE + SAL = REC \tag{6}$$

Before Demski et al. (2006), the above formula is mentioned by Mann (1984) in Chapter 2 as a worksheet for an example company ABC that is imposing articulation between the financial statements.

Note that, with this formula, the sign is already imposed on each variable. For instance, we know from the discussion above that the sign of sales (SAL) will be negative and the sign of Δ accounts receivable (ΔARE) will depend on whether there has been an increase (positive) or a decrease (negative).

The double entry constraint can be shown more readily by taking the right hand side variable to the left. Thus, by placing customer receipts (REC) on the left, the accounting equation will appear as a clearing identity, as follows:

$$\Delta ARE + SAL - REC = 0 \tag{7}$$

As can be seen, the sum of all variables is now equal to zero. That is why the calculated cash flow is preferred to the reported cash flow for use in this thesis. As

was explained in the previous chapter, the current conceptual framework of accounting needs to be refined, as the CFA (2005, p 12) notes that although financial statements must of necessity articulate how individual items flow from one to another of the statements, that articulation is opaque to investors. Therefore, by using the reported cash flow, the double entry constraint would not be observed, whereas the main aim of this study is to use constraint equations in addition to estimation equations to provide an unbiased estimation of a company's future performance.

- ***Purchases and payments to suppliers***

The second variable created in the cash flow statement after customer receipts (REC) is supplier payments (PAY), calculated via the second line-item. Three variables are chosen from the change in balance sheets that are associated with supplier payments (PAY) in the cash flow statement. These are: change in accounts payable (ΔAPA), change in total inventory (ΔINV) and change in accrued payroll (ΔACP). In addition, two items are chosen from the income statement that are related to the previously mentioned items in the cash flow statement. These are: cost of goods sold (CGS) and selling and administrative expenses (SAE). All these variables are applied in the equation below, creating the second variable of cash flow.

$$(\Delta APA + \Delta INV + \Delta ACP) + (CGS + SAE) = PAY \quad (8)$$

By replacing the location of supplier payment (PAY) by the variables on the left, the double entry constraint can be shown as follows:

$$(\Delta APA + \Delta INV + \Delta ACP) + (CGS + SAE) - PAY = 0 \quad (9)$$

- ***Other items in the income statement and other payments and receipts***

The third relevant item in the cash flow statement is other operating cash flow (OOC), which is the result of incorporating several items from the change in balance sheets and the income statement. Change in other current liability (ΔOCL), change in other current assets (ΔOCA), change in prepaid expenses (ΔPRE), change in differed income (ΔDIF), change in other long-term assets (ΔOLA) and change in other long-term liability (ΔOLL) are the items from changes in balance sheets, and other operating expenses (OOE), depreciation, depletion and amortisation (DDA), discontinued operations (DOP), extraordinary credit (ECR), extraordinary charges (ECH), other income (expenses) after tax (OIA), equity interest earnings (EIE) and other income (expense) net (OIE) are the items from the income statement. The equation below shows how other operating cash flow (OOC) is calculated.

$$(\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN + \Delta OLA + \Delta OLL) + (OOE + DDA + DOP + ECR + ECH + OIA + EIE + OIE = OOC \quad (10)$$

As before, in order to show the double entry constraint, the location of other operating cash flow needs to be changed, as below:

$$(\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN + \Delta OLA + \Delta OLL) + (OOE + DDA + DOP + ECR + ECH + OIA + EIE + OIE - OOC = 0 \quad (11)$$

- ***Income tax and tax payments***

The fourth and last variable in the operating part of the cash flow statement is tax payments and provision settlements (TXP). Related items from the change in balance sheet statements are: change in deferred tax (ΔDTX), change in income tax payable (ΔITP) and change in provision for risks and charges (ΔPRC), while the only item from the income statement is income taxes (ITX). Therefore the equation can be written as follows:

$$(\Delta DTX + \Delta ITP + \Delta PRC\ ) + (ITX) = TXP \quad (12)$$

By replacing the location of tax payments and provision settlements (TXP), the double entry constraint can be seen, as below:

$$(\Delta DTX + \Delta ITP + \Delta PRC\ ) + (ITX) - TXP = 0 \quad (13)$$

- ***Gain and loss on sale of assets and capital expenditure net of disposal***

The fifth variable. and the only one showing the investing activities in the cash flow statement. is capital expenditure net of disposal (CED). Change in property, plants and equipment (ΔPPE), change in long-term receivables (ΔLTR), change in other investments (ΔOIN) and change in investments in unconsolidated subsidiaries (ΔIUS) are the items from the change in balance sheets, while the only item from the income statement is extraordinary items and gain (loss) on sale of assets (EGL). Therefore the equation can be written as:

$$(\Delta PPE + \Delta LTR + \Delta OIN + \Delta IUS) + (EGL) = CED \quad (14)$$

As before, by changing the location of capital expenditure net of disposal (CED), the double entry constraint can be observed, as below:

$$(\Delta PPE + \Delta LTR + \Delta OIN + \Delta IUS) + (EGL) - CED = 0 \quad (15)$$

- ***Interest and debt issue (net of debt repayment)***

The sixth and the seventh variables of the cash flow statement are allocated to financing activities. The sixth variable is debt issue (net of debt repayment) (DIR). The applied variables from the change in balance sheet statements are: change in short-term debt (ΔSTD), change in long-term debt (ΔLTD) and change in cash and short-term investment (ΔCSI), while interest expense on debt (IED) and interest income (IIN) are the variables from the income statement. Incorporating all these variables we have:

$$(\Delta STD + \Delta LTD + \Delta CSI) + (IED + IIN) = DIR \quad (16)$$

By replacing the location of debt issue (net of debt repayment) (DIR), the double entry constraint can be shown as below:

$$(\Delta STD + \Delta LTD + \Delta CSI) + (IED + IIN) - DIR = 0 \quad (17)$$

- ***Net income and new capital (net of dividends)***

The last variable of the cash flow statement, and also the last variable of financing cash flows is new capital (net of dividends) (NCD). Items from the change in balance sheet statements are: change in shareholders' equity (ΔSEQ), change in preferred stock (ΔPST), change in non-equity reserves (ΔNER) and change in minority interest (ΔMIN). Net income (NI), minority interest in earnings (MIE) and pre-tax equity interest earnings (PEI) are the items from the income statement. Therefore we have:

$$(\Delta SEQ + \Delta PST + \Delta NER + \Delta MIN) + (NI + MIE + PEI\ ) = NCD \quad (18)$$

By replacing the location of new capital (net of dividends) (NCD), the double entry constraint for the final line-items can be seen as:

$$(\Delta SEQ + \Delta PST + \Delta NER + \Delta MIN) + (NI + MIE + PEI\ ) - NCD = 0 \quad (19)$$

3-2-2 Articulated cash flow statement variables: mean values

The primary aim of this section is to move one step further by providing a comprehensive example for an average company (mean of 21,698 financial statements in the prediction of future cash flows) to show how, given the definition of the variables, different accounting numbers from the various financial statements are articulated with relation to each other.

The comprehensive articulation table in prediction of future cash flow is divided into seven horizontal categories or, in other words, into seven line-items in order to assign the interrelated items to the same categories. The table shows the articulation of financial statements. The first line-items relate to the variables creating customer receipts (*REC*).

Table 3-10 Articulation of financial statement variables in the prediction of future cash flows

Δ Accounts receivable	**Δ*ARE***	**6,705,832.30**	**Sales**	***SAL***	**-440,982,275.27**	**Customer receipts**	***REC***	-434,276,442.97	**0.00**
Δ Payables net of inventory		**-217,350.40**	**Purchases**		**377,133,747.59**				
Δ Accounts payable	Δ*APA*	-3,691,604.05	Cost of goods sold	*CGS*	283,079,122.56	**Supplier payments**	***PAY***	376,916,397.19	**0.00**
Δ Total inventory	Δ*INV*	4,688,963.84	Selling and administrative expenses#	*SAE*	94,054,625.03				
Δ Accrued payroll	Δ*ACP*	-1,214,710.19							
Δ Other items		**4,260,432.26**	**Other operating and non-operating expenses and depreciation**		**26,126,544.10**	**Other operating cash flow**	***OOC***	30,386,976.36	**0.00**
Δ Other current liabilities	Δ*OCL*	-4,439,444.67	Other operating expense	*OOE*	5,052,213.55				
Δ Other current assets	Δ*OCA*	1,666,743.98	Depreciation, depletion and amortisation	*DDA*	16,203,502.65				
Δ Prepaid expenses	Δ*PRE*	250,598.04	Discontinued operations	*DOP*	62,296.40				
Δ Deferred income	Δ*DIN*	-778,768.71	Extraordinary credits	*ECR*	-532,418.19				
Δ Other long-term assets	Δ*OLA*	9,523,992.69	Extraordinary charges	*ECH*	7,289,417.84				
Δ Other long-term liabilities	Δ*OLL*	-1,962,689.07	Other income (expenses) – after tax	*OIA*	0.00				
			Equity interest earnings	*EIE*	-797,967.59				
			Other (income) expense – net	*OIE*	-1,150,500.56				
Δ Taxes and provisions		**-1,353,295.75**	**Income taxes**	***ITX***	**13,040,134.49**	**Tax payments and provision settlements**	***TXP***	11,686,838.74	**0.00**
Δ Deferred tax	Δ*DTX*	514,122.94							
Δ Income tax payable	Δ*ITP*	-8,043.11							
Δ Provision for risks and charges	Δ*PRC*	-1,859,375.58							
Δ Long-term assets		**9,408,952.59**	**Extraordinary items & gain (loss) on sale of assets**	***EGL***	**0.00**	**Capital expenditure (net of asset disposals)**	***CED***	9,408,952.59	**0.00**
Δ Property plant and equipment	Δ*PPE*	7,235,843.45							
Δ Long-term receivables	Δ*LTR*	737,382.53							
Δ Other investments	Δ*OIN*	812,611.89							
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	623,114.72							
Δ Financial items		**-3,775,837.11**	**Increase (decrease) net interest**		**2,874,409.77**	**Debt issues (net of debt repayment)**	***DIR***	-901,427.34	**0.00**
Δ Short-term debt	Δ*STD*	-2,841,324.27	Interest expense on debt	*IED*	4,987,181.04				
Δ Long-term debt	Δ*LTD*	-7,240,398.80	Interest income	*IIN*	-2,112,771.27				
Δ Cash and short-term investments	Δ*CSI*	6,305,885.96							
Δ Equities		**-15,028,733.89**	**Increase (decrease) current profits (losses)**		**21,807,439.33**	**New capital (net of dividends)**	***NCD***	6,778,705.44	**0.00**
Δ Non-equity reserves	Δ*NER*	0.00	Net income (before preferred dividends)	*NI*	21,479,476.40				
Δ Minority interest in net assets	Δ*MIN*	-354,067.99	Minority interest in earnings	*MIE*	327,962.93				
Δ Preferred stock	Δ*PST*	-4,928,679.83	Equity interest earnings	*PEI*	0.00				
Δ Shareholders' equity	Δ*SEQ*	-9,685,539.31							
Δ Dividend payable	Δ*DPA*	-60,446.76							
Total		**0**	**Total**		**0**	**Total**		**0**	

3-2-2-1 Customer receipts (REC)

Customer receipts (*REC*), as the first item of the cash flow statement, is calculated by taking the difference between the change in accounts receivable (ΔARE) in the balance sheet and sales (*SAL*) in the income statement.

The value of 6,705,832.30 for change in accounts receivable (ΔARE) indicates the increase in this variable or, in other words, the increase in the debit items from the last year to the current year. In addition, in the income statement, the value of sales (*SAL*) is shown to be -440,982,275.27, signifying that this is a credit variable. Therefore, making the reconciliation between these two variables leads to customer receipts (*REC*) having the value -434,276,442.97. As can be seen in Table 3-10, customer receipts (*REC*) is displayed with a negative value as a result of the overvaluation of the credit item (sales) compared to the debit item (increase in accounts receivable).

With regard to the definition of the variables, accounts receivable (*ARE*) represents the amounts due to the company resulting from the sale of goods and services on credit to customers after applicable reserves. These assets should reasonably be expected to be collected within a year, or within the normal operating cycle of a business. Additionally, sales (*SAL*) represents gross sales and other operating revenue less discounts, returns and allowances.

Ultimately, as can be seen, the summation of the line-item variables is equal to zero, indicating the double entry equilibrium.

In the following section, the method of calculating supplier payments (*PAY*), as the second line-item, will be demonstrated.

3-2-2-2 Supplier payments (PAY)

In the second line-item, it is demonstrated how to create supplier payments (*PAY*) using the variables of the other financial statements' variables. Indeed, supplier payments indicates the difference between *purchases* in the income statement and change in *payables net of inventory* in the balance sheet.

Purchases contains two components: cost of goods sold (*CGS*) and selling and administrative expenses (*SAE*). In addition, change in payables net of inventory has three components, which are: change in accounts payable (ΔAPA), change in total inventory (ΔINV) and change in accrued payroll (ΔACP). The reason for putting these three items next to each other in the balance sheet categorisation is because they are interrelated with purchases in the income statement.

For instance, increasing *CGS* is associated with decreasing inventories. For the same reason, cost of goods sold (*CGS*) and selling and administrative expenses (*SAE*) are in the same group, as they are interrelated with change in payables net of inventory in the balance sheet.

The issue that has to be addressed is that, as a result of placing similar items in the same group, it is possible to place liabilities beside assets as long as they are interrelated. For instance, in Table 3-10, both change in accounts payable (ΔAPA)

and change in accrued payroll (ΔACP) are from liabilities; however, change in total inventory (ΔINV) is from assets.

For this reason, in the balance sheet section of the articulation in Table 3-10, both positive signs and negative signs are presented next to each other. Increases in liabilities are shown with a negative sign, while this sign appears as positive for increases in assets. As can be seen, the value for change in accounts payable is -3,691,604.05, presented with a negative sign, implying an increase in liabilities from the previous year to the current year; however, this value is presented with a positive sign as 4,688,963.84 for change in total inventory (ΔINV), which again indicates an increase, but this time for assets. The same situation is also observed for change in accrued payroll (ΔACP), with the value of -1,214,710.19 representing an increase in liabilities. In the income statement both CGS and SAE are presented with a positive sign, indicating that these are expenses for the company, and given conventional accounting rules they are kept on the debit side of accounting journals.

In the end, the overall value of change in payables net of inventory is presented as -217,350.40, implying an overvaluation of credit items to debit items, or, in other words, an overvaluation of the increase in liabilities compared to the increase in assets. In the income statement, purchases, with a value of 377,133,747.59, indicates the existence of the debit items.

Therefore, supplier payment (PAY), with the value of 376,916,397.19, is calculated by making the reconciliation between purchases and the change in

payables net of inventory. The positive sign of supplier payment (*PAY*) represents the overvaluation of debit items to credit items. It is important to recall that in the previous section, customer receipts (*REC*) was presented with a negative sign; however, in this part, supplier payments (*PAY*) is presented with a positive sign. In this thesis all receipts are shown with a negative sign and all payments are shown with a positive sign.

In the WorldScope database, accounts payable (*APA*) represents the claims of trade creditors for unpaid goods or services that are due within the normal operating cycle of the company. Meanwhile, accrued payroll (*ACP*) represents payroll expenses that are incurred within the normal operating cycle of the company but are not paid, so it is expected that they will be paid in the following year. Total inventory (*INV*) is defined as tangible items or merchandise net of advances and obsolescence, acquired for either 1) direct resale or 2) inclusion in the production of finished goods manufactured for sale in the normal course of operations. In manufacturing companies, depending upon the stage of completion in the manufacturing process, this item is defined as follows:

1) Finished goods, consisting of products ready for sale;

2) Work in process, consisting of products in various stages of production;

3) Raw materials and supplies, consisting of items that will enter directly or indirectly into the production of finished goods.

In non-manufacturing companies, finished goods bought for resale form the major portion of the inventories.

Ultimately, as can be seen, the summation of all the line-item variables is equal to zero, owing to the double entry equilibrium.

In both the previous section and the current section, customer receipts (*REC*) and supplier payments (*PAY*) have been explained. Two other components of the cash flow statement still remain in order to shape the operating cash flow (*OCF*) completely. They are: other operating cash flow (*OOC*) and tax payments and provision settlements (*TXP*). The third item, which will be explained in the next section, is other operating cash flow (*OOC*).

3-2-2-3 Other operating cash flow (OOC)

The aim of this part is to gather together *other items,* whether in the balance sheet statement or in the income statement. Consequently, other operating cash flow (*OOC*) in the cash flow statement can be calculated by taking the difference between *change in other items* in the balance sheet and *other operating and non-operating expenses and depreciation* in the income statement.

Change in other items in the balance sheet is comprised of six components: change in other current liabilities (ΔOCL), change in other current assets (ΔOCA), change in prepaid expenses (ΔPRE), change in deferred income (ΔDIN), change in other long-term assets (ΔOLA) and change in other long-term liabilities (ΔOLL). As can be seen, most of these items start with the word 'other', as they are related to the other activities of the company. The two variables that do not start with this word (change in deferred income (ΔDIN) and change in prepaid expenses (ΔPRE))

cannot be allocated to the other line-items, which is why it has been decided to include them in the Δ other items section in the balance sheet.

As with the second line-items, a combination of assets and liabilities is presented here. Therefore, increases in assets are displayed with a positive sign, while increases in liabilities are displayed with a negative sign. For instance, the value for change in other current liabilities (Δ*OCL*) in Table 3-10 is shown with a negative value as -4,439,444.67, demonstrating an increase in liabilities in comparison to the previous year; however, change in other current liabilities (Δ*OCA*) is presented with a positive sign, indicating an increase, but this time in assets.

This also applies to *other operating and non-operating expenses and depreciation* in the income statement, as this comprises eight different components: other operating expenses (*OOE*), depreciation, depletion and amortisation (*DDA*), discontinued operations (*DOP*), extraordinary credit (*ECR*), extraordinary charges (*ECH*), other income (expenses) after tax (*OIA*), equity interest earnings (*EIE*) and finally other (income) expenses net (*OIE*). As can be seen, some of these variables, as with the balance sheet changes, are presented with the word 'other' at the start, which is why they are included in this section of the income statement.

An operating expense is an ongoing cost for running a product, business or system. Operating expenses include depreciation of plants and machinery that are used in the production process. Although these operating expenses in the income statement are associated with long-term assets such as property, plants and

equipment (*PPE*) (which is known as the investing variable) in the balance sheet, in the articulation table, putting depreciation, depletion and amortisation (*DDA*) in the same line-item, next to property, plants and equipment (*PPE*) leads to an overvaluation of investing cash outflows and consequently an undervaluation of the operating cash inflows.

The rest of the variables cannot be allocated to the other line-items, which is why they are all gathered together in this section. Regarding the issue of signs, some of the mentioned variables in the income statement are presented with a positive sign and, as noted earlier, they are expenses and debit; however, items displayed with a negative sign are earnings and credit.

Finally, the sum of the *change in other items* in the balance sheet is calculated as 4,260,432.26 with a positive sign, indicating an overvaluation of the debit items compared to the credit items, or, in other words, that the increase in assets is greater than the increase in liabilities. Moreover, the value of *other operating and non-operating expenses and depreciation* in the income statement is 26,126,544.10, indicating an overvaluation of the debit items compared to the credit items, or, in other words, that the expenses are greater than the earnings. By taking the difference between the change in *other items* in the balance sheet and *other operating and non-operating expenses and depreciation* in the income statement, other operating cash flows (*OOC*) in the cash flow statement can be calculated with the value of 30,386,976.36. The positive sign implies an overvaluation of the debit items compared to the credit items, and therefore it is a payment.

In the WorldScope database, prepaid expenses (*PRE*) is defined as payments that are made in advance for services that are to be received within the normal operating cycle of the company, and includes different items, such as: prepaid insurance expenses, prepaid property taxes, prepaid rent, prepaid income taxes, prepaid supplies, prepaid employee benefits and advance corporation tax recoverable within one year. Other current assets (*OCA*) in the balance sheet represents all other current assets besides cash and equivalents, receivables, inventories and prepaid expenses, and includes all deferred expenses included in current assets, property and assets held for disposal, deposits and advances to others, program rights, accrued ad valorem taxes, felling rights (if shown as a current asset) and deferred or future income tax benefits (if shown as a current asset). Other current liabilities (*OCL*) represents all other current liabilities besides accounts payable, short-term debt, accrued payroll, income taxes payable and dividends payable, and includes product guarantees and warranties, contingency reserves, taxes other than income taxes, payroll taxes, ad valorem taxes, restructuring expenses, pensions and dividends to minority shareholders of subsidiaries. Deferred income (*DIN*) represents income received but not yet reported as income. As the income is earned, the deferred income is debited and the revenue is credited. Deferred income may also be referred to as deferred credit or deferred revenue. It includes fees received but not yet earned (e.g. advertising fees), revenue normally included in income but deferred until earned and matched with expenses (e.g. long-term magazine subscriptions) and prepaid income. Other long-term assets (*OLA*) represents all other assets of the company besides current assets, long-term receivables, investment in unconsolidated subsidiaries, other

investments and net property, plants and equipment. Other long-term liabilities (*OLL*) represents all other long-term liabilities of the company besides current liabilities, long-term debt, provision for risk and charges and deferred taxes.

Discontinued operations (*DOP*) represents the after-tax gain or loss resulting from a company discontinuing a segment of its business. Extraordinary credit (*ECR*) represents any pre-tax extraordinary credit which is infrequent or unusual and which is included in the net income of a company. Extraordinary charges (*ECH*) represents any pre-tax extraordinary charge which is infrequent or unusual and which is included in the net income of a company. Other income (expenses) after tax (*OIA*) represents miscellaneous items reported net of tax and which are included in the company's net income. Equity interest earnings (*EIE*) represents that portion of the earnings or losses of a subsidiary whose financial accounts are not consolidated with the controlling company's assets. Other income/expenses net (*OIE*) represents other pre-tax income and expenses besides operating income, interest income, interest expenses on debt, interest capitalised, pre-tax extraordinary charges, pre-tax extraordinary credit and increases/decreases in reserves. Depreciation, depletion and amortisation (*DDA*) represents the process of allocating the cost of a depreciable asset to the accounting periods covered during its expected useful life for a business. It is a non-cash charge for use and obsolescence of an asset. Finally, other operating expenses (OOE) represents all other operating expenses besides interest expenses, salaries and benefits expenses, equipment expenses and provisions for loan losses.

Ultimately, as can be seen, the summation of all the line-item variables is equal to zero, indicating the double entry equilibrium.

The last operating cash flow variable is tax payments and provision settlements (*TXP*), which will be investigated in the next section.

3-2-2-4 Tax payments and provision settlements (TXP)

In this part, the variables employed for generating tax payments and provision settlements (*TXP*) will be investigated. *Change in taxes and provisions* from the balance sheet statement and *income taxes* (*ITX*) from the income statement are the creative components of the *TXP*.

Change in taxes and provision settlements comprises three components: change in deferred tax (Δ*DTX*), change in income tax payable (Δ*ITP*) and change in provision for risks and charges (Δ*PRC*). All these three variables are from the liabilities section, and consequently it is expected that any *increase* in these variables will be displayed with a negative sign. As can be seen, the value for change in deferred tax (Δ*DTX)* is shown with a positive sign, indicating the *decrease* in this variable in comparison to the previous year; however, the other two variables in that category, change in income tax payable (Δ*ITP*) and change in provision for risks and charges (Δ*PRC*), are presented with a negative sign, indicating the increase in these liabilities in comparison to the previous year. The summation of these three variables is -1,353,295.75, showing an overvaluation of the credit items compared to the debit items, or, in other words, addressing the

issue that the increase in liabilities is greater than the decrease in liabilities. In addition, in the income statement, income taxes (*ITX*) is shown with a positive sign, which signifies that this is an expense for the company. By making the reconciliation between the change in income taxes in the balance sheet and income taxes (*ITX*) in the income statement, tax payments and provision settlements (*TXP*) is calculated with the value of 11,686,838.74. The sign of tax payments and provision settlements (*TXP*) is positive, which signifies that this is a payment for the company.

In fact, allocating the variables to this line-item is more precise than allocating the variables to other line-items, as the words 'tax' or 'provision' almost always exist in the variables' names, whether in the balance sheet or the income statement.

In the WorldScope database, deferred tax (*DTX*) represents the accumulation of taxes that are deferred as a result of timing differences between reporting sales and expenses for tax and financial reporting purposes. It includes investment tax credits and excludes ad valorem taxes, taxes other than income taxes, tax loss carry forward reported as an asset and general and service taxes. It also represents an accrued tax liability that is due within the normal operating cycle of the company, and excludes taxes other than income, ad valorem taxes, value added taxes, general and service taxes, excise taxes and windfall profit taxes.

PRC includes provisions for pension funds, repairs and maintenance and litigation claims. In addition, change in income tax payable (Δ*ITP*) represents an accrued tax liability that is due within the normal operating cycle of the company.

One of the main concerns of this thesis is to put similar variables in the same sections in the articulation table (Table 3-10); however, given the statement of the CFA (2005), some difficulties arise in finding purely similar variables. For instance, the CFA notes that, in the view of financial analysts, one of the most controversial issues concerns the inconsistent pattern of aggregation and netting of items in financial statements, along with their obscured, even opaque, articulation. Analysts consider this a barrier to efficient analysis, leading to what is described as nothing more than a 'best guess'. All the variables coming from the Δ *taxes and provisions* section of the balance sheet, which are change in provision for risks and charges (Δ*PRC*), change in deferred tax (Δ*DTX*) and change in income tax payable (Δ*ITP*), are categorised as the liabilities, while both change in provision for risks and charges (Δ*PRC*) and change in income tax payable (Δ*ITP*) are long-term liabilities. The next variable, income taxes (*ITX*) in the income statement, also represents all income taxes levied on the income of a company by federal, state and foreign governments, and includes: federal income taxes, state income taxes, foreign income taxes, charges in lieu of income taxes, charges equivalent to investment tax credit, income taxes on dividends or earnings of unconsolidated subsidiaries or minority interests, if reported before taxes, and deferred taxation charges.

Ultimately, as can be seen, the summation of all the line-item variables is equal to zero, indicating the double entry equilibrium.

Now we have all the components of operating cash flows (*OCF*), as customer receipts (*REC*), supplier payments (*PAY*), other operating cash flow (*OOC*) and tax payments and provision settlements (*TXP*) have all been calculated. In the following section, the investing cash flow (*ICF*) will be investigated.

3-2-2-5 Capital expenditure (net of asset disposals) (CED)

As the operating part of the cash flow statement was defined in the previous sections, it is now the turn of investing cash flow (*ICF*) to be generated using the balance sheet and the income statement variables. In fact, *ICF* is made by carrying out a reconciliation between *change in long-term assets* in the balance sheet and *extraordinary items and gain (loss) on sale of assets (EGL)* in the income statement..

Change in long-term assets in the balance sheet comprises four different components: change in property, plants and equipment (ΔPPE), change in long-term receivables (ΔLTR), change in other investments (ΔOIN) and change in investments in unconsolidated subsidiaries (ΔIUS). All these four variables are known as assets, and any positive sign for these variables indicates an increase from the previous year to the current year. Moreover, as can be seen, the summation of all these four variables is presented with a positive sign, implying that all of these four items increased in comparison to their previous years.

Based on the WorldScope database, property, plants and equipment (*PPE*) includes land, buildings, machinery, equipment, construction work in progress,

minerals, oil, automobiles and trucks, timberland and timber rights, leasehold improvements, rented equipment, furniture and fixtures, property, plants and equipment leased under capitalised lease obligations, book plates, noncurrent film costs and inventory, broadcasting rights and licenses, franchise rights and licenses, long-term power purchase contracts, publishing rights and licenses, funds held for construction, long-term power purchase contracts and software products.

Long-term receivables (*LTR*) also represents amounts due from customers that will not be collected within the normal operating cycle of the company. It includes loans to other companies, investments in sale-type leases and finance receivables for industrial companies whose financial subsidiaries are consolidated. Regarding other investment (*OIN*), different definitions are given for the different situations presented. In industrial sectors, other investment represents any other long-term investment except for investments in unconsolidated subsidiaries. In banks, it represents time deposits, interest-bearing deposits and other short-term investments. For insurance companies, it represents all other investments of the insurance company besides fixed income and equity securities, real estate investments and mortgage and policy loans. Finally, for other financial companies it represents all investments except real estate assets and loans.

The final item from the balance sheet is investment in unconsolidated subsidiaries (*IUS*), which represents long-term investments and advances in unconsolidated subsidiaries and affiliates with which the company has a business relationship or over which it exercises control. It includes joint ventures,

investment in non-consolidated subsidiaries or associated companies, investment in joint ventures and advances to related companies.

In the income statement, the value of extraordinary items and gain (loss) on sale of assets (*EGL*) is shown as zero. In the following parts of this chapter it is explained that all the observations of non-zero extraordinary items and gain (loss) on sale of assets (*EGL*) have been dropped from the main usable sample, as they bring bias into the size of the coefficients as influential observations. This is because the number of these non-zero observations is very small in comparison to the non-zero variables. Extraordinary items and gain (loss) on sale of assets (*EGL*) represents gains and losses resulting from nonrecurring or unusual events.

Therefore, the value of capital expenditure (net of asset disposals) (*CED*) is equal to the value of the change in long-term assets and is displayed with a positive sign, indicating that this is a payment for the company.

As in the previous sections, the summation of the line-item variables is equal to zero, showing the double entry equilibrium.

Capital expenditure (net of asset disposals) (*CED*) is the only item in investing cash flow (*ICF*), and has already been explained in this section. In the following two sections the components of financing cash flows (*FCF*) will be investigated.

3-2-2-6 Debt issues (net of debt repayments) (DIR)

In the previous sections, both operating cash flow (*OCF*) and investing cash flow (*ICF*) have been reviewed. In both this section and the next section, financing cash flow (*FCF*) will be illustrated.

Debt issues (net of debt repayments) (*DIR*) is the first component of financing cash flow (*FCF*), and is calculated by taking the difference between both *change in financial items* in the balance sheet and *Increase (decrease) net interest* in the income statement.

Change in financial items comprises three components: change in short-term debt (ΔSTD), change in long-term debt (ΔLTD) and change in cash and short-term investments (ΔCSI). Both short-term debt (*STD*) and long-term debt (*LTD*) are liabilities, and consequently it is expected that any increase in these variables will be displayed with a negative sign, or *vice-versa.* The third item from the balance sheet, cash and short-term investments (*CSI*), is categorised as a financial asset.

In order to make the results of this thesis more comparable, renaming the changes of the balance sheet variables is attempted. The balance sheet changes are therefore divided into three components: accruals, changes in cash and short-term investments, and changes in equities. The use of the term *accrual* is driven by the work of Richardson et al. (2005), as they refer to all non-cash and non-equity numbers as accruals. Thus, in *change in financial items*, both accruals (change in

short-term debt (ΔSTD) and change in long-term debt (ΔLTD)) and changes in cash (ΔCSI) are included.

Cash and short term investments (CSI) is an asset, in contrast to short-term debt (STD) and long-term debt (LTD), which are liabilities; therefore, the increase in this variable is displayed with a positive sign. As can be seen in Table 3-10, both change in short-term debt (ΔSTD) and change in long-term debt (ΔLTD) are shown with a negative sign, indicating an increase in these variables in comparison to their previous years; however, change in cash and short term investments (ΔCSI) is displayed with a positive sign, again implying an increase in this variable. The value of change in financial items in the balance sheet is -3,775,837.11 with a negative sign, indicating an overvaluation of the credit items compared to the debit items, or, in other words, it indicates that the increase in liabilities is greater than the increase in assets.

Based on the WorldScope database, short-term debt (STD) represents that portion of debt payable within one year, including the current portion of long-term debt and sinking fund requirements of preferred stocks or debentures, and includes notes payable, arising from short-term borrowings, current maturities of participation and entertainment obligations, contracts payable for broadcast rights, current portions of advances and production payments, current portions of long-term debt that must be paid back over the next 12 months and are included in long-term debt, bank overdrafts and advances from subsidiaries/associated companies. Long-term debt (LTD) represents all interest-bearing financial obligations, excluding amounts due within one year. It includes net of premiums or

discounts, encompassing mortgages, bonds, debentures and convertible debt, among other items. Cash and short-term investments (*CSI*) represents the sum of cash and short-term investments, including cash on hand, undeposited cheques, cash in banks and cheques in transit, among other items.

In the income statement the story is almost the same, as the mixture of debit and credit items is presented as components of the *increase (decrease) net interest*. Interest expenses on debt (*IED*) as the debit item and interest income (*IIN*) as the credit item are the components of *increase (decrease) net interest* in the income statement. Interest expenses on debt (*IED*) is an expense for the company, and thus it is displayed with a positive sign, while interest income (*IIN*) is an earning for the company and is shown with a negative sign.

The overall value of the *increase (decrease) net interest* in the income statement is 2,874,409.77 with a positive sign, indicating an overvaluation of the debit items compared to the credit items, or, in other words, an overvaluation of the interest expenses compared to the interest earnings.

Interest expenses on debt (*IED*) represents the service charge for the use of capital before reductions for interest capitalised. If interest expenses are reported net of interest income, and interest income cannot be found, the net figure is shown, and it includes interest expenses on short-term debt, interest expenses on long-term debt and capitalised lease obligations, amortisation expenses associated with the issuance of debt and similar charges.

The second item is interest income (*IIN*), representing income generated from interest-bearing investments not related to the operating activities of the company, and includes interest on savings and interest on investments.

Finally, the value for *DIR* is presented as -901,427.34 with a negative sign, representing that this is a receipt for the company, like customer receipts (*REC*).

As can be seen, the summation of the variables for this line-item is equal to zero, indicating the double entry equilibrium.

The final line-item in the articulation table, and the final part of financial cash flow (*FCF*) is new capital (net of dividends) (*NCD*), which will be investigated in the next section.

3-2-2-7 New capital (net of dividends) (NCD)

In this section of the study, new capital (net of dividends) (*NCD*), as the second and final part of financing cash flow (*FCF*), and also as the final line-item of the articulation table, shown as Table 3-10, is investigated. New capital (net of dividends) (*NCD*) is calculated by making the reconciliation between *change in equities* in the balance sheet and *increase (decrease) current profits (losses)* in the income statement.

Change in equities comprises five different items: change in non-equity reserves (ΔNER), change in minority interest in net assets (ΔMIN), change in preferred stock (ΔPST), change in shareholders' equity (ΔSEQ) and change in

dividends payable (Δ*DPA*). As can be seen, all the values are from the equity section of the balance sheet, and therefore any increase in these variables is expected to be shown with a negative sign.

As can also be seen, all of the variables are presented as negative, indicating that all the equity components increased from the previous year to the current year. Moreover, for the same reason that was mentioned for extraordinary items and gain (loss) on sale of assets (*EGL*) in the previous section, change in non-equity reserves (Δ*NER*) is presented with a value of zero.

The summation of all of these variables as Δ equities is -15,028,733.89 and is presented with a negative sign, representing an increase in the credit items or, in other words, an increase in equities.

Change in non-equity reserves (*NER*) represents a reserve set aside from shareholders' equity with no designated payee. This item is only applicable to non-US corporations and includes specific reserves (including untaxed reserves) and revaluation reserves, as well as current cost reserves, unless clearly placed in shareholders' equity and electricity supply funds. Minority interest in net assets (*MIN*) represents the portion of the net worth (at par or stated value) of a subsidiary pertaining to shares not owned by the controlling company and its consolidated subsidiaries. If minority interest is shown on the asset side, it is updated as a negative and the total assets are adjusted to exclude it. Preferred stock (*PST*) represents a claim prior to that of the common shareholders on the earnings of a company and on its assets in the event of liquidation. For US

corporations, their value is shown as the total involuntary liquidation value of the number of preferred shares outstanding at year end. If preferred stock is redeemable at any time by the shareholder, it is shown at redemption value, or if the company carries it at a higher value than the involuntary liquidation value, the stated value. Shareholders' equity (*SEQ*) represents common shareholders' investment in a company. Dividends payable (*DPA*) represents a dividend that has been declared but not yet paid, and excludes dividends to minority shareholders of subsidiaries.

In the income statement, *increase (decrease) current profits (losses)* comprises three components: net income (before preferred dividend) (*NI*), minority interest in earnings (*MIE*) and pre-tax equity interest earnings (*PEI*).

Net income (before preferred dividend) (*NI*) represents income after all operating and non-operating income and expenses, reserves, income taxes, minority interest and extraordinary items. Minority interest in earnings (*MIE*) represents the portion of earnings/losses of a subsidiary pertaining to common stock not owned by the controlling company or by other members of the consolidated group. Finally, pre-tax equity interest earnings (*PEI*) represents that portion of the earnings or losses of a subsidiary whose financial accounts are not consolidated with the controlling company's assets.

Pre-tax equity interest earnings (*PEI*), for the same reasons that were mentioned before for extraordinary items and gain (loss) on sale of assets (*EGL*), is presented with a value of zero. Minority interest in earnings (*MIE*) is positive,

indicating that this is an expense for the company and is debit; however, net income (before preferred dividend) (*NI*) is an exception for the company and, as mentioned in the previous chapter, any profit has to be shown with a positive sign and any loss has to be shown with a negative sign. This is because in the income statement equation that was presented in the previous chapter, net income (before preferred dividend) (*NI*) is on the right hand side of the equation as a result of the income statement variables, and in order to keep the double entry equilibrium by presenting a value of zero at the end of the equation, this variable has to move to the left hand side of the equation. Therefore its sign will be different from the rest of the variables of the income statement. The positive sign displayed here for net income (before preferred dividend) (*NI*) represents that on average all US corporations have an income of 21,479,476.40.

Finally, the value of new capital (net of dividends) (*NCD*) is presented as 6,778,705.44, with a positive sign indicating that this is a payment for the company. The summation of the variables also shows the double entry equilibrium, as it is presented with a value of zero.

The sum of debt issues (net of debt repayments) (*DIR*) and new capital (net of dividends) (*NCD*) represents the financing cash flows (*FCF*). The final zero that is presented at the end of the cash flow statement represents *the double entry constraint* that is used as a unique condition in the prediction of both future cash flows and future earnings. In fact, when all the line-item variables are equal to zero and all the financial statement variables are summed to zero, this number is also equal to zero.

In the next section, the articulation table is again presented, but this time for the prediction of future earnings, with just one difference in the other operating cash flow (*OOC*) section, which is divided into two components.

3-2-2-8 Other operating cash flow (OOC) as a special issue

In Table 3-10, in order to calculate the other operating cash flow (*OOC*), reconciliation is made between the Δ *other items* in the balance sheet and *other operating expenses and depreciation* in the income statement. However, using these line-items will cause a problem in the prediction of future earnings, as the employed variables from the income statement (*other operating expenses and depreciation)* comprise both operating and non-operating variables. For instance, there are some variables, such as discontinued operations (*DOC*), that are known as non-operating variables in the income statement, but, consistent with International Accounting Standards Number 7 (IAS7), they have to be in the operating part of the cash flow statement. As another example, in the work of Alexander et al. (2011, p 559), discontinued operation (*DOP)* is shown under operating activities in the cash flow statement.

There are some other items that are in the same position as discontinued operation (*DCO),* which were categorised under the non-operating part of the income statement but which, in fact, consistent with the work of Alexander, Britton et al. (2011, p 547) and IAS 7, have to be allocated to the operating part of the cash flow statement, as it is not possible to allocate them to either the investing part or the financing part of the cash flow statement.

As the aim of the second part of this thesis is to predict future operating and non-operating income, the main concern in this part is to decompose the income statement variables employed in creating the other operating cash flow (*OOC*) into two components: *operating* and *non-operating*. To do so in the new articulation table that is presented for the prediction of future earnings (Table 3-11), *OOC* is divided into two components:

- Other operating cash flow (*OOC*)
- Exceptional cash flow (*ECF*)

As can be seen, the name of the first decomposed variable is the same as before, as it solely represents the operating variables in the income statement: other operating cash flow (*OOE*) and depreciation, depletion and amortisation (*DDA*). In fact, given conventional accounting standards, in order to reach the value of the operating income, five variables are needed: sales (*SAL*), cost of goods sold (*CGS*), selling and administrative expenses (*SAE*) (all of which were introduced in the first and second line-items), other operating cash flow (*OOE*) and depreciation, depletion and amortisation (*DDA*). The seventh column in Table 3-11 shows variable categorisation based on being *operating* or *non-operating*.

For the second component, exceptional cash flow (*ECF*) attempts to bring together all *other income and expenses* that are non-operating, such as discontinued operations (*DOP*), extraordinary credit *(ECR*), extraordinary charges (*ECH*), other income (expenses) after tax (*OIA*), equity interest earnings (*EIE*) and other income (expenses) net (*OIE*).

In addition, the balance sheet variables are decomposed into two components, as shown in Table 3-11. Change in other long-term assets (ΔOLA) and

change in other long-term liabilities (ΔOLL) are used in calculating the exceptional cash flow (ECF); however, the rest of the variables are used in calculating the other operating cash flow (OOC).

In Table 3-10, the articulation table for the prediction of future cash flows, the value of other operating cash flow (OOC) is 30,386,976.36, with the positive sign indicating that this is a payment for the company. Therefore, it is expected that the summation of its components (presented in Table 3-11) will be equal to this value.

In Table 3-11, the articulation table for the prediction of future earnings, the value of OOC and ECF are 17,954,844.84 and 12,432,131.52 respectively, and the summation of these two values is the same as the value of the other operating cash flow (OOC) in Table 3-10.

One of the key concerns in this study is the 'non-articulation' problem that all financial statements are currently facing. Supporting this issue, the CFA (2005) notes that the current conceptual framework of accounting needs to be refined. In fact, the CFA (2005) has made the statement that, in the view of its financial analysts, one of the most controversial issues concerns the unclear and ambiguous articulation among financial statements, which analysts observe as a limitation for efficient empirical accounting analysis, and point out that drawing articulation among the different financial statements is mostly based on 'guesses' instead of a clear and transparent structural framework. As mentioned in Chapter 2, a set of 12 suggestions have been presented by the CFA (2005) in order to improve the

quality of financial statements, and particularly the quality of articulation among the financial statements.

The eighth suggestion is about resolving the weak articulation among the financial statements, and attempts to present a solution for this problem. In this suggestion, a 'good articulation' is first defined as a condition whereby through just looking at the opening and closing balance sheets and the income statement, the cash flow statement will be able to be created. As an example, it is stated how individual items flow from one to another of the statements, such as credit sales, receivables in the balance sheet, cash collections in the cash flow statement, and revenues in the income statement. The CFA notes that currently, this articulation is opaque to investors, as the various statements do not follow a consistent structure or the same pattern of aggregation. As a solution, only when all of the changes to individual items are made clear and the measurement characteristics of items are fully disclosed will investors be able to understand a company's process of articulation. The same statement is present in the CFA's tenth suggestion, as they note that any factor that affects any of the financial statements should be reported on with sufficient disaggregation to allow an understanding of the articulation of the financial statement. The aim of this thesis is to construct an articulation table that draws links between all financial statements (opening and closing balance sheets, income statement and cash flow statement); however, to do so, as mentioned, we have faced difficulties, as the associations between the accounts are almost unclear and imprecise, leading to a simply-drawn articulation that may also be different from another investor's perspective. For instance, consistent with the work of Fischer et al. (2008) in certain aggregations of working capital accounts,

such as changes in other current liabilities, it is difficult to tell whether some of the amounts might be related to investing or financing activities rather than operations.

Even though this thesis has faced limitations of 'non-articulation', interestingly, the regression results reported in this thesis will show the superiority of the model that considers this simply-drawn articulation of financial statements by presenting rational signs for estimators (explained in Table 5-9), as well as a greater statistical efficiency, which will be explained in full in Chapter 6.

In the subsequent sections of this chapter, the sample selection process and industry distribution for the usable sample will be illustrated. At the end of the chapter the descriptive statistics table will also be presented.

Table 3-11 Articulation of financial statement variables in the prediction of future earnings

Δ Accounts receivable	**Δ*ARE***	**6,705,832.30**	**Sales**	***SAL***	**-440,982,275.27**	**Operating Income**	**Customer receipts**	***REC***	-434,276,442.97	**0.00**
Δ Payables net of inventory		**-217,350.40**	**Purchases**	***PUR***	**377,133,747.59**					
Δ Accounts payable	Δ*APA*	-3,691,604.05	Cost of goods sold	*CGS*	283,079,122.56	**Operating Income**	**Supplier payments**	***PAY***	376,916,397.19	**0.00**
Δ Total inventory	Δ*INV*	4,688,963.84	Selling and administrative expenses#	*SAE*	94,054,625.03					
Δ Accrued payroll	Δ*ACP*	-1,214,710.19								
Δ Other items		**-3,300,871.36**	**Other operating expenses and depreciation**	***OOD***	**21,255,716.20**					
Δ Other current liabilities	Δ*OCL*	-4,439,444.67	Other operating expenses#	*OOE*	5,052,213.55	**Operating Income**	**Other operating cash flow**	***OOC***	17,954,844.84	**0.00**
Δ Other current assets	Δ*OCA*	1,666,743.98	Depreciation, depletion and amortisation	*DDA*	16,203,502.65					
Δ Prepaid expenses	Δ*PRE*	250,598.04								
Δ Deferred income	Δ*DIN*	-778,768.71								
Δ Other non-current items		**7,561,303.62**	**other non-operating income (expenses)**	***ONE***	**4,870,827.90**					
Δ Other long-term assets	Δ*OLA*	9,523,992.69	Discontinued operations	*DOP*	62,296.40					
Δ Other long-term liabilities	Δ*OLL*	-1,962,689.07	Extraordinary credits	*ECR*	-532,418.19	**Non-operating income**	**Exceptional cash flow**	***ECF***	12,432,131.52	**0.00**
			Extraordinary charges	*ECH*	7,289,417.84					
			Other income (expense) – after tax	*OIA*	0.00					
			Equity interest earnings	*EIE*	-797,967.59					
			Other (income) expense – net	*OIE*	-1,150,500.56					
Δ Taxes and provisions	**Δ*TAP***	**-1,353,295.75**	**Income taxes**	***ITX***	**13,040,134.49**					
Δ Deferred tax	Δ*DTX*	514,122.94				**Non-operating income**	**Tax payments and provision settlements**	***TXP***	11,686,838.74	**0.00**
Δ Income tax payable	Δ*ITP*	-8,043.11								
Δ Provision for risks and charges	Δ*PRC*	-1,859,375.58								
Δ Long-term assets	**Δ*LTA***	**9,408,952.59**	**Extraordinary items & gain (loss) on sale of assets**	***EGL***	**0.00**					
Δ Property plant and equipment	Δ*PPE*	7,235,843.45				**Non-operating income**	**Capital expenditure (net of asset disposals)**	***CED***	9,408,952.59	**0.00**
Δ Long-term receivables	Δ*LTR*	737,382.53								
Δ Other investments	Δ*OIN*	812,611.89								
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	623,114.72								
Δ Financial items	**Δ*FIN***	**-3,775,837.11**	**Increase (decrease) net interest**	***ICH***	**2,874,409.77**					
Δ Short-term debt	Δ*STD*	-2,841,324.27	Interest expense on debt	*IED*	4,987,181.04	**Non-operating income**	**Debt issues (net of debt repayments)**	***DIR***	-901,427.34	**0.00**
Δ Long-term debt	Δ*LTD*	-7,240,398.80	Interest income	*IIN*	-2,112,771.27					
Δ Cash and short-term investments	Δ*CSI*	6,305,885.96								
Δ Equities		**-15,028,733.89**	**Increase (decrease) current profits (losses)**	***IDP***	**21,807,439.33**					
Δ Non-equity reserves	Δ*NER*	0.00	Net income (before preferred dividends)	*NI*	21,479,476.40	**Non-operating income**	**New capital (net of dividends)**	***NCD***	6,778,705.44	**0.00**
Δ Minority interest in net assets	Δ*MIN*	-354,067.99	Minority interest in earnings	*MIE*	327,962.93					
Δ Preferred stock	Δ*PST*	-4,928,679.83	Equity interest earnings	*PEI*	0.00					
Δ Shareholders' equity	Δ*SEQ*	-9,685,539.31								
Δ Dividend payable	Δ*DPA*	-60,446.76								
Total		**0**	**Total**		**0**		**Total**		**0**	

3-2-3 Operating, investing and financing equations

Up to this point, *seven* different variables have been introduced as the components of the cash flow statement. It is now necessary to move one step

further to allocate these seven variables to the three major identities representing a firm's activities: operating, investing and financing. Consequently, three linear regressions for each of these three identities are given, in order to show the predictive ability of the current-year components of each of these identities that come from both the change in the balance sheets and the income statement, with regard to future cash flows one year ahead. Moreover, eight different industries, based on the *ICB* codes, are used as dummy variables for each of these regressions. These will be explained in a later section.

Operating cash flow

The first equation is related to operating cash flows. In financial accounting, operating cash flow (*OCF*), i.e. cash flow provided by operations, or cash flow from operating activities, refers to the amount of cash a company generates from the revenues it brings in, excluding costs associated with long-term investments in capital items or investments in securities. As mentioned earlier, operating cash flows (*OCF*) are obtained from four different components: customer receipts (*REC*), payments to suppliers (*PAY*), other operating cash flow (*OOC*) and tax payments and provision settlements (*TXP*). Therefore, operating cash flows can be displayed as the sum of these four variables, as below:

$$OCF \equiv REC + PAY + OOC + TXP \tag{20}$$

By decomposing the above equation, we have:

$$OCF \equiv (\Delta ARE + SAL) + \big((\Delta APA + \Delta INV + \Delta ACP) + (CGS + SAE)\big) + \big((\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN + \Delta OLA + \Delta OLL) + (OOE + DDA + DOP + ECR + ECH + OIA + EIE + OIE + (\Delta DTX + \Delta ITP + \Delta PRC + ITX) \quad (21)$$

In this thesis, the autoregressions are fully subdivided into lagged components so that we can obtain unique estimates from the marginal responses on each of the variables.

There is strong theoretical argumentation that describes how future accounting aggregates can be explained through the lagged contribution of their subdivided flows. The intuition behind the regressions in this thesis, for instance in eq(26), eg(29) and eq(32), for using the *lagged* variables is highly motivated by the work of Ohlson (1999). The seminal work by Ohlson (1999) develops a *first* order autoregressive process of earnings by collapsing lagged earnings into transitory and permanent components.

Lagged levels are expected to act as instruments to the endogenous regressors, since it is reasonable to expect that past economic values are correlated to current observed values and, by definition, are predetermined and therefore uncorrelated to the fitted error term at time *t*.

Therefore, the first major regression of this thesis in the prediction of future operating cash flows can be written as follows:

$$OCF_{it} = \alpha_1 + \beta_1 \Delta ARE_{i\,t-1} + \beta_2 SAL_{i\,t-1} + \beta_3 \Delta APA_{i\,t-1} + \beta_4 \Delta INV_{i\,t-1} + \beta_5 \Delta ACP_{i\,t-1} + \beta_6 CGS_{i\,t-1} + \beta_7 SAE_{i\,t-1} + \beta_8 \Delta ITP_{i\,t-1} + \beta_9 \Delta PRC_{i\,t-1} + \beta_{10} \Delta DTX_{i\,t-1} + \beta_{11} ITX_{i\,t-1} + \beta_{12} \Delta OCA_{i\,t-1} + \beta_{13} \Delta OCL_{i\,t-1} + \beta_{14} \Delta OLA_{i\,t-1} + \beta_{15} \Delta OLL_{i\,t-1} + \beta_{16} \Delta PRE_{i\,t-1} + \beta_{17} \Delta DIN_{i\,t-1} + \beta_{18} OOE_{i\,t-1} + \beta_{19} DOP_{i\,t-1} + \beta_{20} ECR_{i\,t-1} + \beta_{21} ECH_{i\,t-1} + \beta_{22} EIE_{i\,t-1} + \beta_{23} OIE_{i\,t-1} + \beta_{24} DDA + \beta_{25} OIA + \sum_{k=26}^{33} \beta_k IND_{i\,t} + \varepsilon_1 \quad (22)$$

The second equation is about investing activities, and is explained in the following section.

Investing cash flow

In the second major equation, which is about a firm's investing cash flow, the only presented variable is capital expenditure (net of asset disposals) (CED). Thus, we have:

$$ICF \equiv CED \quad (23)$$

By decomposing the above equation, it can be written as:

$$ICF \equiv (\Delta PPE + \Delta LTR + \Delta OIN + \Delta IUS) + (EGL) \quad (24)$$

Consequently, the second major regression of this study, showing the prediction of future investing cash flow, will be as follows:

$$ICF_{it} = \alpha_2 + \beta_{26}\Delta PPE_{i\,t-1}\,\beta_{27}\Delta LTR_{i\,t-1} + \beta_{28}\Delta OIN_{i\,t-1} + \beta_{29}\Delta IUS_{i\,t-1} + \beta_{30}\Delta EGL_{i\,t-1} + \sum_{k=31}^{38}\beta_k IND_{i\,t} + \varepsilon_2 \quad (25)$$

The final equation is about financing cash flows, and is explained in the following section.

Financing cash flow

As mentioned above, financing cash flow is obtained from two sources. These are: debt issues (net of debt repayments) (*DIR*) and new capital (net of dividends) (*NCD*). Therefore, financing cash flow is written as:

$$FCF \equiv DIR + NCD \quad (26)$$

By decomposing the above equation, it can be displayed as follows:

$$FCF \equiv \big((\Delta STD + \Delta LTD + \Delta CSI) + (IED + IIN)\big) + \big((\Delta SEQ + \Delta PST + \Delta DPA + \Delta NER + \Delta MIN + NI + MIE + PEI \quad (27)$$

Consequently, the third major equation of this thesis regarding the prediction of future financing cash flow can be written as follows:

$$FCF_{it} = \alpha_3 + \beta_{31}\Delta STD_{i\,t-1} + \beta_{32}\Delta LTD_{i\,t-1} + \beta_{33}\Delta CSI_{i\,t-1} + \beta_{34}IED_{i\,t-1} + \beta_{35}IIN_{i\,t-1} + \beta_{36}\Delta NER_{i\,t-1} + \beta_{37}\Delta MIN_{i\,t-1} + \beta_{38}\Delta SEQ_{i\,t-1} + \beta_{39}\Delta PST_{i\,t-1} + \beta_{40}\Delta DPA_{i\,t-1} + \beta_{41}NI_{i\,t-1} + \beta_{42}MIE_{i\,t-1} + \beta_{43}PEI_{i\,t-1} + \sum_{k=44}^{51}\beta_k IND_{i\,t} + \varepsilon_3 \quad (28)$$

3-2-4 Imposing constraints on cash flow predictions

Up to this point, in the multiple situations described, the terminology of *double entry constraint* has been applied to specify that the sum of several variables is equal to zero. For instance, in the financial statements, after imposing the relevant signs, all variables are summed to zero.

One further step has been presented, as the double entry constraint terminology has also been applied to show that the sum of variables in each line-item creating cash flow statement components is equal to zero. For instance, the summation of sales, accounts receivable and customer receipts is equal to zero.

However, in fact, the major *constraint* equation is a mixture of the above constraints, when the sum of *all* variables in *all* financial statements is equal to zero. This issue only can happen when the sum of all seven *created* variables of the cash flow statement is equal to zero, as each one of the seven variables is calculated using the change in balance sheet items and income statement items.

The main concern of this thesis, in comparison to previous studies on the prediction of future cash flows, is to employ this major constraint equation as the *fourth* equation, in addition to the three previously presented equations regarding

different aspects of cash flows (operating, investing and financing cash flow equations).

Therefore, this *major constraint* equation can be written as follows:

$$OCF + ICF + FCF \equiv 0 \qquad (29)$$

By decomposing the above equation, we have:

$$(REC + PAY + OOC + TXP) + (CED) + (DIR + NCD) = 0 \qquad (30)$$

Finally, the constraint equation can be written as follows:

$$(\Delta ARE + SAL) + \big((\Delta APA + \Delta INV + \Delta ACP) + (CGS + SAE)\big) + \big((\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN + \Delta OLA + \Delta OLL + OOE + DDA + DOP + ECR + ECH + OIA + EIE + OIE + (\Delta DTX + \Delta ITP + \Delta PRC + ITX) + \Delta PPE + \Delta LTR + \Delta OIN + \Delta IUS + EGL + \Delta STD + \Delta LTD + \Delta CSI + IED + ICP + IIN + \Delta SEQ + \Delta PST + \Delta NER + \Delta MIN + NI + MIE + PEI = 0 \qquad (31)$$

In the following section it will be clarified how the above equation can be used in addition to the set of simultaneous linear regression equations in the Seemingly Unrelated Regression matrix by giving a constraint to the system that the sum of all the variables' coefficients is equal to zero.

In other words, the above equation notes that while three *separate* equations have been presented to predict future cash flows, all the applied variables come

from the change in balance sheets and the income statement. Consequently, the created cash flow statement items are related to each other, and therefore each of these equations is associated with the other equations.

In this thesis, given the use of the *SUR* system, following Zellner (1962), and as the predictor variables in every single equation (OPC, ICF and FCF) are taken from the same set of contemporaneous financial statements, this leads to contemporaneously correlated residuals between each of the regressions, i.e. the regressions OPC and ICF, $E(u_{\boldsymbol{OPC}it}, u_{\boldsymbol{ICF}it}) \neq 0$, the regressions OPC and FCF, $E(u_{\boldsymbol{OPC}it}, u_{\boldsymbol{FCF}it}) \neq 0$, and finally the regressions ICF and FCF, $E(u_{\boldsymbol{ICF}it}, u_{\boldsymbol{FCF}it}) \neq 0$. In fact, the standalone *OLS* regression is unable to model the deterministic character of contemporaneous financial statement variables as it does not consider a generalised system of simultaneous equations that can be applied to estimations, and consequently does not consider cross-equation error correlation between the residuals, as it looks into each of the equations individually. This issue matters, as will be explained in section 5-1-1, due to the inconsistent regression results for the *OLS* regression.

In addition, in section 5-1-5, the cross-equation error correlation will be examined using the *Breusch and Pagan Lagrange Multiplier test*, indicating that all residuals are significantly correlated with each other, which addresses the usefulness of the *SUR* system.

In the following section, the regressions that apply to the prediction of future earnings (operating and non-operating income) will be explained.

3-3 Applying the constrained prediction of future earnings

The first part of this thesis is about the prediction of future cash flows, and all equations applied have already been explained. The aim of the second part of this thesis is to investigate the prediction of future earnings. To do this, the applied equations, including the *constraint* equations, will be presented in this section.

The main difference between this section and the section regarding the prediction of future cash flows is the number of calculated variables of the cash flow statement. While the number of calculated cash flow variables in the prediction of future cash flows is seven, in order to differentiate between a small number of applied operating and non-operating variables from the income statement in creating other operating cash flow (*OOC*), it is decomposed into two components. Therefore, the number of calculated cash flow variables in the prediction of future earnings increases to eight. This issue will be explained in detail later, while only the applied equations will be explained in this section.

As has been mentioned, the applied income statement variables in calculating the other operating cash flow (*OOC*) are listed as: other operating expenses (*OOE*), depreciation, depletion and amortisation (*DDA*), discontinued operations (*DOP*), extraordinary credit (*ECR*), extraordinary charges (*ECH*), other income (expenses) after tax (*OIA*), equity interest earnings (*EIE*) and other income (expenses) net (*OIE*). The operating and non-operating variables of the income statement are mixed together; therefore, as the next step, in order to decompose them into operating and non-operating categories, other operating cash flow (*OOC*) is divided

into two components: other operating cash flow (OOC) and exceptional cash flow (ECF).

The name 'other operating cash flow' (OOC) remains as before for the first group, but the employed variables of the income statement in the calculation process are limited to just other operating expenses (OOE) and depreciation, depletion and amortisation (DDA), as these are the only operating variables among them.

The equation is presented as:

$$(\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN) + (OOE + DDA) = OOC \qquad (32)$$

As before, in order to show the double entry constraint, the location of the other operating cash flow (OOC) needs to be changed, as below:

$$(\Delta OCL + \Delta OCA + \Delta PRE + \Delta DIN) + (OOE + DDA) - OOC = 0 \qquad (33)$$

The second variable is exceptional cash flow (ECF), which contains those variables from the income statement that are *not* operating, as follows:

$$(\Delta OLA + \Delta OLL) + (DOP + ECR + ECH + OIA + EIE + OIE) = ECF \qquad (34)$$

The double entry constraint can be seen by changing the location of the exceptional cash flow (ECF), as below:

$$(\Delta OLA + \Delta OLL) + (DOP + ECR + ECH + OIA + EIE + OIE) - ECF = 0 \quad (35)$$

In the next part, the equations for operating and non-operating income will be explained.

3-3-1 Operating income

In the prediction of future income two equations are presented for operating and non-operating income. In this section, operating income and its components are reviewed, while the following section will deal with non-operating income.

Operating income in income statements refers to all revenues less cost of goods sold and related operating expenses that apply to the day-to-day operating activities of the company. It excludes financially-related items (i.e. interest income, dividend income and interest expenses), extraordinary items and taxes. Operating income is divided into three major identities: sales (SAL), purchases (PUR) and other expenses and depreciation (OOD), so the operating income equation can be written as:

$$OPI \equiv SAL + PUR + OOD \tag{36}$$

Each of the above variables can be obtained by reconciling the other financial statement variables. For instance, the first variable, sales (SAL), is measured by taking the difference between customer receipts (REC) and change in accounts receivable (ΔARE). Hence sales (SAL) can be displayed as:

$$SAL \equiv REC - \Delta ARE \tag{37}$$

The next identity in the operating income equation is purchases (PUR), and is displayed as:

$$PUR \equiv PAY\ - \Delta APA - \ \Delta INV - \Delta ACP \quad (38)$$

The final identity in the operating income equation is other expenses and depreciation (OOD), shown as:

$$OOD \equiv OOC\ - \Delta OCA - \Delta OCL + \Delta PRE + \Delta DIN \quad (39)$$

The components of operating income would disaggregate as:

$$OPI \equiv (REC - \Delta ARE) + (PAY\ - \Delta APA - \ \Delta INV - \Delta ACP) + (OOC\ - \Delta OCA - \Delta OCL + \Delta PRE + \Delta DIN \quad (40)$$

3-3-2 Non-operating income

Non-operating income represents income that is not derived from the basic business of a company, and is the second equation in the prediction of future income. Its components are listed as: other non-operating income (expenses) (ONE), income tax (ITX), extraordinary items and gain (loss) on sale of assets (EGL), increase (decrease) net interest (ICH) and minority interest in earnings (MIE). They can be structured in an equation as follows:

$$NOI \equiv ONE + ITX + EGL + ICH + MIE \quad (41)$$

As before, each of the above components can be driven by taking the difference between the cash flow statement variables and the change in balance sheet statement variables. For instance, the first component, other non-operating income (expenses) (ONE), is calculated as:

$$ONE \equiv ECF - \Delta OLA - \Delta OLL \quad (42)$$

The next variable, income tax (ITX), is presented as:

$$ITX \equiv TXP - \Delta DTX - \Delta ITP - \Delta PRC \quad (43)$$

Extraordinary items and gain (loss) on sale of assets (EGL), as the third variable, is displayed as:

$$EGL \equiv ADN - \Delta PPE - \Delta LTR - \Delta OIN \quad (44)$$

Increase (decrease) in net interest (ICH) is shown as:

$$ICH \equiv DIR - \Delta STD - \Delta LTD - \Delta CSI \quad (45)$$

The final component of non-operating income is minority interest in earnings (MIE), which can be calculated as:

$$MIE \equiv NCD - \Delta MIN - \Delta PST - \Delta SEQ - \Delta DPA - NI \quad (46)$$

Finally, the non-operating income equation can be disaggregated as:

$$NOI \equiv (ECF - \Delta OLA - \Delta OLL) + (TXP - \Delta DTX - \Delta ITP - \Delta PRC) + (ADN - \Delta PPE - \Delta LTR - \Delta OIN) + (DIR - \Delta STD - \Delta LTD - \Delta CSI) + (NCD - \Delta MIN - \Delta PST - \Delta SEQ - \Delta DPA - NI \quad (47)$$

3-3-3 Imposing constraints on the prediction of future earnings

Consistent with the presented *constraint* equation in the prediction of future cash flows, here another constraint equation is presented to show that the sum of the income statement variables, decomposed into operating and non-operating income plus net income, is equal to zero. So we have:

$$OPI + NOI \equiv NI \quad (48)$$

Subsequently, disaggregating the above equation can be written as follows:

$$(SAL + PUR + OOD) + (ONE + ITX + EGL + ICH + MIE) \equiv NI \quad (49)$$

In addition to the above equation, the sum of minority interest in earnings (*MIE*) and net income (*NI*) can be named as the increase (decrease) in current profits (losses) (*IDP*). A detailed explanation of this is presented in Table 3-11.

By changing the location of net income (NI) in the above equation to the left hand side, it can be written as follows:

$$(SAL + PUR + OOD) + (ONE + ITX + EGL + ICH + IDP) \equiv 0 \qquad (50)$$

Moving one step further in disaggregating the above equation, it can be expressed as follows:

$$((REC - \Delta ARE) + (PAY - \Delta APA - \Delta INV - \Delta ACP) + (OOC - \Delta OCA - \Delta OCL + \Delta PRE + \Delta DIN) + ECF - \Delta OLA - \Delta OLL + TXP - \Delta DTX - \Delta ITP - \Delta PRC + ADN - \Delta PPE - \Delta LTR - \Delta OIN + DIR - \Delta STD - \Delta LTD - \Delta CSI + NCD - \Delta MIN - \Delta PST - \Delta SEQ - \Delta DPA \equiv 0 \qquad (51)$$

As can be seen, the above equation shows the double entry constraint for all variables, indicating that the sum of the income statement items is equal to zero. In other words, the sum of *all* the variables in the cash flow statement and the change in balance sheet statements is equal to zero. In the following section an explanation is given of how this additional equation in both the operating and non-operating equations leads to an unbiased estimation.

It is now time to demonstrated the regression equations in the prediction of operating and non-operating income for one year ahead, considering industries as dummy variables.

- **Operating income regression**

$$OPI = \alpha_1 + \beta_1 REC_{i\,t-1} + \beta_2 \Delta ARE_{i\,t-1} + \beta_3 PAY_{i\,t-1} + \beta_4 \Delta APA_{i\,t-1} + \beta_5 \Delta INV_{i\,t-1} + \beta_6 \Delta ACP_{i\,t-1} + \beta_7 OOC_{i\,t-1} + \beta_8 \Delta OCA_{i\,t-1} + \beta_9 \Delta OCL_{i\,t-1} + \beta_{10} \Delta PRE_{i\,t-1} + \beta_{11} \Delta DIN_{i\,t-1} + \sum_{k=12}^{19} \beta_k IND_{i\,t} + \varepsilon_1 \quad (52)$$

- **Non-operating income regression**

$$NOI = \alpha_2 + \beta_{12} TXP_{i\,t-1} + \beta_{13} \Delta DTX_{i\,t-1} + \beta_{14} \Delta ITP_{i\,t-1} + \beta_{15} \Delta PRC_{i\,t-1} + \beta_{16} ECF_{i\,t-1} + \beta_{17} \Delta OLA_{i\,t-1} + \beta_{18} \Delta OLL_{i\,t-1} + \beta_{19} CED_{i\,t-1} + \beta_{20} \Delta PPE_{i\,t-1} + \beta_{21} \Delta LTR_{i\,t-1} + \beta_{22} \Delta OIN_{i\,t-1} + \beta_{23} DIR_{i\,t-1} + \beta_{24} \Delta STD_{i\,t-1} + \beta_{25} \Delta LTD_{i\,t-1} + \beta_{26} \Delta CSI_{i\,t-1} + \beta_{27} NCD_{i\,t-1} + \beta_{28} \Delta MIN_{i\,t-1} + \beta_{29} \Delta PST_{i\,t-1} + \beta_{30} \Delta SQE_{i\,t-1} + \beta_{31} \Delta DPA_{i\,t-1} + \sum_{k=32}^{39} \beta_k IND_{i\,t} + \varepsilon_2 \quad (53)$$

In the following section of the study, an Ordinary Least Square (*OLS*) Regression and a Seemingly Unrelated Regression (*SUR*) are compared with each other.

3-4 The structural statistical model

The key proposition of this thesis lies in the treatment of financial statements as a matrix of codetermined information bounded by the double entry constraint. To account for the highly-structured information set, a generalised structural system for estimating future cash flows and future earnings within the deterministic relationships of double entry is presented. The main concern in this section is to compare this generalised structural system, as the Seemingly

Unrelated Regression (*SUR*) system, with the more traditional method that is shown to violate basic double entry rules, the Ordinary Least Square (*OLS*) Regression, and to show that when using *SUR*, the empirical applications yield increased precision and improved efficiency in estimation.

3-4-1 The generalised framework: Seemingly Unrelated Regression

In the literature, the nature of the variance-covariance matrix of the *SUR* estimator and the likely magnitude of the efficiency gain have been investigated by Binkley (1982), Zellner (1962) and finally Kmenta (1997).

Following the most recent paper by Kmenta (1997), in this study, a structural system of regressions is used, where the endogenous variables are defined as those drawn from the financial statements of firm *i* and year *t*, and where the estimator variables are defined either as lagged endogenous variables from past financial statements, or as other exogenous covariates that are observed outside the accounting system.

In the previous section, it was demonstrated that the reconciliation of financial statements is used to show that all accounting items are contemporaneously codetermined through the resolution of multiple accounting identities using the equations m = 1,2,…,M.

M is the equation number; for instance, β_{M3} means the third *beta* in the equation number M, or, as another example, β_{13} means the third *beta* in the first equation *OPC*. Therefore, as it is possible to draw a linear combination that describes the endogenous relationship between these accounting variables, a

structural regression system for specifying the expected variation of financial variables can take the following form:

$$\begin{bmatrix} y_{1it} \\ y_{2it} \\ y_{3it} \\ . \\ . \\ . \\ y_{Mit} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}_{11}^{s} \boldsymbol{\beta}_{12}^{s} \dots \boldsymbol{\beta}_{1L}^{s} \\ \boldsymbol{\beta}_{21}^{s} \boldsymbol{\beta}_{22}^{s} \dots \boldsymbol{\beta}_{2L}^{s} \\ \boldsymbol{\beta}_{31}^{s} \boldsymbol{\beta}_{32}^{s} \dots \boldsymbol{\beta}_{3L}^{s} \\ . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \\ . \quad . \quad . \\ \boldsymbol{\beta}_{M1}^{s} \boldsymbol{\beta}_{M2}^{s} \dots \boldsymbol{\beta}_{ML}^{s} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_{1it-s} \\ \boldsymbol{y}_{2it-s} \\ \boldsymbol{y}_{3it-s} \\ . \\ . \\ . \\ \boldsymbol{y}_{Lit-s} \end{bmatrix} + \begin{bmatrix} 1 \\ x_{1it} \\ x_{2it} \\ x_{3it} \\ . \\ . \\ . \\ x_{Kit} \end{bmatrix} + \begin{bmatrix} u_{1it} \\ u_{2it} \\ u_{3it} \\ . \\ . \\ . \\ u_{Mit} \end{bmatrix} \tag{54}$$

In this thesis, dependent variables can be estimated using other information that comes from the lagged variables *L*, and from exogenous variables *K*, which are associated with the firm's environment, for instance the industry in which the firm works, and are not derived from double entry.

In the structural system, endogenous accounting variables y_{mit} are jointly estimated. The first row of the above matrix represents a regression of OPC_{it} on other variables. It is expected that OPC_{it} will be seen as the first elements in the vector of endogenous variables, so we have $y_{1\,it} = OPC_{it}$. Consequently, as the second and the third rows of the matrix have to be allocated to ICF_{it} and FIC_{it}, we have $y_{2\,it} = ICF_{it}$ and $y_{3\,it} = FCF_{it}$.

As all the explanatory variables for the regressions are comprised of only predetermined covariates, the only link between the equations is channelled through their error terms. This case is known as the Seemingly Unrelated Regression (*SUR*). This system can be explained as follows:

$$\boldsymbol{y}_{m\times 1}^{t} = \boldsymbol{\beta}_{m\times l}^{s} \boldsymbol{y}_{l\times 1}^{t-s} + \boldsymbol{x}_{k\times 1}^{t} + \boldsymbol{u}_{m\times 1}^{t} \tag{55}$$

The superscripts show the time dimension, and the subscripts the matrix dimension. The vector $\boldsymbol{y}_{m\times 1}^{t}$ relates to the endogenous variables, $\boldsymbol{y}_{l\times 1}^{t-s}$ relates to the lagged endogenous variables of order s, $\boldsymbol{x}_{k\times 1}^{t}$ shows the strictly exogenous variables, and finally $\boldsymbol{u}_{m\times 1}^{t}$ presents the disturbance terms. There are unique stochastic disturbances u_{mit} that are assigned to each regression m, which satisfy the standard normal assumption that error terms have a mean zero; that is, $Eu_{1it} = Eu_{2it} = 0$.

To ensure that the matrixes are not singular, the sum of the structural coefficients must be set to zero. The structural variance-covariance matrix (VCV) of disturbance terms $\boldsymbol{u}_{m\times 1}^{t}$is also given as:

$$\boldsymbol{\sigma}_{m\times m} \otimes \boldsymbol{I}_N = E[\boldsymbol{u}_{m\times 1}^{t}\boldsymbol{u}_{m\times 1}^{\prime t}] \tag{56}$$

Where $\otimes$ is the Kronecker product and $\boldsymbol{I}_N$ is an identity matrix of order $N=I\times T$ that gives a spherical form to the VCV. I and T are the appropriate weight functions for the identity matrix $\boldsymbol{I}_N$. The identity matrix $\boldsymbol{I}_N$ applies a spherical form to the variance-covariance matrix, i.e. it assumes no heteroscedasticity or autocorrelation. At this stage, we are not concerned with investigating a more precise empirical approximation to the variance-covariance matrix, but we are well aware that a firm-year dataset is likely to be contaminated with heteroscedasticity and/or autocorrelation, and it is possible to correct for either by substituting $\boldsymbol{I}_N$ for some appropriate weight function.

In the prediction of future cash flows, we have the Seemingly Unrelated Regression (*SUR*), as shown below:

$$\boldsymbol{y}_{4\times1}^{t} = \boldsymbol{\beta}_{4\times43}\boldsymbol{y}_{43\times1}^{t-1} + \boldsymbol{x}_{k\times1}^{t} + \boldsymbol{u}_{4\times1}^{t} \tag{57}$$

Nonetheless, it is expected that a significant contemporaneous cross-equation correlation will be seen within $\boldsymbol{u}_{4\times1}^{t}$ as the result of commonly omitted information. The *VCV* is also reduced to $\boldsymbol{\sigma}_{m\times m} \otimes \boldsymbol{I}_N = E[\boldsymbol{u}_{4\times1}^{t}\boldsymbol{u}_{4\times1}^{\prime t}]$ and the parameters are estimated to be as follows:

$$\widehat{\boldsymbol{\beta}}_{SUR} = (\boldsymbol{y}_{43\times1}^{\prime t-1}(\widehat{\boldsymbol{\sigma}}_{4\times4}^{-1} \otimes \boldsymbol{I}_N)\boldsymbol{y}_{43\times1}^{t-1})^{-1}(\boldsymbol{y}_{43\times1}^{\prime t-1}(\widehat{\boldsymbol{\sigma}}_{4\times4}^{-1} \otimes \boldsymbol{I}_N)\boldsymbol{y}_{4\times1}^{t}) \tag{58}$$

The estimated $\hat{\gamma}_{SUR}$ is similarly derived. Zellner (1962) notes that $\widehat{\boldsymbol{\beta}}_{SUR}$ is more efficient and more precise than *OLS* when *N*>20. Moreover, it is important to note that $\widehat{\boldsymbol{\beta}}_{SUR}$ in the above equation may be the same as the single-equation *OLS* estimator only when:

- There is no contemporaneous correlation between disturbance terms;
- The expectation function across the equations is exactly the same;
- As noted by Zellner (1962), the repressors in one equation are a nested set of those in another equation, in which case the nested parameters can be equivalently recovered by *OLS*. Furthermore, in another study Srivastava and Giles (1987) suggest that even in an equivalent condition where the *SUR* system is equal to *OLS*, the *SUR* system design is still superior, as it allows for testing cross-equation hypotheses (which are referred to as

'unconstrained models' in this thesis), and for the imposition of cross-equation parameter constraints (which are referred to as 'constraint models' in this thesis).

3-4-2 Predetermined information

In this thesis, two different sets of predictor variables are presented in the forecast of both cash flows and earnings: endogenous accounting variables and exogenous variables. The discussion below is based on the above matrix of the generalised structural framework.

- Based on the inherent economic character of a company, endogenous accounting variables are expressed as a function of their lagged values $\boldsymbol{y}_{l\,it-s}$ for the variables $l = 1,2,\dots,\mathrm{L} \geq M$ and lag s = 1,2,..., T -1. These lagged values are by definition predetermined and at least weakly exogenous. It is also important to note that $\boldsymbol{y}_{l\,it-s}$ is an inner vector of dimension $s \times 1$, and its respective structural coefficient $\boldsymbol{\beta}_{ml}^{s}$ is an inner vector of dimension $1 \times s$.
- Moreover, there are other variables that are exogenous: $\boldsymbol{x}_{k\,it}$ with *k = 1,2,...,k* . In fact, these are the variables outside the accounting system and it is assumed that they help to explain$\boldsymbol{y}_{m\,it}$. Industry binary indicators are applied in this thesis as exogenous variables; furthermore, the structural coefficients on the strictly exogenous variables $\boldsymbol{y}_{mk}$, also allow for the inclusion of the model intercepts $\boldsymbol{y}_{m0}$, as is shown by the unit constant as the first element in the vector of $\boldsymbol{x}_{k\,it}$ variables.

In the prediction of future cash flows, four key identities have been applied: operating cash flows (*OPC*), investing cash flows (*ICF*), financial cash flows (*FCF*) and finally the cash flow equilibrium condition:

$$OPC + ICF = -FCF \tag{59}$$

As a result of the constraint equation, presented in the previous chapter as:

$$OPC + ICF + FCF = 0 \tag{60}$$

These identities are presented in the different rows of the generalised structural system for imposing the double entry constraint on the parameters, so the coefficients on the lagged accounting variables add up to zero: $\Sigma \boldsymbol{\beta}_{ml}^{s} \equiv 0$.

Ultimately, there are several important issues with regard to the error terms assumptions that have to covered.

3-4-3 Estimation matrix for future cash flows

A joint autoregression of order one among operating cash flows (*OPC*), investing cash flows (*ICF*) and financing cash flows (*FCF*), using information from their respective lagged accounting identity components within a structural system is presented in this section. As OPC, ICF and FCF are jointly determined, it is expected that a negative contemporaneous cross-equations correlation will be seen between *FCF* and both *OPC* and *ICF*. Each of these equations separately omits

useful information, and is absorbed by their linked error terms, such as their codetermination of both income statements and changes in balance sheet statements. Therefore, all three equations have to be estimated jointly as a comprehensive system, indicating that all the covariates are linearly related through a double entry system. The structural system for the prediction of future cash flows is presented below:

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}OCF_{it}\\ICF_{it}\\FCF_{it}\\0\end{bmatrix}=\begin{bmatrix}\beta_{11}\beta_{12}\beta_{13}\dots\beta_{123}\beta_{124}\beta_{125}&0&0&0&0&0&0&0&0\dots&0&0&0\\0&0&0\dots&0&0&0&\beta_{226}\beta_{227}\beta_{228}\beta_{229}\beta_{230}&0&0&0\dots&0&0&0\\0&0&0\dots&0&0&0&0&0&0&0&\beta_{331}\beta_{332}\beta_{333}\dots&\beta_{341}\beta_{342}\beta_{343}\\\beta_{11}\beta_{12}\beta_{13}\dots&\beta_{123}\beta_{124}\beta_{125}\beta_{226}\beta_{227}\beta_{228}\beta_{229}\beta_{230}\beta_{331}\beta_{332}\beta_{333}\dots&\beta_{341}\beta_{342}\beta_{343}\end{bmatrix}\begin{bmatrix}\Delta ARE_{i\,t-1}\\SAL_{i\,t-1}\\\Delta APA_{i\,t-1}\\\vdots\\EIE_{i\,t-1}\\OIE_{i\,t-1}\\DDA_{i\,t-1}\\\Delta PPE_{i\,t-1}\\\Delta LTR_{i\,t-1}\\\Delta OIN_{i\,t-1}\\\Delta IUS_{i\,t-1}\\\Delta STD_{i\,t-1}\\\Delta LTD_{i\,t-1}\\\Delta CSI_{i\,t-1}\\\vdots\\\Delta DPA_{i\,t-1}\\NI_{i\,t-1}\\MIE_{i\,t-1}\end{bmatrix}+\begin{bmatrix}\gamma_{10}&\gamma_{1j}\\\gamma_{20}&\gamma_{2j}\\-\gamma_{30}&-\gamma_{3j}\\0&0\end{bmatrix}\begin{bmatrix}1\\J_{it}\end{bmatrix}+\begin{bmatrix}u_{1\,it}\\u_{2\,it}\\u_{3\,it}\\0\end{bmatrix} \quad (61)$$

The three equations, OPC_{it}, $ICF_{it}$and FCF_{it}, are named m = 1,2 and 3, with structural coefficients $\alpha_{11} = \alpha_{22} = \alpha_{33} = 1$. All off-diagonal vectors are also set to zero as no endogenous variables are applied in the explanatory section.

Regarding the endogenous variables, the $\boldsymbol{y}_{l\,it-s}$ vectors are only the first order lagged values, containing 39 variables and related to the $\boldsymbol{\beta}_{ml}$ coefficients; for instance, in equation *m=1 (OPC_{it}) l = 1,2,3,…,25*, in equation *m=2 (ICF_{it}) l=26 ,27 ,28, 29,30* and finally in equation *m=3 (FCF_{it}) l=31, 32, 33, …, 43.*

Regarding the exogenous variables, $\boldsymbol{x}_{k\,it}$ shows the list of these variables. In this thesis, intercepts are entered as$\boldsymbol{\gamma}_{10}$, $\boldsymbol{\gamma}_{20}$ and $\boldsymbol{\gamma}_{30}$, which are indicated by unity plus industry fixed effects $\boldsymbol{J}it$ and with coefficients of γ_{1j}, γ_{2j} and γ_{3j}. Including the $\boldsymbol{J}it$ ensures that any fixed difference that may exist between the *j* industries in the prediction of future cash flows is considered. It is important that both intercept and industry coefficients are the same across the three equations, but with positive signs for *m=1,2* and a negative sign for *m=3.* This constraint ensures that the fixed effects are estimated at the same level, a treatment that is consistent with the fact that industry effects for all three equations are at the same magnitude that is reflected with different signs.

The final equation, *m=4,* relates to the double entry constraint that imposes the restriction that the sum of all $\boldsymbol{\beta}_{ml}$ coefficients is equal to zero, and consequently the zero value of the $\boldsymbol{y}_{l\,it-s}$ vector of endogenous variables. In the end, as the result of the equilibrium condition, the sign of one of the equations needs to be inverted, and here are the signs for FCF_{it} (from$-\beta_{331}$, $-\beta_{332}$…), therefore the double entry holds and the fitted marginal responses converge with higher precision to their theoretically expected values. We make sure to inform the estimation that the coefficients $\boldsymbol{\gamma_{10}}$, $\boldsymbol{\gamma_{20}}$ and $\boldsymbol{\gamma_{1j}}$, $\boldsymbol{\gamma_{2j}}$ are estimated at the same level across the three autoregressions but with opposite signs, i.e. $-\boldsymbol{\gamma_{30}}$ and $-\boldsymbol{\gamma_{3j}}$. This treatment is consistent with the fact that equation (61) describes mirror autoregressions with opposite sign mean effects that are captured by the intercepts, but also with $\boldsymbol{J}it$ characterising fixed factors that affect *OPC, ICF* and *FCF*

in the same manner, which in turn jointly give rise to the equilibrium between all of them.

3-4-4 Estimation matrix for future earnings

In this section the joint autoregression of order one in operating income (*OIN*) and non-operating income (*NOI*), as with the previous section, is presented. It is expected that a negative contemporaneous cross-equations correlation will be seen between *OPI* and *NOI*, as they are jointly determined.

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}OPI_{it}\\NOI_{it}\\0\end{bmatrix}=\begin{bmatrix}\beta_{11}&\beta_{12}&\beta_{13}&\dots&\beta_{19}&\beta_{110}&\beta_{111}&0&0&0&\dots&0&0&0\\0&0&0&\dots&0&0&0&\beta_{212}&\beta_{213}&\beta_{214}&\dots&\beta_{229}&\beta_{230}&\beta_{231}\\\beta_{11}&\beta_{12}&\beta_{13}&\dots&\beta_{19}&\beta_{110}&\beta_{111}&\beta_{212}&\beta_{213}&\beta_{214}&\dots&\beta_{229}&\beta_{230}&\beta_{231}\end{bmatrix}\begin{bmatrix}REC_{it-1}\\\Delta ARE_{it-1}\\PAY_{it-1}\\\vdots\\\Delta OCL_{it-1}\\\Delta DIN_{it-1}\\\Delta PRE_{it-1}\\TXP_{it-1}\\\Delta DTX_{it-1}\\\Delta ITP_{it-1}\\\vdots\\\Delta PST_{it-1}\\\Delta SEQ_{it-1}\\\Delta DPA_{it-1}\end{bmatrix}+\begin{bmatrix}\gamma_{10}&\gamma_{1j}\\-\gamma_{20}&-\gamma_{2j}\\0&0\end{bmatrix}\begin{bmatrix}1\\J_{it}\end{bmatrix}+\begin{bmatrix}u_{1it}\\u_{2it}\\0\end{bmatrix} \quad (62)$$

In the above matrixes, in equation *m=1* (OPI_{it}) *l = 1,2,3,…,11* and equation *m=2* (NOI_{it}) *l=11, 12, 13, …,31,* intercepts are entered as $\boldsymbol{\gamma}_{10}$ and $\boldsymbol{\gamma}_{20}$, which are indicated by unity plus industry fixed effects $\boldsymbol{J}_{it}$ and with coefficients of $\gamma_{1j}$and γ_{2j}. Both intercept and industry coefficients are the same across the two equations, but with a positive sign for *m=1* and a negative sign for *m=2.* The final equation, *m=3,* as before, is related to the double entry constraint.

In the remainder of this thesis, the *OLS* regression that does not observe the equations simultaneously is named the *OLS* model (1), the Unrestricted *SUR* model that does not observe the double entry constraint is named the *SUR* model (2), and finally the double entry constraint *SUR* model is named the *decSUR* Model (3), and observes both the double entry constraint as an endogenous constraint, and intercepts and industries as exogenous constraints.

The aim of the next chapter is to present both the sampling distribution and industry distribution.

CHAPTER 4 - DATA AND SAMPLE SELECTION

CHAPTER 4- Data and sample selection

The aim of the Chapter three was to illustrate the logic of imposing signs onto variables in the different financial statements, showing that the sum of variables in each financial statement is equal to zero as a result of the double entry system. In the first part of this chapter, the sample selection process for the US market will be described, and, in the second part of the chapter, the industry distribution for the selected sample will be considered. Finally, in the third section, the descriptive statistics will be presented.

4-1 Sample selection

The aim of this section is to show how the usable sample has been selected for the US market. The data used in this thesis were downloaded from the WorldScope database, which is available as part of the Thomson One Banker online resource.

Before explaining the validity of the articulation exercise in the next section, it is first necessary to understand the implications of missing values in our data. As a set of financial statement variables is investigated in this thesis, the existence of even one missing value leads to a missing value in their summation. For instance, if just one line-item in assets is given in the data source as a missing value, it may be expected that a missing value would also be observed in their summation as total assets. However, sometimes these missing values are replaceable with zero, in cases where, after replacement, the relevant accounting identity holds.

In WorldScope, there are several items that tend to be missing, such as reserves in income statements. The problem is that we do not know whether these items were originally zero or missing. To find out the nature of these missing values, a simple procedure is applied. Firstly, it is assumed that all the missing values are zero, so they are replaced with zero. Next, the sum of these values is calculated. The calculated summations after treating all missing values as zero are then compared with the actual downloaded aggregate from the dataset. For instance, it is evident that the summation of cash and short-term investment (*CSI*), accounts receivable (*ARE*), inventories (*INV*), prepaid expenses (*PRE*) and other current assets (*OCA*) has to be equal to current assets (*TCA*). Therefore, if items like prepaid expenses (*PRE*) and other current assets (*OCA*) are missing values, it is not reasonable to take these observations out of our usable sample without validation. Instead, what we have done is to replace these two missing values (*PRE* and *OCA*) with zero, in order to see whether the calculated summation is still equal to the original value of total current assets (*TCA*). This helps to avoid losing many observations, especially in the case of this study, as there are many variables involved. This procedure has been applied to maximise the completeness of all balance sheet variable sets and all income statement variable sets.

4-1-1 Sample selection: validating the articulation

The initial usable sample, referred to in Table 4-1, is 31,355 firm-year observations, which includes all firm-year panels that:

- are between the years 1995 and 2009, and

- do not include missing values for either the dependent or independent variables.

In the second line in Table 4-1, it can be seen that 2,398 observations are dropped, as they are categorised as observations with nonsensical signs. For instance, companies with original negative sales (*SAL*) are dropped from the usable sample in order to prevent bias in the estimation. By definition, a company with negative *SAL* has no sales for the current year in addition to the sales return from the previous year. A list of the rational signs for the variables is presented in Table 4-2.

The next three lines in Table 4-1 refer to the double entry constraint. As mentioned before, the first step is to show that the sum of each financial statement is zero. For instance, in the balance sheet, all liabilities, equities and assets are set to zero. After imposing the signs onto the variables given their accounting identities (*Dr* or *Cr*), the summation of each financial statement is calculated. A firm-panel is included only if the financial statements sum to zero.

Interestingly, we find that in some cases the summation is not precisely equal to zero (zero dollars), although it is close to zero. The third line in Table 4-1 reflects the accounting identity in Equation (9) in Chapter 3, and implies that the sum of the income statement variables is zero dollars, or very close to zero dollars (between +$10 and -$10). The latter cases are included in the usable sample of 25,368 observations, as they represent rounding problems and may be caused by downloading based on different scales; however, observations for which the sum of their income statement variables is not between +$10 and -$10 are taken out of the usable sample (3,589 observations).

This situation, in which the sum of the variables does not pass the zero sum test, implies that the accounting journals are not balanced, which may have happened because of data-provider errors. Hribar and Collins (2002) make reference to the same issue when they note that various measurement errors may arise due to the computation, of accruals in their case, from transactions recorded under *incomplete* double entry.

The fourth line shows the same situation, but this time for the summation of the balance sheet statement. 124 observations have been found that have to be dropped under these circumstances, as the sum of their balance sheet variables is not between +$10 and -$10. A further 90 observations have also been found whose cash flow statement variables are not in the mentioned range, as the overall identity does not hold. Therefore, the final usable sample after this step comprises 25,154 observations.

Ultimately, it is important to address the issue of rounding, as the condition that the sum of each financial statement has to be precisely equal to zero requires that rounding differences between +$10 and -$10 must be allocated to some other variable. That is, by doing this, the summation of each financial statement will reach the exact value of zero, which is the requirement of the double entry constraint. In the income statement, the balance sheet and the cash flow statement, such differences are allocated to *OOE*, *OLA* and *ECF* respectively.

Table 4-1 Construction of the estimation sample

Initial usable sample	31355	
Net of nonsensical signs	28957	-2398
For those observations where income statement ID holds	25368	-3589
For those observations where balance sheet ID holds	25244	-124
For those observations where the full accounting ID hold	**25154**	-90

Table 4-2 Rational signs for the usable sample

SAL<=0	*PRE*>=0	*PPE*>=0	OCL<=0	OLL<=0
CGS>=0	*OCA*>=0	*OLA*>=0	STD<=0	NER<=0
SAE>=0	*TCA*>=0	*ACP*<=0	APA<=0	MIN<=0
CSI>=0	*LTR*>=0	*ITP*<=0	LTD<=0	PST<=0
ARE>=0	*IUS*>=0	*DPA*<=0	PRC<=0	
INV>=0	*OIN*>=0	*OCL*<=0	DIN<=0	

4-1-2 Sample selection: identifying extreme values

The main concern of this section is to identify the statistically influential values in the US sample that may unnecessarily distort the estimation. In fact, in linear regression, such outliers are often defined as any observation with a large residual. In this thesis, following studies such as Bollen and Jackma (1985), Kim and Kross (2005) and Wu and Xu (2008), we have applied the studentised residual technique in order to detect outliers. By using this technique, an outlier is defined as an observation whose dependent-variable value is unusual given its values on the predictor variables. In other words, studentised residuals are helpful in identifying outliers that do not appear to be consistent with the rest of the data. In the case of studentised residuals, large deviations from the regression line are identified. Since the residuals from a regression will generally not be independently or identically distributed (even if the disturbances in the regression model are), it is advisable to weight the residuals by their standard deviations (this

is what is meant by studentisation). In summary, a studentised residual is the quotient resulting from the division of a residual by an estimate of its standard deviation. In this thesis, any observations whose absolute studentised residuals are two or more in regressions of current operating, investing and financing cash flows on their lagged estimators, or regressions of current operating and non-operating income on their lagged estimators, are identified as outliers.

There was some concern with respect to dropping the outliers that have been identified by using the studentised residuals technique, as there would be a possibility that by dropping those observations, our regression results in both the prediction of future earnings or the prediction of future cash flows would be biased. To resolve this concern about the prediction of future earnings, it was decided to run a profitability analysis to establish if these outliers are significantly different from the rest of the observations with respect to their profitability ratios. If so, they need to be dropped.

In fact, the profitability analysis run in this section is done not only to accept whatever results that the studentised residual technique, as an econometric tool, gives us, but also monitor this issue from an accounting perspective. In doing so, a two-sample T-test has been applied to compare the outliers identified by the studentised residual technique with the rest of the observations. If there are significant differences (at %1 level) between them for the mean of profitability ratios, then including the outliers in the sample leads to a bias in the estimation of earnings, and outliers need to be dropped.

As Table 4-4 shows, there are significant differences between the mean of almost all applied profitability ratios, which signifies that the outliers have to be dropped. We could continue the comparison between outliers and the rest of the observations by comparing other factors like cash flows (operating, investing and financing cash flows); however the result of Table 4-4 convinces us about the necessity of dropping the outliers.

Furthermore, all the ratios that have been applied in the comparison (Table 4-4) are well-known ratios in the profitability analysis literature.

Table 4-3 shows the number of outliers that have been found for each regression in both the prediction of future cash flows and the prediction of future earnings.

Table 4-3 Number of outliers

Prediction of future cash flows	**Outliers**	**Non-outliers**	**Total sample in Table 4-1**
Number of outliers – Operating cash flow (*OPC*)	1443		
Number of outliers – Investing cash flow (*ICF*)	1384		
Number of outliers – Financing cash flow (*FCF*)	307		
Total	**3134**	**22020**	**25154**
Prediction of future earnings	**Outliers**	**Non-outliers**	**Total sample in Table 4-1**
Number of outliers – Operating income (*OPI*)	1505		
Number of outliers - Non-operating income (*NOI*)	1210		
Total	**2715**	**22439**	**25154**

As can be seen, the number of outliers that have been identified using the studentised residual technique for the regressions in the prediction of future cash flows are 1443, 1384 and 307 for *OPC, ICF* and *FCF* respectively. Therefore, the total number of outliers in the prediction of future cash flows is the summation of

these three figures, which is 3134. Of the rest of the observations, 22020 (25154 – 3134) are those that the *outliers* (3134 observations) need to be compared with.

In the same circumstances, 1505 and 1210 observations have been found as the outliers for the *OPI* regression and the *NOI* regression respectively in the prediction of future earnings. Thus, the total number of outliers in this part is 2715. Furthermore, the rest of the observations that are not outliers number 22439 (25154 – 2715).

Table 4-4 Two-sample t-test for samples including and excluding outliers

Variables	Prediction of future cash flows		Prediction of future earnings	
	t-statistic	p-value	t-statistic	p-value
Return on equity	-1.5809	0.114	2.6211	0.009
Financial leverage	0.2655	0.791	3.5789	0.000
Net financial liabilities	18.1262	0.000	11.3843	0.000
Total equity	6.7893	0.000	-11.6889	0.000
Spread	-3.3823	0.001	-7.363	0.000
Return on assets	-3.6414	0.000	-6.6118	0.000
Profit margin	1.8285	0.068	3.4063	0.001
Operating income	-10.4605	0.000	-37.0933	0.000
Sales	-45.5992	0.000	-54.5829	0.000
Asset turnover	0.0403	0.968	0.4046	0.686
Operating assets	-14.3867	0.000	15.0936	0.000
Operating liabilities	10.1811	0.000	9.8102	0.000
Net borrowing costs	0.7312	0.465	3.6424	0.000
Net financial expenses	-13.5622	0.000	-22.3592	0.000
Interest expenses on debt	-9.5991	0.000	-19.9913	0.000
Non-operating interest income	10.6903	0.000	7.6003	0.000
Financial liabilities (obligation)	14.7217	0.000	6.7101	0.000
Short-term debt	11.1004	0.000	16.8634	0.000
Total debt	9.9293	0.000	-4.9422	0.000
Financial assets	-10.4134	0.000	-8.2635	0.000

Another issue that is important, and which encourages the dropping of these outliers, is the *multicollinearity* issue. In fact, including these outliers leads to a

large multicollinearity problem that biases the results. For instance, in the prediction of future earnings, the ratio of $1/VIF$, which measures the mulitcollinearity problem, is 0.0003, 0.0005, 0.0015, 0.0038 and 0.0041 for *CED*, *PPE*, *OIN, LTR* and *IUS* respectively, indicating a significant problem resulting from including these observations in the final usable sample.

Multicollinearity happens when two or more predictor variables in a multiple regression model are highly correlated. In this situation, the coefficient estimates may change erratically in response to small changes in the model or the data. Multivariate regressions, such as those considered here, attempt to estimate the independent effect of each explanatory variable, but the more correlated are the explanatory variables, the more difficult it is to determine the independent explanatory effect, leading to inflated estimated variances, imprecision and higher rates of rejection. In addition, this problem may not give valid results about any individual predictor, or about which predictors are redundant with respect to others.

A variance inflation factor (*VIF*) is a measure of collinearity for the explanatory variable k, and is equal to $1/(1-R_k^2)$, where R_k^2 is obtained from regressing k on the other explanatory variables. It follows that the inverse of the variance inflation factor,$(1-R_k^2)$, indicates the fraction of the variance of k that is independent from the variance of the rest. Chatterjee et al. (2000), Baum (2003) and others recommend a minimum tolerance of $(1-R_k^2)$ >10%, otherwise the estimates are too sensitive (i.e. unstable) to even small changes in the data. For this thesis, we consider all collinear variables with the even more cautious threshold of $(1-R_k^2)$< 25%. In any case, it should be noted that multicollinearity does not

invalidate inference, and still provides unbiased estimates only with larger estimated variances. Ultimately, it has been decided to drop these observations from the usable sample; however, approximately 10% to 15% of observations have decreased as a result. After dropping these outliers the multicollinearity problem has been solved.

Tables 4-5 and 4-6 show the sample selection processes for the both the prediction of future cash flows and the prediction of future earnings after dropping the outliers.

Table 4-5 Sample selection – prediction of future cash flows

Initial usable sample	*31355*	
Nonsensical signs	*28957*	*-2398*
Observations where income statement id does not hold	*25368*	*-3589*
Observations where balance sheet id does not hold	*25244*	*-124*
Observations where cash flow statement id does not hold	*25154*	*-90*
Outliers – Operating cash flow (*OPC*)	**23711**	**-1443**
Outliers – Investing cash flow (*ICF*)	**22327**	**-1384**
Outliers – Financing cash flow (*FCF*)	***22020***	**-307**

Table 4-6 Sample selection – prediction of future earnings

Initial usable sample	*31355*	
Net of nonsensical signs	*28957*	*-2398*
For those observations where income statement id holds	*25368*	*-3589*
For those observations where balance sheet id holds	*25244*	*-124*
For those observations where cash flow statement id holds	*25154*	*-90*
Outliers – Operating income (*OPI*)	**23649**	**-1505**
Outliers – Non-operating income (*NOI*)	***22439***	**-1210**

As can be seen, the usable sample before dropping the outliers is the same in both tables, presented as 25154 observations. However, in the next step, as a result of applying different regressions, the number of outliers changes and the number

of observations for the usable sample for the prediction of future cash flows is finally identified as 22020, while it is 22439 for the prediction of future earnings.

4-1-3 Sample selection: determining influential observations

In Chapter 3, it was explained how the different sets of variables from the different financial statements were employed to predict the future performance of a company. However, the problem at this point is that four of the mentioned variables – *EGL*, *OIA, NER* and *PEI* – are different from the rest of the variables with regard to the number of non-zero values. For instance, in the prediction of future cash flows, there are just 29 observations that have a non-zero value for *EGL*, in comparison to 21991 (22020–29) observations with a zero value for *EGL*. This small number of non-zero observations leads to a bias in the results, and therefore it has been decided to drop any observation that has a non-zero value for each of these four variables. Table 4-7 shows the number of observations that have been dropped from the usable sample.

Table 4-7 Variables with influential non-zero values

Variables	**Prediction of future cash flows**	**Prediction of future earnings**
Observations with non-zero *EGL*	29	36
Observations with non-zero *OIA*	98	106
Observations with non-zero *NER*	179	188
Observations with non-zero *PEI*	16	16
Total	**322**	**346**

As can be seen, the total number of observations that have been dropped from the usable sample in both the prediction of future cash flow and the prediction of future earnings are 322 and 346 respectively, which are not very large numbers.

At the end of this section, Tables 4-8 and 4-9 present the process of reaching the final usable sample for both the predictions.

Table 4-8 Final sample selection: prediction of future cash flows

Initial usable sample	*31355*	
Net of nonsensical signs	*28957*	*-2398*
For those observations where income statement id holds	*25368*	*-3589*
For those observations where balance sheet id holds	*25244*	*-124*
For those observations where cash flow statement id holds	*25154*	*-90*
Number of outliers – operating cash flow (OPC)	*23711*	*-1443*
Number of outliers – investing cash flow (ICF)	*22327*	*-1384*
Number of outliers – financing cash flow (FCF)	*22020*	*-307*
Observations with non-zero *EGL*	**21991**	**-29**
Observations with non-zero *OIA*	**21893**	**-98**
Observations with non-zero *NER*	**21714**	**-179**
Observations with non-zero *PEI*	***21698***	**-16**

As can be seen in Table 4-8, the number of observations in the final usable sample for the prediction of future cash flow is 21698, while this number is 22093 for the prediction of future earnings. As mentioned earlier, the reason for presenting two different usable samples is because of the different regression equations, and consequently the different number of outliers.

Table 4-9 Final Sample selection: prediction of future earnings

Initial usable sample	*31355*	
Net of nonsensical signs	*28957*	*-2398*
For those observations where income statement id holds	*25368*	*-3589*
For those observations where balance sheet id holds	*25244*	*-124*
For those observations where cash flow statement id hold	*25154*	*-90*
Number of outliers – operating income (OPI)	*23649*	*-1505*
Number of outliers – non-operating income (NOI)	*22439*	*-1210*
Observations with non-zero *EGL*	**22403**	**-36**
Observations with non-zero *OIA*	**22297**	**-106**
Observations with non-zero *NER*	**22109**	**-188**
Observations with non-zero *PEI*	***22093***	**-16**

The aim of the next part of this chapter is to present a table for the industry distribution of the selected sample and to show how many companies apply individually for each industry and in total.

4-2 Industry distribution

In the previous section, the final number of observations in the usable sample was shown. The aim of this section is to illustrate the distribution of that selected number of observations, based on the Industry Classification Benchmark (*ICB*).

In 2005, the *FTSE* Group and Dow Jones Indexes created a definitive classification system called the Industry Classification Benchmark (*ICB*), covering over 60,000 companies and 65,000 securities worldwide from the Dow Jones and *FTSE* universes. This coverage makes the database a comprehensive tool for global sector analysis. It includes four classification levels: Industries (10), Super sectors (19), Sectors (41) and Subsectors (114). Table 4-12 shows this classification for

Industry level, as this thesis applies only this level of classification for its empirical work.

Table 4-10 Industry Classification

Industry
(1) Oil & Gas
(2) Basic Materials
(3) Industrials
(4) Consumer Goods
(5) Health Care
(6) Consumer Services
(7) Telecommunications
(8) Utilities
(9) Financials
(10) Technology

One of the aims of this thesis is to control the industry fixed effects as the *exogenous* variables in the prediction of the future performance of the firm, in addition to the purely accounting numbers that are only related to the financial statements and are known as the *endogenous* variables. Therefore, in this section all of these industries will be individually introduced.

Ten different industries were initially utilised in this study, but all the observations from the Financial and Utilities industries were then dropped from the usable sample, as the companies in these two industries operate differently from companies in other industries, and the structure of their financial statements is different. Therefore, all the observations in the usable sample in Tables 4-9 and 4-10 are distributed between the remaining eight industries, listed in Tables 4-11 and 4-12, for both the prediction of future cash flows and the prediction of future earnings.

Table 4-11 Industry distribution - prediction of future cash flows

Industry	Number of observations	Number of firms
Basic Materials	926	223
Consumer Goods	2787	471
Consumer Services	2921	594
Health Care	4149	709
Industrials	5184	916
Oil & Gas	743	225
Technology	4800	836
Telecommunications	188	54
Total	**21698**	**4028**

The first column of Table 4-11 represents the eight different relevant industries based on the *ICB* coding. The second column illustrates the number of utilised observations from each industry, for instance, 926 observations (firm-year observations) from the Basic Materials industry. The total number of distributed observations among the industries is also presented at the end of the table, which in this case is 21698 for the prediction of future cash flows. This is the same as the number displayed at the end of Table 4-8. The third column shows the number of companies from each of the industries, for instance, 223 firms work in the Basic Materials industry, which equals 926 firm-year observations. The total number of companies employed in the prediction of future cash flows is 4028.

In the next chapter, all the empirical work of this thesis will be retested for each industry to show the validity of the results.

Table 4-12 Industry distribution - prediction of future earnings

Industry	Number of observations	Number of firms
Basic Materials	953	224
Consumer Goods	2783	470
Consumer Services	3030	602
Health Care	4080	713
Industrials	5231	918
Oil & Gas	1050	269
Technology	4770	834
Telecommunications	196	56
Total	**22093**	**4086**

The same is also applicable to Table 4-12, as it shows the industry distribution for the usable sample, but this time for the prediction of future earnings. The total number of selected observations is 22093, which is also displayed in Table 4-9. The third column presents the number of firms utilised in the prediction of future earnings, which is 4086 firms.

In the next part of this chapter, descriptive statistics are presented for all applied variables in both the prediction of future cash flows and the prediction of future earnings.

4-3 Descriptive statistics

In this section, Tables 4-13 and 4-14 represent the descriptive statistics for all the employed variables in both the prediction of future cash flows and future earnings.

Regarding Table 4-13, means, standard deviations, minimum values and maximum values are the main columns of the table, providing the relevant information for all employed dependent and independent variables in the prediction of future cash flows.

The number of observations, as already discussed, is 21,698. Regarding the means, the value for operating cash flow (*OCF*) is -18,225,330, with the negative sign indicating the firm's receipts from its operating activities; however, the values for investing cash flow (*ICF*) and financing cash flow (*FCF*) are presented as 8,368,586 and 9,856,744, with the positive signs indicating the firm's payments.

The rest of the variables in the table represent both the balance sheet variables and the income statement variables. All the changes in assets or, in other words, all the 'accruals from assets' and 'changes in cash' are presented with positive signs, indicating that all of these items on average have increased from the previous year to the current year. With regard to the changes in liabilities or, in other words, all 'accruals from liabilities' and 'changes in equities' it is also applicable, as the most of the variables (except ΔDTX) are presented with negative signs, indicating that all of these variables on average have increased from the previous year to the current year.

With regard to the income statement, as mentioned before, all credit items such as sales (*SAL*) have to be displayed with a negative sign. In the table, *SAL* and cost of goods sold (*CGS*) are the largest means, which is reasonable as in general the largest cost in companies is *CGS,* while *SAL* represents a firm's main operating activities.

As mentioned in Table 4-2, a total number of 2398 observations have been dropped from the sample as they are categorised as observations with 'nonsensical signs'. Following this issue, in Table 4-13, the minimum values for both cost of goods sold (*CGS*) and selling and administrative expenses (*SAE*) are presented as zero, because in Table 4-2 it is defined that any observation with negative *CGS* or *SAE* has to be dropped from the sample.

For example, with regard to *CGS,* purchasing raw materials has to be debited into *CGS* and credited into any interrelated account; however, on some occasions, a company may face some unusual circumstances leading to not only stopping buying raw materials for an entire specific year, but also starting to return some of those raw materials to the suppliers. In this unusual situation we decided to drop any observations like this from the sample, and that is why in Table 4-13 the minimum value for *CGS* is zero. In addition, the same situation can be observed for *SAE*, as any observation with negative *SAE* has been dropped from the sample. In fact, there is a possibility that a company, again in unusual circumstances, not only has no selling and administrative expenses for a specific year but may also have a return of the previous year's selling and administrative expenses.

Moreover, there are some variables in Table 4-13 with a minimum value of zero which are originally presented with this value, if we did not introduce a range of accepted values for them in Table 4-2. For instance, in this thesis we do not have any observations in the US sample with negative depreciation, depletion and amortisation (*DDA*), and this is why the minimum value for *DDA* is zero.

Table 4-13 Descriptive statistics - prediction of future cash flows

Variables		Obs.	Mean	Std. Dev.	Min.	Max.
Operating cash flow	*OCF*	21,698	-18,225,330	212,447,297	-6,040,414,308	4,856,309,724
Δ Accounts receivable	Δ*ARE*	21,698	6,705,832	53,610,382	-1,138,685,898	1,914,475,108
Sales	*SAL*	21,698	-440,982,275	1,693,231,651	-42,381,216,619	0
Δ Accounts payable	Δ*APA*	21,698	-3,691,604	35,011,530	-1,194,272,106	895,691,063
Δ Total inventory	Δ*INV*	21,698	4,688,964	44,937,381	-1,958,709,620	1,367,232,038
Δ Accrued payroll	Δ*ACP*	21,698	-1,214,710	12,336,858	-395,629,027	389,069,616
Cost of goods sold	*CGS*	21,698	283,079,123	1,227,785,503	0	34,146,493,723
Selling and administrative expenses	*SAE*	21,698	94,054,625	367,035,526	0	8,189,184,490
Δ Deferred tax	Δ*DTX*	21,698	514,123	44,193,758	-2,488,780,570	2,853,121,233
Δ Income tax payable	Δ*ITP*	21,698	-8,043	26,654,060	-1,714,649,155	1,545,186,954
Δ Provision for risks and charges	Δ*PRC*	21,698	-1,859,376	74,838,679	-8,248,157,402	1,244,986,043
Income taxes	*ITX*	21,698	13,040,134	75,540,799	-1,743,696,748	2,560,133,137
Δ Other current assets	Δ*OCA*	21,698	1,666,744	25,466,629	-709,705,105	1,215,857,986
Δ Other current liabilities	Δ*OCL*	21,698	-4,439,445	38,751,253	-1,364,394,972	886,375,885
Δ Other long-term assets	Δ*OLA*	21,698	9,523,993	117,153,139	-3,342,849,212	3,625,231,991
Δ Other long-term liabilities	Δ*OLL*	21,698	-1,962,689	53,255,675	-3,134,065,212	2,970,339,524
Δ Prepaid expenses	Δ*PRE*	21,698	250,598	17,331,976	-1,015,774,138	1,015,774,138
Δ Deferred income	Δ*DIN*	21,698	-778,769	21,644,452	-2,044,245,907	451,434,671
Other operating expense	*OOE*	21,698	5,052,214	59,715,906	-78,595,027	2,465,705,548
Depreciation, depletion and amortization	*DDA*	21,698	16,203,503	74,712,769	0	2,832,164,902
Discontinued operations	*DOP*	21,698	62,296	2,431,406	-51,743,193	180,602,816
Extraordinary credits	*ECR*	21,698	-532,418	9,425,942	-862,554,004	0
Extraordinary charges	*ECH*	21,698	7,289,418	74,294,408	-2,617,301	5,413,319,423
Equity interest earnings	*EIE*	21,698	-797,968	21,633,153	-1,495,277,806	114,165,167
Other (income) expense – net	*OIE*	21,698	-1,150,501	42,343,348	-4,629,868,201	610,052,946
Investing cash flow	*ICF*	21,698	8,368,586	116,614,863	-3,501,704,485	4,994,309,439
Δ Property plant and equipment	Δ*PPE*	21,698	7,235,843	76,356,843	-2,332,549,750	4,448,186,407
Δ Long-term receivables	ΔLTR	21,698	737,383	26,837,569	-438,903,508	1,740,023,745
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	21,698	623,115	20,952,239	-598,262,896	1,129,171,959
Δ Other investments	Δ*OIN*	21,698	812,612	50,813,329	-2,858,533,097	3,334,010,427
Financing cash flow	*FCF*	21,698	9,856,744	227,442,046	-6,503,246,795	6,196,997,288
Δ Short-term debt	Δ*STD*	21,698	-2,841,324	84,162,942	-6,421,374,408	1,709,966,797
Δ Long-term debt	Δ*LTD*	21,698	-7,240,399	109,770,421	-4,632,756,815	4,531,297,096
Δ Cash and short-term investments	Δ*CSI*	21,698	6,305,886	114,018,811	-4,656,957,815	5,415,990,791
Interest expense on debt	*IED*	21,698	4,987,181	27,546,417	0	997,663,872
Interest income	*IIN*	21,698	-2,112,771	14,240,787	-577,010,450	1,043,297
Δ Minority interest in net assets	Δ*MIN*	21,698	-354,068	23,417,468	-1,662,340,155	933,420,250
Δ Preferred stock	Δ*PST*	21,698	-4,928,680	509,472,973	-61,025,601,890	651,572,564
Δ Shareholders' equity	Δ*SEQ*	21,698	-9,685,539	538,044,214	-8,242,597,874	61,436,890,201
Δ Dividend payable	Δ*DPA*	21,698	-60,447	2,670,035	-105,251,232	159,168,976
Net income (before preferred dividends)	*NI*	21,698	21,479,476	181,655,029	-3,056,164,145	5,046,745,304
Minority interest in earnings	*MIE*	21,698	327,963	7,090,565	-30,760,067	682,634,672

In Table 4-14, the descriptive statistics for all the variables employed in the prediction of future earnings are presented. The first piece of information given in this table is the number of observations, which is 22,093. Moreover, in this table the number of variables is less than the number of variables that was employed in the prediction of future cash flows, as has been explained above. In Table 4-13, the largest variables were *SAL, CGS* and *SAE*; however, in Table 4-14 the largest variables are customer receipts (*REC*) and supplier payments (*PAY*), which are created as the result of the articulation process that used *SAL, CGS* and *SAE* as the components of *REC* and *PAY*.

Furthermore, both the values of operating income (*OPI*) and non-operating income (*NOI*) are presented as -48,751,504 and 24,448,963 respectively. Therefore, *OPI* represents the *income* for the company because of the negative sign; however, *NOI* represents a non-operating *loss* for the company because of the positive sign.

Moreover, all variables from the cash flow statement, such as *PAY, OOC, TXP, ECF, CED* and *NCD*, are presented with positive signs, indicating that these are payments for the company; however, other variables such *REC* and *DIR* are displayed with negative signs, indicating that on average firms had receipts, whether from their customers or from debt issues.

Table 4-14 Descriptive statistics – prediction of future earnings

Variables		Obs	Mean	Std.Dev.	Min	Max
Operating income	*OPI*	22,093	-48,751,504	268,866,098	-8,347,157,110	2,106,245,146
Customer receipts	*REC*	22,093	-464,583,072	1,755,692,675	-41,781,905,311	20,656,224
Δ Accounts receivable	Δ*ARE*	22,093	8,085,952	82,125,960	-1,138,685,898	6,164,645,894
Supplier payments	*PAY*	22,093	397,985,798	1,518,399,107	-169,887,847	39,465,545,825
Δ Accounts payable	Δ*APA*	22,093	-4,032,205	41,925,425	-1,194,272,106	2,010,925,690
Δ Total inventory	Δ*INV*	22,093	5,424,572	45,039,314	-692,174,950	1,460,679,364
Δ Accrued payroll	Δ*ACP*	22,093	-1,496,901	14,358,033	-460,829,815	389,069,616
Other operating cash flow	*OOC*	22,093	20,792,722	144,694,227	-3,920,814,282	6,470,621,240
Δ Other current assets	Δ*OCA*	22,093	2,433,114	39,064,332	-709,705,105	2,568,651,239
Δ Other current liabilities	Δ*OCL*	22,093	-6,114,584	82,186,571	-4,489,545,696	4,986,666,685
Δ Deferred income	Δ*DIN*	22,093	-1,182,714	28,009,871	-2,044,245,907	451,434,671
Δ Prepaid expenses	Δ*PRE*	22,093	309,894	19,653,895	-1,015,774,138	1,145,149,840
Non-operating income	*NOI*	22,093	24,448,963	126,320,139	-3,591,664,860	6,284,124,499
Tax payments and provision settlements	*TXP*	22,093	11,703,677	115,437,021	-6,898,498,582	3,154,617,422
Δ Deferred tax	Δ*DTX*	22,093	-376,069	55,347,551	-2,488,780,570	2,853,121,233
Δ Income tax payable	Δ*ITP*	22,093	43,278	24,034,382	-556,656,418	1,545,186,954
Δ Provision for risks and charges	Δ*PRC*	22,093	-2,508,535	89,127,066	-8,248,157,402	1,244,986,043
Exceptional cash flow	*ECF*	22,093	21,157,259	260,074,983	-4,477,367,805	16,631,255,281
Δ Other long-term assets	Δ*OLA*	22,093	17,297,459	203,425,277	-4,334,574,726	9,891,154,551
Δ Other long-term liabilities	Δ*OLL*	22,093	-1,246,605	103,163,659	-2,822,455,366	10,079,868,099
Capital expenditure (net of asset disposals)	*CED*	22,093	14,620,700	204,281,071	-7,183,407,353	8,477,752,444
Δ Property plant and equipment	Δ*PPE*	22,093	12,271,761	148,949,395	-2,332,549,750	8,468,595,111
Δ Long-term receivables	Δ*LTR*	22,093	-64,597	67,326,546	-4,913,777,687	1,740,023,745
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	22,093	856,615	32,743,560	-1,374,192,662	2,108,850,635
Δ Other investments	Δ*OIN*	22,093	1,556,920	92,485,990	-5,092,426,817	4,560,959,892
Debt issues (net of debt repayments)	*DIR*	22,093	-3,525,790	180,069,074	-6,529,521,826	5,821,199,132
Δ Short-term debt	Δ*STD*	22,093	-2,944,615	89,570,616	-3,052,032,723	2,638,739,967
Δ Long-term debt	Δ*LTD*	22,093	-12,368,499	144,891,413	-6,706,975,786	2,150,000,286
Δ Cash and short-term investments	Δ*CSI*	22,093	8,654,347	141,002,195	-4,656,957,815	6,187,548,819
New capital (net of dividends)	*NCD*	22,093	1,848,706	247,028,036	-11,215,874,806	7,430,992,998
Δ Minority interest in net assets	Δ*MIN*	22,093	-381,911	25,151,279	-1,662,340,155	933,420,250
Δ Preferred stock	Δ*PST*	22,093	-3,975,827	495,013,460	-61,025,601,890	651,572,564
Δ Shareholders' equity	Δ*SEQ*	22,093	-20,179,190	554,783,902	-12,025,680,895	61,436,890,201
Δ Dividend payable	Δ*DPA*	22,093	-61,662	2,948,274	-105,251,232	159,168,976

Before moving to the next chapter, it is important to address the way that the data has been deflated. Consistent with the work of Christodoulou and McLeay (2010), the summation of the absolute values of all variables has been applied as the deflator. This type of deflator also recognises that all flows would originally appear with positive values and, therefore, all variables are now interpreted as percentages of the total volume of all transactions that flow within a year.

Our results are sensitive to the choice of deflator, and it is attempted to apply a deflator which tends to lessen the degree of multicollinearity. Consistent with the work of Christodoulou and McLeay (2010), it is preferred to give bounded support to the distribution by deflating all variables through the summation of the absolute values for all flows that give rise to the equilibrium, and as the result of using this kind of deflator, the lowest level of multicollinearity has been monitored, whether in the prediction of future cash flows or the prediction of future earnings. Therefore, all variables are now interpreted as percentages of the total size of transactions that flow within a year. This means that the variables are transformed into percentages (from -100% to 100%) to the total size of transactions that flow in a period.

In empirical financial analysis research it is common to create common-size statements. For example, dividing all components on the balance sheet by total assets creates a vertical common-size balance sheet, so that the resulting percentages eliminate the size-effect and enable comparative study. However, deflation via total assets does not create proportional measurements for a reconciliation statement, since sales and expenses are often much larger.

Therefore, in order to transform all items taken from a reconciliation statement proportionally, we use a measure of the total size of transactions that flow within a year, i.e. the addition of all variables that define the equilibrium level of financial statements.

The aim of the next chapter of this thesis is to provide the empirical results for both the prediction of future cash flows and the prediction of future earnings by using both the *OLS* and *SUR* regressions, and to compare the predictive ability and the explanatory power of the accruals concerning their level of reliability in the prediction of future earnings.

CHAPTER 5 - RESULTS

CHAPTER 5: Results

The aim of this chapter is to present and interpret the results of the thesis. This chapter is divided into three sections. The first section provides the results of the prediction of future cash flows, the second section is with regard to the results of the prediction of future earnings, and finally, in the third section the reliability of accruals in the prediction of future earnings will be discussed.

5-1 Prediction of future cash flows

The main concern of this section is to provide the results concerning the issue of whether imposing the double entry constraint onto the simultaneous linear equations provides a superior model.

With this in mind, three different statistical models are compared with each other from their different viewpoints, and the comparative results are distributed throughout the models of *OLS* (model 1), unconstrained *SUR* (model 2) and constrained *SUR* (model 3).

To this end, first, all models are compared with regard to presenting *reasonable signs* for their coefficients, e.g. it is expected that an increase in accounts receivables (*ARE*) in the present year would be consistent with an increase in customer receipts (*REC*) in the following year. The results reveal that both models (1) and (2) are not capable of showing *reasonable signs* for the coefficients in the prediction of future cash

flows; however, when the double entry constraint is added to the set of simultaneous equations as an extra equation, the coefficients show a set of *reasonable signs*. In addition to this issue, the summation of the coefficients is also tested using the *Wald linear coefficient test*. The test results reveal that in the *decSUR* model (3), the summation of the coefficients is equal to zero.

Moreover, it is assessed whether the standard errors of predictor variables are lower in the case of a more powerful model (and, consequently, whether the t-statistics are higher). Again, the result confirms the superiority of the *decSUR* model (3).

In addition, one of the contributions of this thesis is to show that financial statement items that incorporate both *cash and accrual* information make a higher contribution to prediction (i.e. they are more significant and larger) than those that are either just cash amounts or accrual amounts. For example, as the sales figure contains both cash and accrual information, sales may be expected to contribute more to cash flow prediction compared to receivables, which is based only on accrual information. In fact, all of the three models are capable of showing the superiority of most of the income statement variables in comparison to the all the accruals, changes in cash and changes in equities; however, as mentioned above, models (1) and (2) are not able of showing the correct *signs* for the coefficients.

Later in this chapter, the industry sensitivity test shows that our findings are also appropriately distributed among the different industries, particularly when checked for the prediction of future earnings.

5-1-1 Coefficient signs in the prediction of future cash flows

The aim of this section is to find the model that appropriately gives signs to the coefficients in the prediction of future cash flows. The *OLS* model (1) is the first model that is tested at this stage, and later its results will be compared with models (2) and (3). In the *OLS* model (1), it is assumed that each of the linear equations describing the cash flow components (operating, investing and financing) is unrelated to the other two. Consequently, the results of each of these linear equations are presented separately in a separate table. Table 5-1 describes the regression results in the prediction of future operating cash flow (*OCF*) using the *OLS* model (1).

In addition to the first and the second columns of Table 5-1 illustrating the variable names and their abbreviations, the third column represents the expected signs for the estimators' coefficients. As can be seen, the expected signs for all variables representing *changes in balance sheets* are negative (-), and the expected signs for all the variables in the *income statement* are positive (+).

Rational signs for balance sheet variables (assets)

In Table 3-5, it was demonstrated that increases and decreases in assets need to be shown with positive and negative signs; however, increases and decreases in cash flow are presented with negative and positive signs. In addition, it is expected that increases in assets in the current year will be associated with increases in future cash flows, and that decreases in current assets will be associated with decreases in future cash flows.

To show the above statement regarding the association between the increase in current assets and the increase in future cash flows, we can write it as:

$$(increase\ in\ cash\ flows)_{t+1} = (increase\ in\ assets)_t$$

In the estimation model, in order to hold the above accounting logic, a 'negative' coefficient is required to be imposed onto the variables representing increases in assets. Furthermore, to show the association between decreases in current assets and decreases in future cash flows, it can be written as below:

$$(decrease\ in\ cash\ flows)_{t+1} = (decrease\ in\ assets)_t$$

However, to hold the above accounting logic, a 'negative' coefficient is required to be imposed onto the variables representing decreases in assets. At the end of this section, it can be concluded that for any variable representing changes in assets (whether increases or decreases), a negative sign as a coefficient is required to be imposed.

Rational signs for balance sheet variables (liabilities)

With respect to liabilities in the balance sheet, Table 3-5 demonstrated that increases and decreases in these variables need to be shown with negative and positive signs. Furthermore, it is expected that increases in liabilities in the current year will be associated with decreases of future cash flows, and decreases in current liabilities will be associated with increases in future cash flows. To show this statement regarding the

association between increases in current liabilities and decreases in future cash flows, it can be written as:

$$(decrese\ in\ cash\ flows)_{t+1} = (increase\ in\ libilities)_t$$

To hold the above accounting logic, a 'negative' coefficient is required to be imposed onto the variables representing increases in liabilities. To show the association between decreases in current liabilities and increases in future cash flows, it can be written as below:

$$(increase\ in\ cash\ flows)_{t+1} = (decrease\ in\ liabilities)_t$$

However, to hold the above accounting logic, a 'negative' coefficient is required to be imposed onto the variables representing decreases in liabilities. Therefore, the rational coefficient sign for both increases and decreases in liabilities is negative, which is consistent with the coefficient sign for variables representing changes in assets.

Rational signs for income statement variables (revenues)

Table 3-5 demonstrated that increases and decreases in revenues need to be shown with negative and positive signs. In addition, it is expected that increases in revenues in the current year will be associated with increases of future cash flows, and decreases in current revenues will be associated with decreases of future cash flows.

To show the above statement regarding the association between increases in current revenues and increases in future cash flows, it can be written as:

$$(increase\ in\ cash\ flows)_{t+1} = (increase\ in\ revenues)_t$$

To hold the above accounting logic, a 'positive' coefficient is required to be imposed onto the variables representing increases in revenues. To show the association between decreases in current revenues and decreases in future cash flows, it can be written it as below:

$$(decrease\ in\ cash\ flows)_{t+1} = (decrease\ in\ revenues)_t$$

To hold the above accounting logic, again a 'positive' coefficient is required to be imposed onto the variables representing decreases in revenues. At the end of this section, it can be concluded that in the model a positive coefficient sign needs to be imposed onto both increases and decreases of revenues.

Rational signs for income statement variables (expenses)

In the end, Table 3-5 demonstrated that increases and decreases in expenses need to be shown with positive and negative signs. In addition, it is expected that increases in expenses in the current year will be associated with decreases of future cash flows, and decreases in current expenses will be associated with increases of future cash flows. To show this statement regarding the association between increases in current expenses and decreases in future cash flows, it can be written as:

$$(decrease\ in\ cash\ flows)_{t+1} = (increase\ in\ expenses)_t$$

To hold the above accounting logic, a 'positive' coefficient is required to be imposed onto the variables representing increases in expenses. To show the association

between decreases in current expenses and increases in future cash flows, it can be written as below:

$$(increase\ in\ cash\ flows)_{t+1} = (decrease\ in\ expenses)_t$$

To hold the above accounting logic, a 'positive' coefficient is required to be imposed onto the variables representing decreases in expenses. It can be concluded that a positive coefficient sign is required to be imposed on all variables from the income statement.

Although all variables (except ΔPRE) are significant in explaining the future OCF, not enough evidence is provided to accept the OLS Model (1) as an accurate model for the estimation of future OCF, because it is not able to show the proper signs for the estimators.

Table 5-1 Regression results for *OLS* Model (1) – operating cash flow

Variables		Expected sign	Actual sign	Coefficients	Standard errors	t-statistic	p-value
Δ Accounts receivable	Δ*ARE*	-	+	0.121	0.015	7.83	0.000
Sales	*SAL*	+	+	0.746	0.005	139.2	0.000
Δ Accounts payable	Δ*APA*	-	+	0.160	0.018	8.76	0.000
Δ Total inventory	Δ*INV*	-	+	0.190	0.020	9.49	0.000
Δ Accrued payroll	Δ*ACP*	-	+	0.228	0.035	6.51	0.000
Cost of goods sold	*CGS*	+	+	0.664	0.008	86.45	0.000
Selling and administrative expenses	*SAE*	+	+	0.664	0.007	89.13	0.000
Δ Income tax payable	Δ*ITP*	-	+	0.158	0.061	2.59	0.010
Δ Provision for risks and charges	Δ*PRC*	-	+	0.341	0.065	5.27	0.000
Δ Deferred tax	Δ*DTX*	-	+	0.332	0.040	8.37	0.000
Income taxes	*ITX*	+	+	0.600	0.030	19.75	0.000
Δ Other current assets	Δ*OCA*	-	+	0.147	0.033	4.49	0.000
Δ Other current liabilities	Δ*OCL*	-	+	0.184	0.014	13.1	0.000
Δ Other long-term assets	Δ*OLA*	-	+	0.084	0.010	8.73	0.000
Δ Other long-term liabilities	Δ*OLL*	-	+	0.100	0.024	4.17	0.000
Δ Prepaid expenses	Δ*PRE*	-	*#N/A*	0.067	0.064	1.05	0.293
Δ Deferred income	Δ*DIN*	-	+	0.221	0.031	7.07	0.000
Other operating expense	*OOE*	+	+	0.629	0.014	45.59	0.000
Discontinued operations	*DOP*	+	+	0.291	0.078	3.73	0.000
Extraordinary credits	*ECR*	+	+	0.337	0.026	12.89	0.000
Extraordinary charges	*ECH*	+	+	0.181	0.016	11.29	0.000
Equity interest earnings	*EIE*	+	+	0.668	0.054	12.35	0.000
Other (income) expenses – net	*OIE*	+	+	0.411	0.023	17.51	0.000
Depreciation, depletion and amortisation	*DDA*	+	+	0.310	0.024	13.12	0.000

Another issue that has been tested here is the multicollinearity issue, as it is one of the concerns of this thesis that the inappropriate distribution of actual signs may have been triggered by the multicollinearity problem. To check this issue, Table 5-2 presents the results of the multicollinearity test. As can be seen, multicollinearity is not a significant concern as most of the variables have a *VIF* ratio of less than 4 and a 1/*VIF* ratio of greater than 0.25; however, some degree of the multicollinearity problem has

been observed for *CGS* and *SAL,* with *VIF* ratios of 11.24 and 9.34 and 1/*VIF* ratios of 0.09 and 0.11 respectively.

Some degree of multicollinearity is unavoidable, especially in accounting models that rely on such highly structured information. Nevertheless, Binkley (1982) notes the unexpected but welcomed result that such moderate multicollinearity between covariates within each equation will in fact enhance the superiority of *SUR* over *OLS*, given the two stages involved in estimation. While this is a paradoxical result for *OLS*, some multicollinearity is advantageous for a system of *SUR* because it is consumed during the first stage of estimation, and only then the 'cleaned' and more efficient estimates will be employed in the second and final stage.

Binkley (1982) investigates this particular point and proves that the greater the intra-equation multicollinearity, the more likely it is to have a considerable gain in efficiency for the entire system of *SUR*. Using the words of Binkley (1982, p.894): "*Consider a variable in a particular equation. If this variable is not highly correlated with the other variables within the equation, then any correlation between this and variables in a second equation will be fully felt in reducing the efficiency of SUR. On the other hand, if there exists high multicollinearity within the equation, it is quite probable that correlation of variables across the equations will merely mirror that already existing and thus have little effect on SUR variance: most of the damage will have already been absorbed by OLS variance*".

Therefore, following the above statements, it is expected that less of a multicollinearity problem will be seen for both the *SUR* model (2) and the *decSUR* model (3).

Table 5-2 Multicollinearity test – operating cash flow

Variables		**VIF**	**1/VIF**
Cost of goods sold	*CGS*	11.24	0.09
Sales	*SAL*	9.34	0.11
Selling and administrative expenses	*SAE*	3.31	0.30
Income taxes	*ITX*	1.81	0.55
Depreciation, depletion and amortisation	*DDA*	1.31	0.76
Other operating expense	*OOE*	1.28	0.78
Extraordinary charges	*ECH*	1.27	0.79
Δ Other long-term assets	Δ*OLA*	1.25	0.80
Δ Accounts receivable	Δ*ARE*	1.24	0.81
Δ Deferred tax	Δ*DTX*	1.23	0.82
Δ Accounts payable	Δ*APA*	1.19	0.84
Δ Other current liabilities	Δ*OCL*	1.12	0.90
Δ Total inventory	Δ*INV*	1.11	0.90
Δ Other current assets	Δ*OCA*	1.08	0.93
Δ Other long-term liabilities	Δ*OLL*	1.06	0.95
Extraordinary credits	*ECR*	1.04	0.96
Δ Prepaid expenses	Δ*PRE*	1.03	0.97
Other (income) expenses – net	*OIE*	1.03	0.98
Δ Provision for risks and charges	Δ*PRC*	1.03	0.98
Δ Accrued payroll	Δ*ACP*	1.02	0.98
Δ Deferred income	Δ*DIN*	1.02	0.98
Δ Income tax payable	Δ*ITP*	1.02	0.98
Equity interest earnings	*EIE*	1.01	0.99
Discontinued operations	*DOP*	1.01	0.99

As the *OLS* model (1) considers the equations separately, the regression result of the prediction of future investing cash flows is also presented separately in Table 5-3. Two of the four employed variables in the prediction of future investing cash flow, Δ*LTR* and Δ*IUS,* are not significant, resulting in two unknown values for the actual signs. Moreover, Δ*PPE* is presented with a positive actual sign, which is inconsistent with the one that is expected. The only variable that is significant and which performs

reasonably with respect to the signs is Δ*OIN*. Overall, at this point, this model is still not able to show a set of accurate signs for the estimators in both the prediction of future operating cash flows and the prediction of future investing cash flows. However, as with the prediction of future operating cash flow, a multicollinearity test was also undertaken to test whether the inappropriate actual signs were the result of the multicollinearity problem. The results revealed in Table 5-4 find no multicollinearity problem among the estimators.

Table 5-3 Regression results for *OLS* Model (1) – investing cash flow

Variables		Expected sign	Actual sign	Coefficients	Standard errors	t-statistic	p-value
Δ Property plant and equipment	Δ*PPE*	-	+	0.240	0.006	41.83	0.000
Δ Long-term receivables	Δ*LTR*	-	*#N/A*	0.021	0.016	1.29	0.198
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	-	*#N/A*	-0.014	0.018	-0.77	0.441
Δ Other investments	Δ*OIN*	-	-	-0.018	0.010	-1.8	0.072

Table 5-4 Multicollinearity test - investing cash flow

Variables		*VIF*	*1/VIF*
Δ Property plant and equipment	Δ*PPE*	1.03	0.97
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	1	1.00
Δ Other investments	Δ*OIN*	1	1.00
Δ Long-term receivables	Δ*LTR*	1	1.00

The final part of the cash flow estimation using the *OLS* model (1) is the prediction of future financing cash flows. Once again, this model is not able to show a set of reasonable signs for the coefficients in prediction of future financing cash flows, as it gives positive *actual* signs for all variables. Although all of the variables are significant in explaining future financing cash flows, the model is not adequate to be chosen as an appropriate model, as it is not capable of providing a set of reasonable signs for the coefficients.

Table 5-5 Regression results for *OLS* Model (1) – financing cash flow

Variables		Expected sign	Actual sign	Coefficients	Standard errors	t-statistic	p-value
Δ Short-term debt	Δ*STD*	-	+	0.108	0.013	8.56	0.000
Δ Long-term debt	Δ*LT*D	-	+	0.124	0.010	12.36	0.000
Δ Cash and short-term investments	Δ*CSI*	-	+	0.118	0.008	15.21	0.000
Interest expenses on debt	*IED*	+	+	0.699	0.024	29.73	0.000
Interest income	*IIN*	+	+	1.761	0.062	28.33	0.000
Δ Minority interest in net assets	Δ*MIN*	-	+	0.304	0.052	5.87	0.000
Δ Preferred stock	Δ*PST*	-	+	0.231	0.015	15.41	0.000
Δ Shareholders' equity	Δ*SEQ*	-	+	0.293	0.006	45.4	0.000
Δ Dividend payable	Δ*DPA*	-	+	0.176	0.091	1.95	0.052
Net income (before preferred dividends)	*NI*	+	+	0.785	0.005	171.78	0.000
Minority interest in earnings	*MIE*	+	+	0.820	0.123	6.69	0.000

As Table 5-6 demonstrates, as before, no trace of the multicollinearity problem has been found. The largest *VIF* ratio value is for Δ*SEQ* with a value of 2.77 and a 1/*VIF* ratio of 0.36.

Table 5-6 Multicollinearity test – financing cash flow

Variables		***VIF***	**1/*VIF***
Δ Shareholders' equity	Δ*SEQ*	2.77	0.36
Δ Cash and short-term investments	Δ*CSI*	2.15	0.47
Net income (before preferred dividends)	*NI*	1.46	0.69
Δ Preferred stock	Δ*PST*	1.19	0.84
Δ Short-term debt	Δ*STD*	1.19	0.84
Δ Long-term debt	Δ*LTD*	1.17	0.85
Interest expenses on debt	*IED*	1.13	0.89
Interest income	*IIN*	1.1	0.91
Δ Minority interest in net assets	Δ*MIN*	1.01	0.99
Minority interest in earnings	*MIE*	1.01	0.99
Δ Dividend payable	Δ*DPA*	1.01	0.99

Binkley notes the unexpected but welcomed result that such moderate multicollinearity between covariates within each equation (see Tables 5-2, 5-4 and 5-6) will in fact enhance the superiority of *SUR* over *OLS*, given the two stages involved in estimation. While this is a paradoxical result for *OLS*, some multicollinearity is advantageous for a system of *SUR* because it is consumed during the first stage of estimation, and only then the 'cleaned' and more efficient estimates will be employed in the second and final stage.

Moreover, Christodoulou and McLeay (2010) note that regarding the system of Seemingly Unrelated Regression (*SUR*), where intra-equation multicollinearity acts in a beneficial way, the statistical literature has not reached a consensus on how exactly multicollinearity affects the estimation of integrated systems of simultaneous equations. Therefore, here it is assumed that both the *SUR* model (2) and the *decSUR* model (3) are associated with a smaller variance inflation factor (*VIF*) compared with the *OLS* model (1).

Coefficient signs in the prediction of future cash flows, using the SUR model (2)

The main concern in this section is to present the regression results using the *SUR* model (2), where the Seemingly Unrelated Regression is applied to the prediction of future cash flows; however, the double entry constraint is not imposed on the estimation. In fact, in the second model, the three linear cash flow equations (OCF, ICF andFCF) are related to one another as a set of simultaneous equations, through their correlated residuals.

As can be seen, Table 5-7 illustrates the regression results in the prediction of future cash flows (OCF, ICF andFCF) using the *SUR* model (2). All the variables employed in the prediction process are significant in explaining the future *OCF, ICF* and *FCF*. In fact, in comparison to the *OLS* model (1), this model performs even better, as in the *OLS* model (1) variables such as ΔPRE, ΔLTR and ΔIUS are not significant in the prediction, while they are significant in the *SUR* model (2). However, with respect to the sign issue, the problem still exists, as this model is also unable to show a set of *reasonable signs* for most of the employed variables in the prediction of future cash flows. For instance, the sign of ΔARE in the prediction of future operating cash flow is still positive, which is incorrect, as mentioned in the previous section.

Table 5-7 Regression results for *SUR* Model (2)

Variables		Expected sign	Actual sign	Coefficients	Standard errors	t-statistic	p-value
Operating Cash Flow (*OCF*)							
Δ Accounts receivable	Δ*ARE*	-	+	0.200	0.008	25.87	0.000
Sales	*SAL*	+	+	0.819	0.004	218.15	0.000
Δ Accounts payable	Δ*APA*	-	+	0.209	0.009	24.14	0.000
Δ Total inventory	Δ*INV*	-	+	0.210	0.010	21.93	0.000
Δ Accrued payroll	Δ*ACP*	-	+	0.207	0.016	13.33	0.000
Cost of goods sold	*CGS*	+	+	0.829	0.004	200.52	0.000
Selling and administrative expenses	*SAE*	+	+	0.837	0.004	216.22	0.000
Δ Income tax payable	Δ*ITP*	-	+	0.187	0.027	6.98	0.000
Δ Provision for risks and charges	Δ*PRC*	-	+	0.192	0.028	6.84	0.000
Δ Deferred tax	Δ*DTX*	-	+	0.237	0.017	13.61	0.000
Income taxes	*ITX*	+	+	0.852	0.013	65.48	0.000
Δ Other current assets	Δ*OCA*	-	+	0.209	0.015	14.29	0.000
Δ Other current liabilities	Δ*OCL*	-	+	0.199	0.007	28.05	0.000
Δ Other long-term assets	Δ*OLA*	-	+	0.203	0.006	36.82	0.000
Δ Other long-term liabilities	Δ*OLL*	-	+	0.186	0.011	17.1	0.000
Δ Prepaid expenses	Δ*PRE*	-	+	0.203	0.028	7.3	0.000
Δ Deferred income	Δ*DIN*	-	+	0.191	0.014	13.6	0.000
Other operating expenses	*OOE*	+	+	0.827	0.006	129.38	0.000
Discontinued operations	*DOP*	+	+	0.789	0.034	23.4	0.000
Extraordinary credits	*ECR*	+	+	0.750	0.012	63.27	0.000
Extraordinary charges	*ECH*	+	+	0.790	0.007	109.7	0.000
Equity interest earnings	*EIE*	+	+	0.800	0.024	33.92	0.000
Other (income) expenses – net	*OIE*	+	+	0.767	0.011	72.26	0.000
Depreciation, depletion and amort.	*DDA*	+	+	0.821	0.010	81.14	0.000
Investing Cash Flow (*ICF*)							
Δ Property plants and equipment	Δ*PPE*	-	+	0.224	0.005	46.78	0.000
Δ Long-term receivables	Δ*LTR*	-	+	0.110	0.012	9.42	0.000
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	-	+	0.097	0.013	7.52	0.000
Δ Other investments	Δ*OIN*	-	+	0.101	0.007	13.7	0.000
Financing Cash Flow (*FCF*)							
Δ Short-term debt	Δ*STD*	-	+	0.188	0.006	30.81	0.000
Δ Long-term debt	Δ*LTD*	-	+	0.191	0.005	35.91	0.000
Δ Cash and short-term investments	Δ*CSI*	-	+	0.189	0.005	39.84	0.000
Interest expense on debt	*IED*	+	+	0.795	0.010	81.06	0.000
Interest income	*IIN*	+	+	0.959	0.024	39.35	0.000
Δ Minority interest in net assets	Δ*MIN*	-	+	0.216	0.020	10.71	0.000
Δ Preferred stock	Δ*PST*	-	+	0.199	0.007	29.18	0.000
Δ Shareholders' equity	Δ*SEQ*	-	+	0.205	0.004	45.71	0.000
Δ Dividend payable	Δ*DPA*	-	+	0.207	0.035	5.93	0.000
Net income (before preferred divs)	*NI*	+	+	0.810	0.004	221.94	0.000
Minority interest in earnings	*MIE*	+	+	0.810	0.047	17.21	0.000

Coefficient signs in the prediction of future cash flows, using the decSUR model (3)

In this section, coefficient signs in the prediction of future cash flows using the *decSUR* model (3) are explained. As mentioned before, in this model, in addition to assuming the equations are a set of simultaneous equations and assuming that the equations' residuals are correlated, all the coefficients are limited to the value of zero as a result of the imposed *double entry constraint*. In Chapter 4 the issue of how this extra equation, representing the double entry constraint, acts as the fourth equation in the presented structural matrix was fully explained.

Table 5-8 illustrates the regression results using the *decSUR* model (3). Consistent with our expectations, this model performs better with respect to presenting the coefficient signs. In the prediction of future operating cash flow, given the reasons demonstrated at the start of this chapter, it is assumed that all variables from the income statement have to be presented with positive coefficient signs and all the coefficients of variables representing the changes in the balance sheets have to be shown with negative signs.

The results indicate that, in the prediction of future *OPC,* all variables except *DOP* are presented with signs consistent with those expected. In addition, all of these variables are significant in explaining the future *OPC* at a significance level of 1%. This situation was not seen either in the *OLS* model (1) or in the *SUR* model (2).

Regarding the prediction of future *ICF,* it also applies, as all coefficient signs presented are consistent with those expected, except the coefficient sign of ΔPPE, which is positive (+), while its expected sign is negative (-).

Given the previously presented example concerning Δ*ARE*, it is assumed that an increase in these sorts of accruals is associated with an increase in cash in the following year, and this association is signalled with a negative sign. However, the positive sign of Δ*PPE* implies that an increase in these assets at the present time is not only not associated with a cash inflow in the following year but is in fact associated with a cash outflow for the following year. Consistent with the work of Barth et al. (2001), it is possible to partly interpret the explanation for this positive sign. Barth et al. assume that a firm makes such investments because they expect to generate higher cash flows than would be generated from the firm's previously existing asset base over multiple future periods (not necessarily in the short-term). Depreciation and amortisation tend to match the cost of the investments to the profits that are going to be generated by them for the corporation. If matching is achieved and the investment earns a positive return, then the cash inflows associated with the investment will exceed its depreciation or amortisation in each period, even if the rate of return is lower than the firm's cost of capital, but for the first year of prediction it seems that the matching is not achieved and the investment cannot earn a positive return, and consequently the cash inflow associated with the investment will not exceed its depreciation or amortisation expenses. Therefore, as the prediction horizon in this study is limited to just one year, it is only possible to monitor the first year's cash flow and the Δ*PPE* associated with a cash outflow for the first year.

For the final regression concerning the prediction of future *FCF*, again, almost all variables are significant except Δ*SEQ*. In fact, this is the only variable that is not significant in this section. The rest of the variables are consistent with their expected signs, except *MIE*.

As can be seen, almost all the coefficient signs for the *decSUR* model (3) are consistent with their expected signs as a result of imposing a set of constraints.

Table 5-8 Regression results for *decSUR* Model (3)

Variables		Expected sign	Actual sign	Coefficients	Standard errors	t-statistic	p-value
Operating Cash Flow (*OCF*)							
Δ Accounts receivable	Δ*ARE*	-	-	-0.087	0.010	-9.07	0.000
Sales	*SAL*	+	+	0.674	0.004	176.22	0.000
Δ Accounts payable	Δ*APA*	-	-	-0.081	0.011	-7.32	0.000
Δ Total inventory	Δ*INV*	-	-	-0.074	0.012	-6.03	0.000
Δ Accrued payroll	Δ*ACP*	-	-	-0.222	0.021	-10.65	0.000
Cost of goods sold	*CGS*	+	+	0.687	0.004	154.14	0.000
Selling and administrative expenses	*SAE*	+	+	0.717	0.004	178.98	0.000
Δ Income tax payable	Δ*ITP*	-	-	-0.771	0.035	-21.72	0.000
Δ Provision for risks and charges	Δ*PRC*	-	-	-0.820	0.037	-21.98	0.000
Δ Deferred tax	Δ*DTX*	-	-	-0.412	0.023	-18.08	0.000
Income taxes	*ITX*	+	+	0.351	0.017	21	0.000
Δ Other current assets	Δ*OCA*	-	-	-0.280	0.019	-14.55	0.000
Δ Other current liabilities	Δ*OCL*	-	-	-0.077	0.009	-8.92	0.000
Δ Other long-term assets	Δ*OLA*	-	-	-0.021	0.006	-3.4	0.001
Δ Other long-term liabilities	Δ*OLL*	-	-	-0.205	0.014	-14.52	0.000
Δ Prepaid expenses	Δ*PRE*	-	-	-0.791	0.037	-21.4	0.000
Δ Deferred income	Δ*DIN*	-	-	-0.216	0.019	-11.56	0.000
Other operating expenses	*OOE*	+	+	0.678	0.008	83.79	0.000
Discontinued operations	*DOP*	+	-	-0.309	0.045	-6.83	0.000
Extraordinary credits	*ECR*	+	+	0.418	0.016	26.42	0.000
Extraordinary charges	*ECH*	+	+	0.621	0.009	66.88	0.000
Equity interest earnings	*EIE*	+	+	0.157	0.032	4.93	0.000
Other (income) expenses – net	*OIE*	+	+	0.511	0.014	35.94	0.000
Depreciation, depletion and amortisation	*DDA*	+	+	0.635	0.013	47.24	0.000
Investing Cash Flow (*ICF*)							
Δ Property plant and equipment	Δ*PPE*	-	+	0.130	0.005	25.17	0.000
Δ Long-term receivables	Δ*LTR*	-	-	-0.102	0.014	-7.55	0.000
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	-	-	-0.153	0.015	-10.12	0.000
Δ Other investments	Δ*OIN*	-	-	-0.060	0.008	-7.15	0.000
Financing Cash Flow (*FCF*)							
Δ Short-term debt	Δ*STD*	-	-	-0.037	0.007	-5.17	0.000
Δ Long-term debt	Δ*LTD*	-	-	-0.012	0.006	-1.99	0.047
Δ Cash and short-term investments	Δ*CSI*	-	-	-0.038	0.005	-7.77	0.000
Interest expense on debt	*IED*	+	+	0.579	0.013	45.18	0.000
Interest income	*IIN*	+	+	0.556	0.033	17.08	0.000
Δ Minority interest in net assets	Δ*MIN*	-	-	-0.296	0.027	-11.05	0.000
Δ Preferred stock	Δ*PST*	-	-	-0.032	0.008	-3.92	0.000
Δ Shareholders' equity	Δ*SEQ*	-	*#N/A*	0.002	0.004	0.43	0.669
Δ Dividend payable	Δ*DPA*	-	-	-1.112	0.045	-24.55	0.000
Net income (before preferred dividends)	*NI*	+	+	0.705	0.004	191.35	0.000
Minority interest in earnings	*MIE*	+	-	-1.213	0.060	-20.14	0.000

In order to compare models (1), (2) and (3) at a glance, Table 5-9 sets out all the models in a single table. As can be seen, the superiority of the *decSUR* model (3) is more observable in comparison to models (1) and (2).

Table 5-9 A comparison between models (1), (2) and (3)

Variables		Expected sign	Actual sign	*OLS* Model (1)	Expected sign	Actual sign	*SUR* Model (2)	Expected sign	Actual sign	*decSUR* Model (3)
Operating Cash Flow (*OCF*)										
Δ Accounts receivable	Δ*ARE*	-	+	0.1206***	-	+	0.1999***	-	-	-0.0866***
Sales	*SAL*	+	+	0.7464***	+	+	0.8192***	+	+	0.6740***
Δ Accounts payable	Δ*APA*	-	+	0.1597***	-	+	0.2089***	-	-	-0.0807***
Δ Total inventory	Δ*INV*	-	+	0.1896***	-	+	0.2098***	-	-	-0.0743***
Δ Accrued payroll	Δ*ACP*	-	+	0.2284***	-	+	0.2074***	-	-	-0.2218***
Cost of goods sold	*CGS*	+	+	0.6638***	+	+	0.8290***	+	+	0.6871***
Selling and administrative expenses	*SAE*	+	+	0.6644***	+	+	0.8369***	+	+	0.7173***
Δ Income tax payable	Δ*ITP*	-	+	0.1584***	-	+	0.1867***	-	-	-0.7710***
Δ Provision for risks and charges	Δ*PRC*	-	+	0.3405***	-	+	0.1923***	-	-	-0.8200***
Δ Deferred tax	Δ*DTX*	-	+	0.3324***	-	+	0.2367***	-	-	-0.4121***
Income taxes	*ITX*	+	+	0.6000***	+	+	0.8518***	+	+	0.3510***
Δ Other current assets	Δ*OCA*	-	+	0.1472***	-	+	0.2093***	-	-	-0.2800***
Δ Other current liabilities	Δ*OCL*	-	+	0.1839***	-	+	0.1989***	-	-	-0.0773***
Δ Other long-term assets	Δ*OLA*	-	+	0.0836***	-	+	0.2027***	-	-	-0.0212***
Δ Other long-term liabilities	Δ*OLL*	-	+	0.0996***	-	+	0.1862***	-	-	-0.2049***
Δ Prepaid expenses	Δ*PRE*	-	*#N/A*	0.0672	-	+	0.2034***	-	-	-0.7907***
Δ Deferred income	Δ*DIN*	-	+	0.2213***	-	+	0.1909***	-	-	-0.2158***
Other operating expenses	*OOE*	+	+	0.6293***	+	+	0.8270***	+	+	0.6779***
Discontinued operations	*DOP*	+	+	0.2912***	+	+	0.7890***	+	-	-0.3090***
Extraordinary credits	*ECR*	+	+	0.3372***	+	+	0.7498***	+	+	0.4180***
Extraordinary charges	*ECH*	+	+	0.1806***	+	+	0.7903***	+	+	0.6213***
Equity interest earnings	*EIE*	+	+	0.6676***	+	+	0.7998***	+	+	0.1569***
Other (income) expenses – net	*OIE*	+	+	0.4111***	+	+	0.7673***	+	+	0.5112***
Depreciation, depletion and amortisation	*DDA*	+	+	0.3104***	+	+	0.8211***	+	+	0.6350***
Investing Cash Flow (*ICF*)										
Δ Property plant and equipment	Δ*PPE*	-	+	0.2404***	-	+	0.2238***	-	+	0.1300***
Δ Long-term receivables	Δ*LTR*	-	*#N/A*	0.0206	-	+	0.1097***	-	-	-0.1025***
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	-	*#N/A*	-0.0137	-	+	0.0973***	-	-	-0.1527***
Δ Other investments	Δ*OIN*	-	-	-0.0176*	-	+	0.1007***	-	-	-0.0600***
Financing Cash Flow (*FCF*)										
Δ Short-term debt	Δ*STD*	-	+	0.1079***	-	+	0.1881***	-	-	-0.0372***
Δ Long-term debt	Δ*LTD*	-	+	0.1242***	-	+	0.1913***	-	-	-0.0120**
Δ Cash and short-term investments	Δ*CSI*	-	+	0.1179***	-	+	0.1891***	-	-	-0.0380***
Interest expense on debt	*IED*	+	+	0.6989***	+	+	0.7948***	+	+	0.5788***
Interest income	*IIN*	+	+	1.7607***	+	+	0.9591***	+	+	0.5557***
Δ Minority interest in net assets	Δ*MIN*	-	+	0.3042***	-	+	0.2162***	-	-	-0.2958***
Δ Preferred stock	Δ*PST*	-	+	0.2307***	-	+	0.1994***	-	-	-0.0324***
Δ Shareholders' equity	Δ*SEQ*	-	+	0.2930***	-	+	0.2050***	-	*#N/A*	0.0019
Δ Dividend payable	Δ*DPA*	-	+	0.1764*	-	+	0.2069***	-	-	-1.1119***
Net income (before preferred dividends)	*NI*	+	+	0.7846***	+	+	0.8102***	+	+	0.7045***
Minority interest in earnings	*MIE*	+	+	0.8204***	+	+	0.8095***	+	-	-1.2127***

However, some questions remain unanswered, such as: what would be the distribution of the coefficient signs when the sample used is limited to a single industry; what would be the difference of the standard errors and t-statistics among the different models; and what would be the difference between the income statement variables and the variables representing changes in the balance sheets between different models. Additionally, it is still necessary to compare the likelihood ratios, the correlation between the residuals and the *Wald linear coefficient test* between the different models.

Coefficient signs in the prediction of future cash flows based on industry distribution

The aim of this section is to retest all the above regressions to see whether the same results can be obtained. Therefore, instead of the running the regressions for the whole sample, i.e. 21,698 firm-year observations, it has been run for eight employed industries.

In this case, the regression results are only presented for models (2) and (3), as the main aim is to highlight the effect of the double entry constraint.

Table 5-10 represents the regression results in the prediction of future cash flows using the *SUR* model (2), when the double entry constraint is not included in the model. For each industry, a column, A*, represents the actual signs which can be compared with column E*, representing the expected signs. As can be seen, the *SUR* model (2) is unable to present actual signs consistent with the expected signs. It is assumed that this occurs because the double entry constraint has not yet been imposed onto the system;

however, in the following section, when the constraint is observed the problem will be solved. The only distributed sign in Table 5-10 as an actual sign is the positive sign (+), which is not correct, given the reasons presented at the start of this chapter. Moreover, most of the variables are insignificant in the prediction process, particularly for industries with a smaller number of observations, such as the Telecommunication industry with just 188 observations.

Table 5-10 Regression results by industry for the *SUR* Mmodel (2)

Variables		E*	A*	Technology	A*	Industrials	A*	Health Care	A*	Basic Materials	A*	Oil & Gas	A*	Consumer Goods	A*	Telecommunications	A*	Consumer Services
Operating Cash Flow (*OCF*)																		
Δ Accounts receivable	*ΔARE*	-	+	0.18***	+	0.14***	+	0.11***	+	0.16***	+	0.22***	+	0.13***	+	0.21*	+	0.19***
Sales	*SAL*	+	+	0.69***	+	0.75***	+	0.78***	+	0.80***	+	0.82***	+	0.76***	+	0.72***	+	0.78***
Δ Accounts payable	*ΔAPA*	-	+	0.19***	+	0.15***	+	0.12***	+	0.20***	+	0.24***	+	0.13***	+	0.22**	+	0.19***
Δ Total inventory	*ΔINV*	-	+	0.20***	+	0.15***	+	0.12***	+	0.19***	+	0.25***	+	0.14***	*N/A*	0.11	+	0.18***
Δ Accrued payroll	*ΔACP*	-	+	0.16***	+	0.15***	+	0.17***	+	0.22***	*N/A*	0.18	+	0.16***	*N/A*	-0.02	+	0.19***
Cost of goods sold	*CGS*	+	+	0.67***	+	0.74***	+	0.75***	+	0.78***	+	0.82***	+	0.75***	+	0.69***	+	0.77***
Selling and administrative expenses	*SAE*	+	+	0.67***	+	0.74***	+	0.74***	+	0.79***	+	0.82***	+	0.75***	+	0.67***	+	0.77***
Δ Income tax payable	*ΔITP*	-	+	0.17***	+	0.14***	*N/A*	0.09	*N/A*	0.20	*N/A*	0.24	+	0.14**	*N/A*	0.19	+	0.16***
Δ Provision for risks and charges	*ΔPRC*	-	+	0.18**	+	0.16***	*N/A*	0.13	+	0.18*	*N/A*	0.26	+	0.15***	*N/A*	0.69	+	0.17**
Δ Deferred tax	*ΔDTX*	-	+	0.18***	+	0.16***	+	0.10**	+	0.20**	+	0.25***	+	0.16***	*N/A*	0.03	+	0.18***
Income taxes	*ITX*	+	+	0.64***	+	0.72***	+	0.71***	+	0.82***	+	0.81***	+	0.75***	+	0.45*	+	0.72***
Δ Other current assets	*ΔOCA*	-	+	0.19***	+	0.14***	+	0.11***	+	0.19*	+	0.27***	+	0.11***	*N/A*	0.22	+	0.16***
Δ Other current liabilities	*ΔOCL*	-	+	0.19***	+	0.15***	+	0.13***	+	0.19***	+	0.21***	+	0.14***	*N/A*	0.13	+	0.18***
Δ Other long-term assets	*ΔOLA*	-	+	0.17***	+	0.13***	+	0.10***	+	0.18***	+	0.21***	+	0.12***	+	0.13*	+	0.17***
Δ Other long-term liabilities	*ΔOLL*	-	+	0.19***	+	0.13***	+	0.10***	+	0.14***	+	0.22***	+	0.11***	*N/A*	0.11	+	0.18***
Δ Prepaid expenses	*ΔPRE*	-	+	0.18***	+	0.18***	*N/A*	0.08	*N/A*	0.14	+	0.25*	+	0.11*	*N/A*	-0.10	+	0.16***
Δ Deferred income	*ΔDIN*	-	+	0.19***	+	0.14***	+	0.13***	+	0.19**	*N/A*	0.14	+	0.13***	*N/A*	0.24	+	0.22***
Other operating expense	*OOE*	+	+	0.67***	+	0.74***	+	0.73***	+	0.78***	+	0.82***	+	0.76***	+	0.66***	+	0.76***
Discontinued operations	*DOP*	+	+	0.56***	+	0.65***	+	0.72***	+	0.68***	+	0.77***	+	0.76***	*N/A*	0.59	+	0.76***
Extraordinary credits	*ECR*	+	+	0.62***	+	0.70***	+	0.73***	+	0.77***	+	0.77***	+	0.71***	+	0.53***	+	0.71***
Extraordinary charges	*ECH*	+	+	0.60***	+	0.67***	+	0.70***	+	0.73***	+	0.77***	+	0.69***	+	0.54***	+	0.72***
Equity interest earnings	*EIE*	+	+	0.62***	+	0.73***	+	0.71***	+	0.80***	+	0.83***	+	0.74***	*N/A*	1.10	+	0.77***
Other (income) expense – net	*OIE*	+	+	0.64***	+	0.71***	+	0.72***	+	0.73***	+	0.76***	+	0.73***	+	0.58***	+	0.73***
Depreciation, depletion and amortization	*DDA*	+	+	0.61***	+	0.70***	+	0.69***	+	0.76***	+	0.76***	+	0.73***	+	0.55***	+	0.76***
Investing Cash Flow (*ICF*)																		
Δ Property plant and equipment	*ΔPPE*	+	+	0.23***	+	0.17***	+	0.19***	+	0.18***	+	0.26***	+	0.14***	+	0.14***	+	0.20***
Δ Long-term receivables	*ΔLTR*	-	+	0.10***	+	0.09***	*N/A*	0.03	*N/A*	0.08	+	0.24***	+	0.10***	*N/A*	0.28	+	0.12***
Δ Investment in unconsolidated subsidiaries	*ΔIUS*	-	+	0.05**	+	0.09***	*N/A*	0.03	+	0.10**	+	0.17***	+	0.12***	*N/A*	-0.08	+	0.07**
Δ Other investments	*ΔOIN*	-	+	0.09***	+	0.08***	*N/A*	0.01	+	0.10***	+	0.17***	+	0.08***	+	0.36*	+	0.12***
Financing Cash Flow (*FCF*)																		
Δ Short-term debt	*ΔSTD*	-	+	0.19***	+	0.15***	+	0.13***	+	0.18***	+	0.23***	+	0.13***	*N/A*	0.14	+	0.18***
Δ Long-term debt	*ΔLTD*	-	+	0.19***	+	0.15***	+	0.13***	+	0.18***	+	0.24***	+	0.13***	+	0.16**	+	0.18***
Δ Cash and short-term investments	*ΔCSI*	-	+	0.18***	+	0.14***	+	0.12***	+	0.17***	+	0.23***	+	0.14***	+	0.18***	+	0.18***
Interest expense on debt	*IED*	+	+	0.65***	+	0.73***	+	0.75***	+	0.80***	+	0.82***	+	0.75***	+	0.73***	+	0.76***
Interest income	*IIN*	+	+	0.75***	+	0.88***	+	0.96***	+	1.08***	+	0.88***	+	0.82***	*N/A*	0.73	+	0.87***
Δ Minority interest in net assets	*ΔMIN*	-	+	0.18***	+	0.16***	+	0.13***	+	0.20**	+	0.29***	+	0.17**	*N/A*	-0.13	+	0.19***
Δ Preferred stock	*ΔPST*	-	+	0.20***	+	0.15***	+	0.13***	+	0.20***	+	0.23***	+	0.13***	+	0.23***	+	0.19***
Δ Shareholders' equity	*ΔSEQ*	-	+	0.20***	+	0.16***	+	0.14***	+	0.20***	+	0.25***	+	0.14***	+	0.22***	+	0.19***
Δ Dividend payable	*ΔDPA*	-	+	0.21***	+	0.13***	*N/A*	0.12	*N/A*	0.31	*N/A*	0.71	*N/A*	0.17	*N/A*	0.77	*N/A*	0.18
Net income (before preferred dividends)	*NI*	+	+	0.70***	+	0.76***	+	0.79***	+	0.81***	+	0.82***	+	0.77***	+	0.74***	+	0.78***
Minority interest in earnings	*MIE*	+	+	0.72***	+	0.91***	+	0.77***	+	0.94***	+	0.88***	+	0.77***	+	1.89*	+	0.78***
Number of the observations			**4800**		**5184**		**4149**		**926**		**743**		**2787**		**188**		**2921**	

Table 5-11 represents the regression results when the *decSUR* model (3) has been applied. As can be seen, the regression results for each industry are almost the same as for when the regression was run on the basis of the whole sample. The strength of the model is greater for the industries with a larger number of observations. For instance, most of the variables' coefficient signs are consistent with the expected signs for the Industrials and Health Care industries, with 5184 and 4149 observations respectively. Although some degree of inconsistency can be monitored in the table, mostly related to those industries with a smaller number of observations (for instance, for the Telecommunication and Oil & Gas industries, with 188 and 743 observations respectively), this is still better than the *SUR* model (2) which was totally incapable of providing actual signs consistent with expected signs.

Table 5-11 Regression results by industry for the *decSUR* model (3)

Variables		E*	A*	Technology	A*	Industrials	A*	Health Care	A*	Basic Materials	A*	Oil & Gas	A*	Consumer Goods	A*	Telecommunications	A*	Consumer Services
Operating Cash Flow (*OCF*)																		
Δ Accounts receivable	Δ*ARE*	-	-	-0.08***	-	-0.08***	-	-0.06***	+	0.03**	-	-0.14***	*N/A*	0.01	+	0.12**	*N/A*	0.10
Sales	*SAL*	+	+	0.60***	+	0.67***	+	0.60***	+	0.58***	+	0.58***	+	0.69***	+	0.77***	+	0.62***
Δ Accounts payable	Δ*APA*	-	-	-0.07***	-	-0.06***	-	-0.05**	+	0.04**	-	-0.09***	*N/A*	0.01	+	0.18***	*N/A*	0.14
Δ Total inventory	Δ*INV*	-	-	-0.06***	-	-0.10***	-	-0.08***	+	0.05**	-	-0.09***	*N/A*	-0.01	*N/A*	0.15	*N/A*	-0.01
Δ Accrued payroll	Δ*ACP*	-	-	-0.17***	*N/A*	-0.03	-	-0.13***	-	-0.07*	-	-0.20***	*N/A*	0.07	*N/A*	-0.08	*N/A*	-0.25
Cost of goods sold	*CGS*	+	+	0.57***	+	0.61***	+	0.57***	+	0.54***	+	0.54***	+	0.67***	+	0.77***	+	0.56***
Selling and administrative expenses	*SAE*	+	+	0.58***	+	0.60***	+	0.57***	+	0.54***	+	0.55***	+	0.69***	+	0.78***	+	0.55***
Δ Income tax payable	Δ*ITP*	-	-	-0.30***	-	-0.66***	-	-0.41***	-	-0.26***	-	-0.49***	-	-0.70***	*N/A*	-0.01	*N/A*	-0.01
Δ Provision for risks and charges	Δ*PRC*	-	-	-0.21***	-	-3.45***	-	-0.66***	-	-0.33***	-	-0.36***	-	-0.24*	*N/A*	-0.15	*N/A*	-0.67
Δ Deferred tax	Δ*DTX*	-	-	-0.19***	-	-0.41***	-	-0.34***	-	-0.17***	-	-0.38***	-	-0.35***	+	0.19*	*N/A*	-0.05
Income taxes	*ITX*	+	+	0.32***	+	0.29***	+	0.14***	+	0.30***	+	0.23***	+	0.37***	+	0.68***	*N/A*	0.19
Δ Other current assets	Δ*OCA*	-	-	-0.14***	-	-0.27***	-	-0.12***	-	-0.10***	-	-0.31***	-	-0.22*	*N/A*	0.13	*N/A*	-0.17
Δ Other current liabilities	Δ*OCL*	-	-	-0.06***	-	-0.04**	-	-0.04***	+	0.04***	-	-0.13***	*N/A*	0.05	+	0.10**	*N/A*	0.01
Δ Other long-term assets	Δ*OLA*	-	-	-0.05***	-	-0.07***	-	-0.05***	+	0.03***	-	-0.12***	*N/A*	0.02	+	0.14***	*N/A*	0.01
Δ Other long-term liabilities	Δ*OLL*	-	-	-0.11***	-	-0.15***	-	-0.16***	*N/A*	-0.01	-	-0.26***	*N/A*	-0.03	+	0.11*	*N/A*	-0.01
Δ Prepaid expenses	Δ*PRE*	-	-	-0.46***	-	-0.50***	-	-0.50***	-	-0.24***	-	-0.52***	-	-0.42***	*N/A*	0.00	-	-1.96**
Δ Deferred income	Δ*DIN*	-	-	-0.35***	-	-0.04*	-	-0.24***	*N/A*	-0.02	-	-0.37***	*N/A*	-0.04	*N/A*	-0.14	*N/A*	0.05
Other operating expense	*OOE*	+	+	0.56***	+	0.58***	+	0.54***	+	0.52***	+	0.55***	+	0.68***	+	0.76***	+	0.55***
Discontinued operations	*DOP*	+	+	0.22***	*N/A*	-0.07	*N/A*	-0.11	-	-0.86***	+	0.15*	+	0.51*	+	0.61***	*N/A*	0.19
Extraordinary credits	*ECR*	+	+	0.46***	+	0.58***	+	0.36***	+	0.45***	+	0.47***	+	0.64***	+	0.67***	+	0.42**
Extraordinary charges	*ECH*	+	+	0.48***	+	0.53***	+	0.46***	+	0.45***	+	0.42***	+	0.60***	+	0.74***	+	0.37***
Equity interest earnings	*EIE*	+	*N/A*	0.09	+	0.37***	+	0.33***	+	0.19***	+	0.29***	+	0.47***	+	0.73***	*N/A*	-0.19
Other (income) expense – net	*OIE*	+	+	0.52***	+	0.58***	+	0.49***	+	0.48***	+	0.55***	+	0.56***	+	0.69***	+	0.50***
Depreciation, depletion and amortization	*DDA*	+	+	0.45***	+	0.48***	+	0.52***	+	0.44***	+	0.39***	+	0.47***	+	0.68***	+	0.32**
Investing Cash Flow (*ICF*)																		
Δ Property plant and equipment	Δ*PPE*	+	+	0.08***	+	0.13***	+	0.09***	+	0.17***	*N/A*	0.01	+	0.11***	+	0.22***	+	0.11***
Δ Long-term receivables	Δ*LTR*	-	*N/A*	-0.03	-	-0.08***	-	-0.07***	*N/A*	-0.01	-	-0.07**	*N/A*	-0.05	+	0.15*	*N/A*	0.26
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	-	-	-0.16***	-	-0.09***	-	-0.22***	-	-0.05**	*N/A*	-0.04	*N/A*	-0.02	+	0.09*	*N/A*	-0.10
Δ Other investments	Δ*OIN*	-	-	-0.04***	-	-0.07***	-	-0.08***	*N/A*	0.02	-	-0.07***	*N/A*	-0.00	+	0.12***	*N/A*	0.32
Financing Cash Flow (*FCF*)																		
Δ Short-term debt	Δ*STD*	-	*N/A*	-0.01	-	-0.03**	-	-0.02*	+	0.08***	-	-0.08***	*N/A*	0.05	+	0.17***	*N/A*	0.04
Δ Long-term debt	Δ*LTD*	-	*N/A*	-0.01	*N/A*	-0.01	*N/A*	-0.01	+	0.07***	-	-0.06***	+	0.05**	+	0.19***	*N/A*	0.10
Δ Cash and short-term investments	Δ*CSI*	-	-	-0.04***	-	-0.03***	-	-0.02*	+	0.05***	-	-0.08***	+	0.05**	+	0.17***	+	0.13*
Interest expense on debt	*IED*	+	+	0.57***	+	0.60***	+	0.50***	+	0.51***	+	0.48***	+	0.73***	+	0.77***	+	0.61***
Interest income	*IIN*	+	+	0.49***	+	0.85***	*N/A*	-0.07	+	0.49***	*N/A*	-0.12	+	0.48**	+	0.76***	*N/A*	-0.43
Δ Minority interest in net assets	Δ*MIN*	-	-	-0.12*	-	-0.13***	-	-0.09*	-	-0.34***	-	-0.66***	*N/A*	0.10	+	0.24***	*N/A*	-0.97
Δ Preferred stock	Δ*PST*	-	*N/A*	-0.01	-	-0.03**	*N/A*	-0.01	+	0.08***	-	-0.16***	+	0.10***	+	0.19***	+	0.21**
Δ Shareholders' equity	Δ*SEQ*	-	*N/A*	-0.01	*N/A*	-0.01	*N/A*	0.01	+	0.09***	-	-0.04***	+	0.08***	+	0.21***	+	0.16***
Δ Dividend payable	Δ*DPA*	-	-	-0.37***	-	-1.03***	-	-1.68***	-	-0.27***	-	-1.25***	-	-6.06***	-	-13.38***	-	-2.29**
Net income (before preferred dividends)	*NI*	+	+	0.65***	+	0.72***	+	0.65***	+	0.61***	+	0.63***	+	0.75***	+	0.80***	+	0.71***
Minority interest in earnings	*MIE*	+	-	-3.53***	*N/A*	-0.11	-	-0.56***	-	-4.19***	+	0.29***	-	-0.93**	+	0.60**	*N/A*	-0.19
Number of the observations				**5184**		**4149**		**2921**		**4800**		**2787**		**926**		**743**		**188**

At the end of this section, it is possible to make the decision that the *decSUR* model (3) is superior to models (1) and (2) with respect to presenting reasonable signs in the prediction of future cash flows. Moreover, in a further industry analysis that limits the number of observations employed to those firms operating in specific industries, almost identical results were obtained. In the following section the summation of the coefficients will be compared between the three models in order to find the superior model.

5-1-2 Testing the zero sum of coefficients using the Wald statistic

In the previous sections, the regression results were compared in order to find the superior model with respect to presenting the rational signs in the prediction of future cash flows. However, the main aim of this section is to provide other supportive evidence in order to show the superiority of the *decSUR* model (3) in comparison to models (1) and (2). The summation of the coefficients is the issue that will be addressed in this section, and will be compared between the three equations. Using the double entry constraint it is assumed that, as all the creative components of *OCF, ICF* and *FCF* are employed in the prediction of future cash flows, the sum of their coefficients will also be set to zero. This issue was explained in full in Chapter 4 using the structural matrixes.

Table 5-12 compares the summation of the coefficients for the *OLS* model (1). As can be seen, the first three rows represent the summation of the coefficients for each of

the *OCF, ICF* and *FCF* regressions. The last row is the total of the three equations' coefficients.

Table 5-12 Summation of coefficients for the *OLS* model (1)

Sum of the coefficients for the *OLS* Model (1) in prediction of future *OCF*	7.8344
Sum of the coefficients for the *OLS* Model (1) in prediction of future *ICF*	0.2297
Sum of the coefficients for the *OLS* Model (1) in prediction of future *FCF*	5.4188
Sum of the coefficients across the three equations in *OLS* Model (1)	**13.4829**

As can be seen, in the absence of the double entry constraint and considering the equations as simultaneous equations, the summation of the coefficients for the *OLS* model (1) is equal to 13.4829, which is far from the value of zero.

Using the *SUR* model (2), the sum of the coefficients is equal to 16.8155, as the double entry constraint has not yet been imposed. Moreover, the *Wald linear coefficient test* is applied to check this value. The result of the test reveals that it is possible to strongly reject the null hypothesis, at a significance level of 1% (p-value = *0.000*), that the sum of the coefficients is set to zero. The value of the *f-statistic* of 11801 also confirms this issue. However, when the double entry constraint is imposed onto the simultaneous equations by using the *decSUR* model (3), the sum of the coefficients will be zero, which is confirmed by the p-value of *1.000* and the *F-statistic* of *0,* which strongly signifies that it is *impossible* to reject the null hypothesis that the sum of the coefficients is set to zero.

Therefore, it can be concluded that the *decSUR* model (3) completely observes the effect of imposing the double entry constraint, as the summation of the variables' coefficients is set to zero.

The next issue that will be examined between the three models is the distribution of the standard errors and t-statistics.

5-1-3 Predictive ability and estimation errors

Giles (2002) and Ozuna and Gomez (1994) note that the gain in asymptotic efficiency associated with *SUR* estimation is reflected in smaller standard errors in comparison with equation-by-equation estimation under *OLS*. An older study by Carlson (1978) suggests that *SUR* would decrease standard errors by up to 20% in comparison with *OLS*. Consequently, in this thesis the aim is to demonstrate the efficiency improvements that would be reflected in the lower standard deviation of the sampling distribution that is associated with each of the estimation methods used in predicting future cash flows. From an accounting perspective, it is expected that, by considering all cash flow equations (i.e. the accounting identities of operating, investing and financing cash flows) at the same time, as simultaneous double entry equations, this will be associated with a smaller standard deviation of the sampling distribution in each equation, regardless of the fact that the residual terms in these linear equations are predictably correlated with each other, in accordance with the underlying accounting relations.

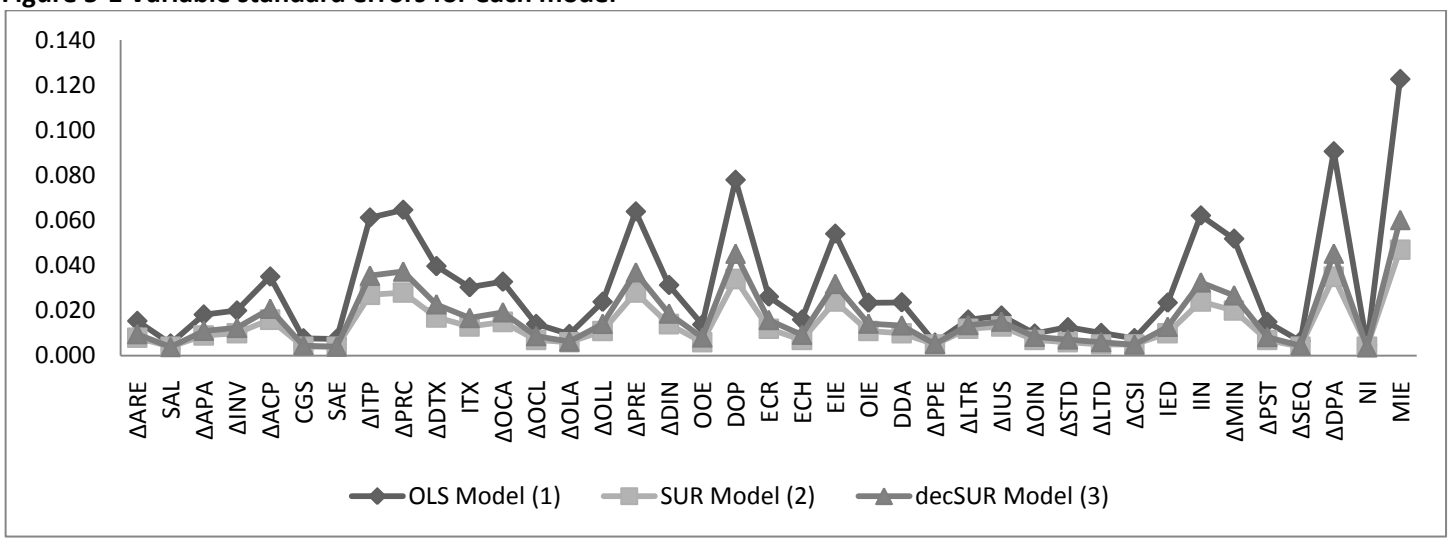
Therefore, the main concern in this section is to provide the evidence that the *decSUR* model (3) is superior to the other models due to presenting smaller values of standard error for predictor variables and consequently larger values of t-statistics.

As Figure 5-1 implies, the *OLS* model (1) is the worst model of all with regard to presenting small values for standard errors. For instance, the standard error for Δ*ITP* in the *OLS* model (1) in Table 5-1 is presented with the value of 0.061, compared to 0.027 and 0.035 for models (2) and (3) in Tables 5-7 and 5-8 respectively.

As can be seen, the index line for the *OLS* model (1) is higher than for the other two models, representing the fact that this model is the worst of the models.

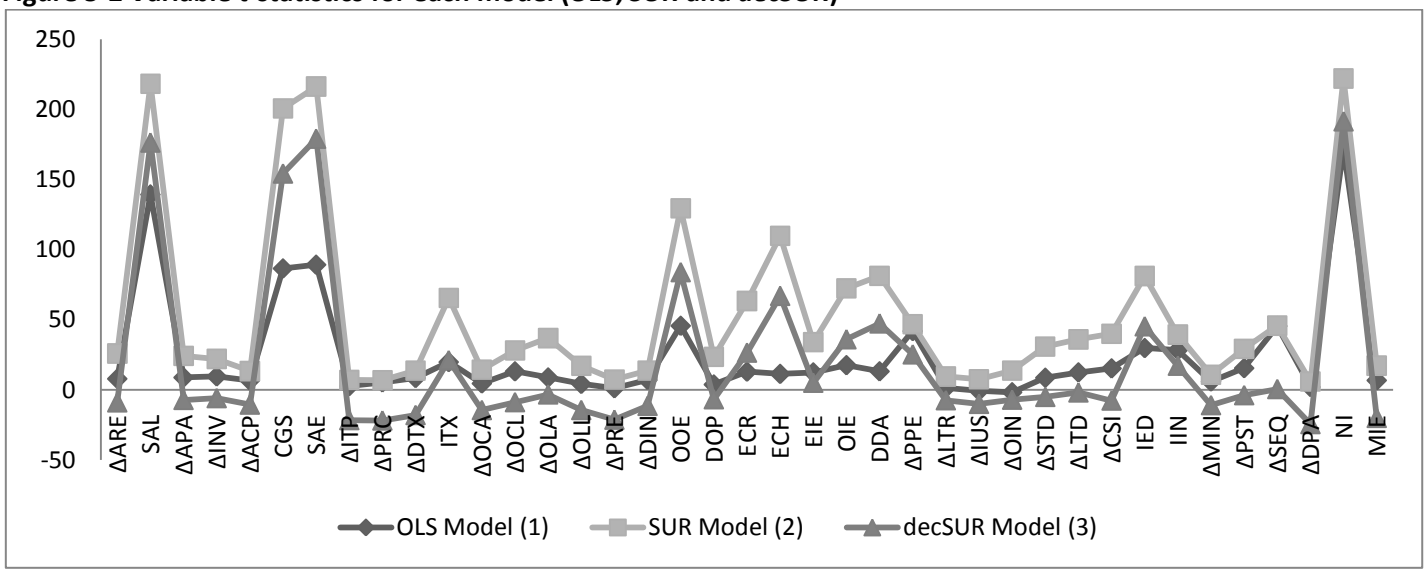
The *SUR* model (2) almost performs better than the *decSUR* model (3), as in the *decSUR* model (3) a set of constraints has been imposed onto the regression which leads to a small increase oin standard errors, but overall the *SUR* system acts better than the *OLS* system.

Figure 5-1 Variable standard errors for each model



In Figure 5-2, t-statistics are compared with each other in order to find the best model. At the first stage, the *OLS* model (1) again performs worse than the others, as it is associated with the smallest numbers for t-statistics. Consistent with Figure 5-1, for the same reason, the *SUR* model (2) performs slightly better than the *decSUR* model (3).

Figure 5-2 Variable t-statistics for each model (*OLS*, *SUR* and *decSUR*)



Overall, both models (2) and (3), applying the *SUR* regression, are superior to the *OLS* model (1). Moreover, the *SUR* model (2) is slightly better than the *decSUR* model (3), and, as mentioned above, it is assumed that this small superiority is due to the imposition of a set of constraints. It is important to address the issue that we cannot choose the *SUR* model (2) as a superior model to the *decSUR* model (3) because of its presentation of slightly smaller values for standard errors and slightly larger values for t-statistics, as previously the failings of the *SUR* model (2) concerning the sign issue were confirmed.

5-1-4 Comparing likelihood ratios

In addition to testing the capability of the models concerning presenting reasonable signs in predictions, as well as presenting a zero value for the summation of the coefficients, and, finally, comparing the standard errors and t-statistics, it is now the turn of the remaining issue, the log of likelihood, which must also be compared between the models in order to identify the model that performs best.

In fact, the further inclusion of the double entry bookkeeping identity as a parameter constraint is expected to ensure estimates that converge with higher precision to their theoretical counterparts.

The first penalised function of the likelihood for model selection is the Akaike information criterion (*AIC*), which is a measure of the relative goodness of fit of a statistical model. This measure offers a relative measure of the information lost when a given model is used to describe reality. It can be said to describe the trade-off between bias and variance in model construction, or, loosely speaking, between accuracy and complexity in a model. Given a dataset, several candidate models may be ranked according to their *AIC* values, and finally the model with the lowest *AIC* is the best model. From the *AIC* values one may also infer that, for example, the top two models are roughly tied, and the rest are far worse, which is exactly the case in this thesis, as both models that apply the *SUR* system are roughly in a tie, and the model that applies *OLS* is far worse with respect to both log-likelihood and *AIC* measures. This issue is shown in Table 5-13.

Here, the second applied penalised function of the likelihood for model selection is the Bayesian information criterion (*BIC*). In statistics, the Bayesian information criterion (*BIC*) is a criterion for model selection among a finite set of models. When fitting models, it is possible to increase the likelihood by adding parameters, but doing so may result in overfitting. The *BIC* resolves this problem by introducing a penalty term for the number of parameters in the model. Therefore, as with the *AIC*, the model with the smallest *BIC* is the best model. The same results are observed in Table 5-13 for *BIC*, as again both SUR models are roughly in a tie, and again *OLS* is far worse as it represents the largest values for *BIC*.[2]

Table 5-13 Log-likelihood, *AIC* and *BIC*

	***OLS* Model 1**			***SUR* Model (2)**	***decSUR* Model (3)**
	OCF	***ICF***	***FCF***		
Log-likelihood	39,358	55,052	36,632	188,094	156,476
AIC – Akaike information criterion	-78,653	-110,079	-73,225	-376,080	-312,846
BIC – Bayesian information criterion	-78,397	-109,984	-73,073	-375,649	-312,423

As mentioned in Section 3-4-2, in addition to the double entry constraint, other types of parameter constraints have been incorporated into the model. For instance, it was explained how to impose cross-equation constraints for recovering same-level coefficients with respect to industry-specific fixed effects. Nevertheless, irrespective of the type of parameter constraint, it should be noted that the imposition of an implied admissible range of parameter space should not alter the fit of the regression given the

[2] *AIC* and *BIC* are superior criteria to the simple *R2*, which entirely ignores the complexity of the model and overfits additive specifications, but also to the *adjusted-R2*, which is not severe enough in its penalisation for lost degrees of freedom.

data. This means that the likelihood recovered by the unconstrained model should not be greatly diminished when moving to the constrained version of the model, and this is the issue that is displayed in Table 5-13 as, by moving from the *SUR* model (2) to the *decSUR* model (3), the log-likelihood regression decreases by 31,618 (188,094 - 156,476) units, which is not a great deduction when compared with the difference between the *SUR* model (2) and the *OLS* model (1).

In fact, at a first glance we may conclude that the *SUR* model (2) is superior to both the *OLS* model (1) and the *decSUR* model (3) as it presents considerably better log-likelihood, and both its penalised functions, *AIC* and *BIC*, are better than the *OLS* model (1), and a bit better than the *decSUR* model (3). Although we can admit the issue of superiority of the *SUR* model (2) to the *OLS* model (1), we cannot choose the *SUR* model (2) as the superior model to the *decSUR* model (3) because of it presenting slightly better log-likelihood, *AIC* and *BIC*, as this model was completely unable to provide a set of rational signs for coefficients in the estimation of cash flows as, well as being unable to provide the requirement of the *Wald linear coefficient test* about the double entry equilibrium, explained in Section 5-1-2.

5-1-5 Breusch-Pagan Lagrange Multiplier test of residual correlation

For the fundamental assumption of cross-equation error correlation, this is empirically verified using the asymptotic *Breusch and Pagan Lagrange Multiplier test.*[3]

Tables 5-14 and 5-15 illustrate the cross-equation error correlations using models (2) and (3) respectively. As can be seen, there is a positive correlation between the *OCF* residuals and the *ICF* residuals; however, the correlation between the *FCF* residuals and both the *OCF* and *ICF* residuals is negative.

Table 5-14 Cross-equation error correlation – *SUR* model (2)

	OCF	*ICF*	*FCF*
OCF	1		
ICF	0.0412	1	
FCF	-0.8401	-0.4208	1

Table 5-15 Cross-equation error correlation – *decSUR* model (3)

	OCF	*ICF*	*FCF*
OCF	1		
ICF	0.0641	1	
FCF	-0.7830	-0.4078	1

Furthermore, the Breusch and Pagan test for both models (2) and (3) verifies the strong contemporaneous dependence between the cross-equation error terms at p-values <0.0001. The above correlations between the equations' residuals validate the use of a system of *SUR* for estimating the joint autoregression in the prediction of future

[3] The Breusch and Pagan test places the $\mathrm{m}(\mathrm{m}-1)/2$ off-diagonal unknown error terms in a single vector and uses the asymptotically efficient estimated residuals to test the null hypothesis of no correlation between the error term in equation *m* and the error term in equation *(m-1)*. This is a Lagrange Multiplier test that is distributed as $\mathrm{x}^2_{\mathrm{m}}(\mathrm{m}-1)/2$.

cash flows. In fact, the *OLS* regression assumes there is no correlation between the equations' residuals; however, the existence of the correlation between the residuals has been verified here.

So far, to identify the superior model, different issues have been compared among the models, such as: (1) rational signs consistent with expectations, not only in the entire sample but also in each industry; (2) testing that double entry holds, based on the Wald test; and (3) testing the model's improvement by way of a higher likelihood and greater precision attached to predictor variables. Moreover, in this section the cross-equation error correlation has been tested by the Breusch-Pagan Lagrange Multiplier test, indicating the importance of the *SUR* system in comparison to the *OLS*. However, in the next section, a further issue will also be tested among the models, comparing whether the financial statement items that incorporate both *cash* and *accrual* information make a higher contribution to prediction (i.e. they are more significant and larger) than those that are either just cash amounts or accrual amounts.

5-1-6 Explanatory power of accruals and earnings

Dechow et al. (1998) built their prediction models on the assumption that sales are more able to forecast future cash flow than current cash flow and accruals, because each sales contract determines both the timing and the amount of the subsequent cash inflows (and often the related cash outflows), and the period in which earnings are recognised. As the sales contract specifies when and under what conditions the customer has to pay, whether in the current period or in the following year, it follows

that the conditions of sale will determine when a future cash inflow is actualised and, in turn, the inclusion of the transaction in earnings along with any associated cash outflows in the form of matched costs. Given that such conditions determine the pattern of cash receipts, the sales contract can be said to be more 'primitive' than both cash receipts and changes in accounts receivable.

Note also that earnings suffer less from these timing and matching problems, as they are associated with both accruals and cash information. Overall, we can say that, between accruals and income statement items, it is income statement items that are expected to have a higher predictive ability for future cash flow, as current cash flow information and accruals information are both embedded in them.

The main concern in this section is to find the model that is able to show the superiority of the *income statement variables* to the variables representing the *changes in balance sheets* by showing a larger significant coefficient, indicating a higher predictive ability and explanatory power.

For the *OLS* model (1), as can be seen in Table 5-1, if we make a comparison between the two related variables, *SAL* and Δ*ARE*, which are in the same line-item, *SAL* is associated with a larger significant coefficient in comparison with Δ*ARE*: the values are 0.746 and 0.121 respectively, indicating that for $1 change in *SAL* in the current year, the future *OCF* will change about 75 cents; however, the future *OCF* will change only about 12 cents when *ARE* changes by $1. Moving to the other line-item variables, the same situation can be observed. For instance, in a comparison between the variables of the second line-item, both *CGS* and *SAE* from the income statement are again presented with larger significant coefficients in comparison with accruals such as Δ*APA,*

Δ*INV* and Δ*ACP*. This condition is the same for almost all line-items. Although the *OLS* model (1) is able to show the superiority of the income statement variables in comparison with accruals, changes in cash and changes in equities, the other two models are also capable of showing this difference, as Tables 5-7 and 5-8 demonstrate.

In fact, this comparison is not that significant in determining the best model, as all the models are capable of differentiating between the income statement variables and the variables representing the changes in the balance sheet.

So far, different issues have been compared among the models in order to find the superior model in the estimation of future cash flows. The first and most important issue that was compared among the models is the model's capability in providing rational coefficient signs consistent with the expected signs. For instance, from an accounting perspective, it is expected that an increase in current accounts receivable (*ARE*) will be seen as cash receipts from customers (*REC*) in the following year.

Neither the *OLS* model (1) or *SUR* model (2) is capable of providing rational signs in prediction of future cash flow, as, for instance, they present the above statement adversely, which is wrong from the accounting perspective. The only model that is capable of providing the rational signs is the *decSUR* model (3). This has also been tested through each single industry and similar results have been found. Therefore, from an accounting viewpoint, any model that is not capable of providing these basic accounting definitions can be rejected; however, we are still interested in monitoring the results of other comparison tools as they highlight more strongly the role of the double entry constraint.

The second issue tests the models' accuracy with respect to the double entry equilibrium issue by using the *Wald linear coefficient test*. The result of this test shows that the summation of all coefficients is set to zero only in the *decSUR* model (3), which has also been confirmed by a very significant *f*-test.

As the next step, all models have been checked in a way that assumes financial statement items that incorporate both cash and accrual information make a higher contribution to prediction (i.e. they are more significant and larger) than those that are either just cash amounts or accrual amounts. The result of this comparison shows that all models are capable of differentiating between these sorts of items, which is also confirmed with significant *f*-tests. Therefore, this comparison tool does not score in our marking.

The next two issues that are compared among the models are the variables' standard errors and log-likelihood (as well as their penalised functions, *AIC* and *BIC*). This shows the superiority of both *SUR* models to the *OLS* model; however, both the *SUR* models, which are the *SUR* model (2) and the *decSUR* model (3), are roughly in a tie, and the *OLS* model (1) is far worse. As is noted in Section 3-4-2, in addition to the double entry constraint, other sorts of constraints have also been included in our models, such as industry-specific fixed effects. Nevertheless, irrespective of the type of parameter constraint, it has to be noted that the imposition of a set of constraints should not alter the fit of the regression given the data, which signify that the likelihood recovered by the *SUR* model (2) should not be greatly decreased by moving to the *decSUR* model (3), and our results show a little improvement for the *SUR* model (2) in comparison to the *decSUR* model (3) with respect to both variables' standard errors and log-likelihood (as

well as their penalised functions, *AIC* and *BIC*). Although we can clearly choose the *SUR* model (2) as the superior model to the *OLS* model (1) because of the very large difference in the mentioned scores (for the result of the comparison, see Sections 5-1-3 and 5-1-4), we cannot accept the superiority of the *SUR* model (2) to the *decSUR* model (3) because of its slightly better results for the mentioned scores, as it was not capable of providing rational signs for coefficients consistent with the basic accounting definitions.

5-2 Prediction of future earnings

As the second step of this chapter, as the *decSUR* model (3) based on the double entry constraint *SUR* was identified as a superior cash flow prediction model, its demonstrated accounting validity should support the use of such a model in the prediction of future *earnings.* In this respect, therefore, the aim of the second part of this thesis is to compare the explanatory power of cash flows and accruals in the simultaneous prediction of future income, which, in this case, is disaggregated into future *operating income* and future *non-operating income.* As noted before, the motivation is to unravel the confusion relating to the relative power of accrual versus cash flow predictors, which emphasise either the relative strength of accrual information (e.g. Chan et al. (2004)) or the greater predictive power of cash flow information (e.g. Arthur et al. (2010) and Sloan (1996)). The results reported in this thesis suggest that, using two simultaneous equations based on the accounting identities for operating and non-operating income, together with double entry constraints on the predictor variables, cash flow items prove to be significantly more powerful in predicting future earnings than accruals and changes in equities.

Moreover, as the superiority of the *decSUR* model (3) to models (1) and (2) has been confirmed in the previous section, the remaining empirical work in the prediction of future earnings is based only on the *decSUR* model (3). Table 5-16 illustrates the results of the prediction of future operating and non-operating incomes.

We start with the results of the prediction of future *OPI* and the first two variables, *REC* and Δ*ARE,* which are in the first line-item. As the main concern of this section is to compare the predictive ability and explanatory power of cash variables versus accrual variables, it is necessary to compare their significant coefficients and, following that, their standard errors and t-statistics. As can be seen in Table 5-16, the significant coefficient of *REC* is 0.885, which is -0.629 for Δ*ARE,* indicating that for $1 change in the *REC, OPI* will change by 88 cents; however, by changing $1 of Δ*ARE, OPI* will change by 63 cents, about 25 cents less than *REC*. Thus, it can be concluded that, of the cash variables, *REC* is more able to explain the future *OPI* that the accrual variable Δ*ARE*.

Not only is the significant coefficient of *REC* larger than the significant coefficient of Δ*ARE,* the standard error of *REC* is about one-third of the standard error of Δ*ARE,* at 0.003 and 0.011 respectively. This implies greater precision in estimation for the cash variable, *REC*, in comparison to the accrual variable, Δ*ARE,* as a result of providing the smaller standard error. The superiority of cash versus accruals is also indicated by the very large t-statistic of 334.67 for *REC,* about six times larger than the t-statistic of Δ*ARE,* which is -57.71.

Moving to the second line-items, the same situation can also be observed. The difference between the second line-items and the first line-items is the number of presented accrual variables, as just one, Δ*ARE,* is presented in the first line-items, while

the second line-items present three: Δ*APA,* Δ*INV* and Δ*ACP.* As before, the cash variable, *PAY,* is more capable of explaining the future *OPI* than all the accrual variables, Δ*APA,* Δ*INV* and Δ*ACP*. As can be seen, the coefficient of *PAY* is 0.91, which is larger than all significant coefficients for accruals, Δ*APA,* Δ*INV* and Δ*ACP,* which are -0.53, -0.61 and -0.23 respectively. Therefore, *PAY* is more able than accruals to explain the future *OPI,* as by changing $1 of *PAY*, the future *OPI* will change by 91 cents, compared to 53, 61 and 23 cents for Δ*APA,* Δ*INV* and Δ*ACP* respectively. Following this clear superiority shown by the larger significant coefficient, the standard error of *PAY* has a value of 0.003, which is smaller than the standard errors of accruals, with values of 0.013, 0.015 and 0.025 for Δ*APA,* Δ*INV* and Δ*ACP* respectively. Finally, both the larger significant coefficient and the smaller standard error for *PAY* are concluded in a very large t-statistic of 318.59, compared to -40.19, -41.15 and -9.5 for Δ*APA,* Δ*INV* and Δ*ACP* respectively.

Almost the same situation can be observed for the third line-items, as the cash variable *OOC* is more able than the accrual variables, Δ*OCA,* Δ*OCA* and Δ*OCA,* in explaining the future *OPI*. The coefficient of *OOC* is 0.76, which is larger than the coefficients Δ*OCA,* Δ*OCA* and Δ*OCA*, which have values of -0.08, -0.47 and -0.31 respectively. However, the size of the coefficient for Δ*PRE,* with a value of 1.15, is larger than the coefficient of *OOC*. In fact, this is the only accrual in the prediction of the future *OPI* that is associated with a larger significant coefficient in comparison to the cash variables. Comparing the standard errors in this section, the absolute superiority of the cash variable, *OOC*, is again clarified, as it is 0.007 in comparison to the accruals, Δ*OCA,* Δ*OCA,* Δ*OCA* and Δ*PRE,* which have values of 0.024, 0.012, 0.024 and 0.045 respectively. Moreover, the only accrual variable, Δ*PRE,* with a coefficient larger than that of the cash

variable, *OOC,* is here associated with the largest standard error of 0.045. Finally, the t-statistic of 102.89 confirms the superiority of *OOC,* compared to -3.28, -38.72, -13.3 and 25.59 for Δ*OCA,* Δ*OCA,* Δ*OCA* and Δ*PRE* respectively.

In conclusion, as a result of employing the *decSUR* model (3), which imposes the double entry constraint onto the *SUR* system and considers all equations as simultaneous equations, our findings solve the current confusion relating to the relative power of accrual versus cash flow predictors, as all cash variables, *REC, PAY* and *OOC,* are associated with the largest significant coefficients among almost all the predictor variables in the forecast of *OPI,* with values of 0.88, 0.91 and 0.76 respectively. Moreover, they are also associated with the smallest values of standard errors among all the predictor variables, which are 0.003, 0.003 and 0.007 respectively, and, finally, with the largest t-statistics of 334.67, 318.59 and 102.89 respectively.

The next simultaneous regression in this section relates to the prediction of future non-operating income, *NOI*.

Table 5-16 Prediction of future earnings

Variables		Coefficients	Standard errors	t-statistic	p-value
Operating Income					
Customer receipts	*REC*	0.885	0.003	334.67	0.000
Δ Accounts receivable	Δ*ARE*	-0.629	0.011	-57.71	0.000
Supplier payments	*PAY*	0.911	0.003	318.59	0.000
Δ Accounts payable	Δ*APA*	-0.528	0.013	-40.19	0.000
Δ Total inventory	Δ*INV*	-0.611	0.015	-41.15	0.000
Δ Accrued payroll	Δ*ACP*	-0.235	0.025	-9.5	0.000
Other operating cash flow	*OOC*	0.761	0.007	102.89	0.000
Δ Other current assets	Δ*OCA*	-0.078	0.024	-3.28	0.001
Δ Other current liabilities	Δ*OCA*	-0.466	0.012	-38.72	0.000
Δ Deferred income	Δ*OCA*	-0.313	0.024	-13.3	0.000
Δ Prepaid expenses	Δ*PRE*	1.150	0.045	25.59	0.000
Non-operating Income					
Exceptional cash flow	*ECF*	0.099	0.006	16.08	0.000
Δ Other long-term assets	Δ*OLA*	0.130	0.008	15.61	0.000
Δ Other long-term liabilities	Δ*OLL*	0.167	0.013	12.52	0.000
Tax payments and provision settlements	*TXP*	0.306	0.013	22.87	0.000
Δ Deferred tax	Δ*DTX*	0.238	0.020	11.88	0.000
Δ Income tax payable	Δ*ITP*	0.274	0.030	9.28	0.000
Δ Provision for risks and charges	Δ*PRC*	0.556	0.034	16.21	0.000
Capital expenditure (net of asset disposals)	*CED*	-0.392	0.015	-25.83	0.000
Δ Property plant and equipment	Δ*PPE*	0.627	0.015	42.55	0.000
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	0.697	0.021	33.17	0.000
Δ Other investments	Δ*OIN*	0.597	0.016	37.21	0.000
Debt issues (net of debt repayments)	*DIR*	0.398	0.010	38.64	0.000
Δ Short-term debt	Δ*STD*	-0.295	0.011	-25.64	0.000
Δ Long-term debt	Δ*LTD*	-0.277	0.011	-24.5	0.000
Δ Cash and short-term investments	Δ*CSI*	-0.258	0.011	-23.55	0.000
New capital (net of dividends)	*NCD*	-0.014	0.002	-6.03	0.000
Δ Minority interest in net assets	Δ*MIN*	0.542	0.024	22.84	0.000
Δ Preferred stock	Δ*PST*	0.189	0.008	24.32	0.000
Δ Shareholders' equity	Δ*SEQ*	0.150	0.004	33.95	0.000
Δ Dividend payable	Δ*DPA*	1.330	0.043	31	0.000

In the prediction of future *NOI*, the superiority of the cash variables to the accruals cannot be observed as was seen in the prediction of future *OPI*. For instance, in the fourth line-items the coefficient of the cash variable *ECF* is 0.10, compared to 0.13 and 0.17 for Δ*OLA* and Δ*OLL* respectively, indicating that for $1 change in *ECF* the future *NOI* will change by 10 cents; however these changes in Δ*OLA* and Δ*OLL* are about 13 and 17

cents for each $1 change in *ECF*. However, the standard error of *ECF,* which is 0.006, is still smaller than the standard errors for both Δ*OLA* and Δ*OLL,* with values of 0.008 and 0.013 respectively. This is also reflected in the larger t-statistic for the cash variables: the t-statistic for *ECF* is 16.08 in comparison with Δ*OLA* and Δ*OLL,* with values of 15.61 and 12.52 respectively. As can be seen, making a decision to choose the cash variable, *ECF*, as the more powerful variable in the prediction of future *NOI* in comparison with accrual variables is difficult, as, on one hand, it is associated with the smallest significant coefficient, and, on the other, it is associated with the smallest standard error and consequently the largest t-statistic. It is therefore evident that the superiority of the cash variable here is not as clear as the superiority of the cash variable in the prediction of future *OPI*.

With respect to the fifth line-items, the cash variable, *TXP,* is associated with almost the largest significant coefficient with a value of 0.306; however, the coefficients for accruals, Δ*DTX,* Δ*ITP* and Δ*PRC,* are 0.24, 0.27 and 0.56 respectively. Therefore, it can be concluded that in this line-item, *TXP* is more able to explain the future *NOI* than all accruals except Δ*PRC*. Moreover, the cash variable, *TXP*, as with all cash variables, is associated with the smallest standard error, which is 0.013; meanwhile, for Δ*DTX,* Δ*ITP* and Δ*PRC* the standard errors are 0.020, 0.030 and 0.034 respectively. Drawing a comparison between the t-statistics of the variables in this line-item highlights the superiority of the *TXP* with a value of 22.87, compared to Δ*DTX,* Δ*ITP* and Δ*PRC* with values of 11.88, 9.28 and 16.24 respectively.

Of all the variables up to this point, the cash variables of *REC, PAY, OOC, ECF* and *TXP* are associated with the smallest standard errors and the largest t-statistics; however some inconsistencies have been observed with respect to the coefficient sizes.

Conversely, among the variables in the sixth line-items, the cash variable, *CED*, is associated with the smallest coefficient of -0.392, in comparison to the related accruals, Δ*PPE,* Δ*IUS* and Δ*OIN,* which have coefficients of 0.63, 0.70 and 0.60 respectively. However, the standard error of the *CED* is still around the smallest at 0.015, in comparison to accruals with standard values of 0.015, 0.021 and 0.016 for Δ*PPE*, Δ*IUS* and Δ*OIN* respectively.

In the seventh line-items, three sorts of variables are compared: cash flow (*DIR*), accruals (Δ*STD* and Δ*LTD*) and changes in cash (Δ*CSI*). As can be seen, *DIR* is associated with the largest significant coefficient, which is 0.40 in comparison to Δ*STD,* Δ*LTD* and Δ*CSI,* with values of -0.29, -0.28 and -0.26 respectively. In addition to presenting the largest values of coefficients among all the variables, the smallest value of standard errors is also allocated to *DIR* at 0.010; however, Δ*STD,* Δ*LTD* and Δ*CSI* all have a value of 0.011. The t-statistic for *DIR* is also 38.64, which is the largest in comparison to those for Δ*STD,* Δ*LTD* and Δ*CSI,* with values of -25.64, -24.5 and -23.5 respectively.

The last line-items compare the cash variable, *NCD,* with changes in equities, Δ*MIN,* Δ*PST,* Δ*SEQ* and Δ*DPA*. As can be seen, all the variables representing changes in equities are more able to explain the future *NOI* as they are associated with larger significant coefficient sizes. For instance, the coefficient for *NCD* is -0.01, compared to 0.54, 0.19, 0.15 and 1.33 for Δ*MIN,* Δ*PST,* Δ*SEQ* and Δ*DPA* respectively. However, as with the other cash variables, *NCD* presents the smallest standard error with a value of 0.002,

in comparison to *changes in equities* with values of 0.024, 0.008, 0.004 and 0.043 respectively. Finally, the t-statistic for *NCD* is the smallest of all with a value of -6.03, compared to 22.84, 24.32, 33.95 and 31 for Δ*MIN,* Δ*PST,* Δ*SEQ* and Δ*DPA* respectively.

In conclusion, for the prediction of future *NOI*, a transparent and clear superiority of the cash flow variables when compared to accruals cannot be observed; however, both *TXP* and *DIR* were partly able to show the superiority of the cash flow variables compared to accruals by showing larger significant coefficients, smaller standard errors and consequently larger t-statistics. Moreover, as with the prediction of future *OPI*, all the cash flow variables are associated with the smallest standard errors in the prediction process. In addition, in a comparison between the cash variables, *NCD* and changes in equities, it can be seen that all the variables representing changes in equities are more able to explain future *NOI* as a result of providing larger significant coefficients.

Overall, for the prediction of future earnings, the findings indicate the superiority of the cash flow variables *REC, PAY, OOC, TXP* and *DIR* to all the accruals, as they provide almost the largest significant coefficients in comparison to accruals, thus representing a higher predictive ability and explanatory power for cash variables in comparison to accruals. Moreover, all cash flow variables are associated with the smallest standard errors, whether in comparison with accruals or in comparison with changes in cash and changes in equities.

In addition to the issue that almost all accrual variables in Table 5-16 are presented with larger significant coefficients and smaller standard errors in comparison to the cash flow variables, representing higher explanatory power and predictive ability,

a *Wald linear coefficient test* is also applied to our empirical work in order to compare these two sorts of variables statistically. The results indicate that we can reject the hypothesis at 1% level that most of the accrual variables are the same as the cash flow variables in the prediction of future earnings. The difference between these two sorts of variables is also confirmed in most of the cases with a very large *f*-statistic. We have not reported the results here as it would be a very large table, given that every single accrual variable needs to be compared with every cash variable.

5-2-1 Industry analysis

In order to be sure about the superiority of the cash variables to the accrual variables, all the regressions have been run again within each industry to see whether the same results will be obtained. As can be seen, the results are robust even within those industries with a very small number of observations, such as Telecommunication, with only 196 observations. This indicates the superiority of cash variables to accrual variables.

For instance, in the prediction of future *OPI*, in a comparison between cash and accruals in the first line-items, *REC* is always larger than ΔARE through all industries. The same situation also can be observed for the second line-items, as *PAY* is always larger than all accruals.

Some inconsistency can be seen in the third line-items, but overall it can be concluded that all the cash flow variables are superior to the accrual variables in

explaining the future cash flows, which has been shown by the larger significant coefficient sizes.

In the prediction of the future *NOI*, of all variables of the fourth line-items, *ECF* is associated with the largest significant coefficient for Health Care, Basic Materials, Oil and Gas, Consumer Goods and Industrials. In the fifth line-items, again, some inconsistencies are observed, as it is not possible to highlight the superiority of the cash variable to accruals.

As before, within the variables from the sixth line-items, accruals are associated with larger coefficients than the cash variable. Moving to the seventh line-items, *DIR* is always associated with a larger significant coefficient than the accruals and changes in cash, and finally, as mentioned before, the coefficients for changes in equities are larger than the coefficient of the cash variable, *NCD*.

Table 5-17 Prediction of future earnings, by industry

Variables		Health Care	Basic Materials	Oil & Gas	Consumer Goods	Technology	Industrials	Telecommunications	Consumer Services
Operating Income									
Customer receipts	*REC*	0.83***	0.86***	0.83***	0.86***	0.81***	0.83***	0.84***	0.84***
Δ Accounts receivable	Δ*ARE*	-0.73***	-0.85***	-0.73***	-0.63***	-0.68***	-0.68***	-0.62***	-0.57***
Supplier payments	*PAY*	0.70***	0.79***	0.79***	0.81***	0.74***	0.76***	0.83***	0.78***
Δ Accounts payable	Δ*APA*	-0.39***	-0.50***	-0.43***	-0.53***	-0.44***	-0.48***	-0.62***	-0.41***
Δ Total inventory	Δ*INV*	-0.59***	-0.60***	-0.31***	-0.65***	-0.44***	-0.60***	-0.18	-0.45***
Δ Accrued payroll	Δ*ACP*	-0.22***	-0.37***	-0.12	-0.13**	0.10	-0.23***	-0.39	-0.23***
Other operating cash flow	*OOC*	0.49***	0.60***	0.61***	0.65***	0.35***	0.66***	0.80***	0.69***
Δ Other current assets	Δ*OCA*	-0.06	-0.17	-0.35***	-0.26***	0.24***	-0.29***	-0.25	-0.22***
Δ Other current liabilities	Δ*OCL*	-0.16***	-0.37***	-0.35***	-0.27***	-0.10***	-0.37***	-0.64***	-0.34***
Δ Deferred income	Δ*DIN*	-0.29***	-0.22*	0.74***	0.80***	0.33***	0.29***	-0.01	-0.10
Δ Prepaid expenses	Δ*PRE*	0.80***	0.45**	0.79***	0.45***	0.84***	0.96***	0.52	0.96***
Non-operating Income									
Exceptional cash flow	*ECF*	0.15***	0.13***	0.12***	0.10***	0.08***	0.12***	0.02	0.10***
Δ Other long-term assets	Δ*OLA*	-0.01	-0.01	0.04	0.04*	0.02*	0.07***	0.10	0.16***
Δ Other long-term liabilities	Δ*OLL*	0.03	0.03	0.02	0.07**	0.11***	0.09***	0.06	0.24***
Tax payments and provision settlements	*TXP*	0.38***	0.22***	0.23***	0.42***	0.29***	0.29***	0.48***	0.21***
Δ Deferred tax	Δ*DTX*	0.05	0.23**	-0.09	0.05	0.05	0.09***	-0.11	0.21***
Δ Income tax payable	Δ*ITP*	0.24***	0.61***	0.33**	0.21**	0.09	0.07	-0.33	0.11**
Δ Provision for risks and charges	Δ*PRC*	2.94***	0.12	0.65***	-0.14**	0.41***	0.08*	1.58**	1.19***
Capital expenditure (net of asset disposals)	*CED*	-0.39***	-0.30***	-0.17***	-0.41***	-0.35***	-0.81***	-0.38*	-0.09***
Δ Property plant and equipment	Δ*PPE*	0.53***	0.43***	0.29***	0.55***	0.56***	1.00***	0.56**	0.30***
Δ Long-term receivables	Δ*LTR*	0.58***	0.42***	0.33***	0.58***	0.62***	1.03***	1.04***	0.00
Δ Investment in unconsolidated subsidiaries	Δ*IUS*	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.37***
Δ Other investments	Δ*OIN*	0.47***	0.37***	0.29***	0.58***	0.46***	0.95***	0.00	0.34***
Debt issues (net of debt repayments)	*DIR*	0.58***	0.50***	0.43***	0.37***	0.42***	0.46***	0.48***	0.39***
Δ Short-term debt	Δ*STD*	-0.54***	-0.48***	-0.39***	-0.33***	-0.39***	-0.39***	-0.37***	-0.22***
Δ Long-term debt	Δ*LTD*	-0.53***	-0.42***	-0.36***	-0.31***	-0.39***	-0.35***	-0.43***	-0.22***
Δ Cash and short-term investments	Δ*CSI*	-0.51***	-0.44***	-0.36***	-0.28***	-0.34***	-0.35***	-0.42***	-0.22***
New capital (net of dividends)	*NCD*	0.02***	0.02***	-0.02***	-0.01	-0.02***	-0.01	-0.01	0.01
Δ Minority interest in net assets	Δ*MIN*	0.22***	0.34***	0.07	0.88***	0.54***	0.63***	0.80	0.61***
Δ Preferred stock	Δ*PST*	0.07***	0.07***	0.13***	0.09***	0.11***	0.16***	0.16**	0.22***
Δ Shareholders' equity	Δ*SEQ*	0.05***	0.05***	0.08***	0.08***	0.09***	0.11***	0.11**	0.16***
Δ Dividend payable	Δ*DPA*	0.85***	4.16***	2.56***	1.97***	0.51***	0.50***	2.20**	1.04***
Number of observations		**4080**	**953**	**1050**	**2783**	**4770**	**5231**	**196**	**3030**

Therefore, as can be seen, even after restricting the number of observations to those within each industry, the same results can still be observed, even in very small industries.

The aim of the following section of this study is to draw a link between the reliability level of accruals and their relevance.

5-3 Reliability of accruals

As the final step of this chapter, a qualitative assessment is carried out of *low and high reliability* accruals in order to assess their predictive ability with respect to future earnings. As mentioned previously, prior work by Lev and Sougiannis (1996) and Richardson et al. (2005) suggests that accruals that are less *reliable* in terms of the greater managerial discretion they afford are expected to be more *relevant*, and accruals that are more *reliable* in terms of the less managerial discretion that they afford are expected to be less *relevant*. However, in contradiction to the work of Richardson et al. (2005), the results show that by using the modelling framework that was built up in the first and second sections of this thesis, the final findings provide evidence that low reliability accruals make a stronger contribution to explaining future income, and high reliability accruals make a less powerful contribution to explaining future income, which offers further insights into the trade-off between the reliability and relevance of accruals.

In fact, the main aim in this section is to draw a natural link between the terminology of *subjectivity* and the well-known accounting concept of *reliability*. Therefore, in this section of the study, reliability is directly linked to empirically observable properties of accounting numbers. It is important to address the fact that reliability is always at the side of relevance, and these are the most important qualities

that make information useful for decision making.[4] Although a great deal of research has been done with regard to the relevance of accounting numbers, there has been relatively little research about reliability.

Lev and Sougiannis (1996) note that if one item is highly relevant, it is to be associated with low reliability. However, Watts (2003) in his work notes that allowing less verifiable and hence less reliable estimates into accounting numbers can seriously compromise their usefulness. In fact, this part of the thesis highlights the crucial trade-off between relevance and reliability.

Regarding the prediction of future *OPI*, there are two accruals that are known to have a low level of reliability: Δ*ARE* and Δ*ACP.* Interestingly, both of them are associated with very large significant coefficients, which are -0.63 and -0.61 respectively. They are almost the largest coefficients of all predictor variables in the forecast of future *OPI (*except in comparison with Δ*PRE).* As mentioned before, consistent with the *Statement of Financial Accounting Concepts 2* of FASB, this thesis poses that an accrual with less reliability is associated with high relevance, which means that a more relevant variable is more capable of 'making a difference' in a decision making, and consequently provides more help to users to form predictions. Clearly this issue, discussed in paragraph 55, shows that information with high relevance is associated with higher 'predictive ability'.

Consequently, the higher predictive ability for less reliable accruals is shown with the larger significant coefficients in comparison with the more reliable accruals and this difference is statistically confirmed by the *Wald linear coefficient test* at the significant

[4] For example, see Fig. 1 in Statement of Financial Accounting Concepts 2,''Qualitative Characteristics of Accounting Information'' (Financial Accounting Standards Board, 2002).

level of 1%. The lower level of reliability can be translated into a higher relevance, which, given to FASB 2, is associated with higher predictive ability, and consequently a larger significant coefficient is presented, indicating the higher explanatory power of these accruals.

Moreover, there are three accruals with a high level of reliability in the prediction of future *OPI,* which are Δ*APA,* Δ*ACP* and Δ*OCL.* It was expected that all these accruals with a high level of reliability will be associated with a low level of relevance and a lower level of predictive ability, which can be concluded by observing the smaller significant coefficients, indicating the weaker explanatory power for these accruals.

This clear difference between accruals with a low level of reliability and accruals with a high level of reliability in the prediction of future *OPI* has also been confirmed by the very high *f-statistics* at the significant level of 1%.

In the prediction of future *NOI* this can also be observed. Δ*PRC* and Δ*PPE* are the predictor variables with a high level of reliability, and, as can be seen, both of these are associated with a higher level of relevance and a higher level of predictive ability, which can be concluded from the larger significant coefficients for both of them in comparison to accruals with a high level of reliability. This implies a higher explanatory power for these accruals.

With respect to the accruals with a high level of reliability, the same situation can be observed, as all these three accruals, Δ*ITP,* Δ*STD* and Δ*LTD,* are associated with a lower level of relevance in comparison to accruals with a low level of reliability. Consequently, this can be translated into weaker predictive ability and smaller significant coefficients, indicating their lower explanatory power.

The difference between the accruals with a low level of reliability and the accruals with a high level of reliability in the prediction of future *NOI* has also been confirmed by the very high *f-statistics* at the significant level of 1%. Our finding are in contradiction to the work of Richardson et al. (2005), as they argue that increased relevance could be at the expense of reliability, and consequently less reliable accrual estimates introduce measurement errors that reduce earnings persistence. In Section 2-3 a comprehensive comparison was carried out, which shows why our results are different from the work of Richardson et al. (2005). The next chapter of this thesis presents the conclusion and provides a summary of the applied methodology and all the findings.

Chapter 6 - CONCLUSION

CHAPTER 6: Conclusion

The 'articulation' of financial statements is an inherent outcome of the double entry system, and, as a result, every transaction that is kept in accounting journals has to be reflected on both the debit side and the credit side. Thus, it is theoretically possible to generate the cash flow statement variables by using the revenues and expenses in the income statement, together with changes in line-items in the balance sheet.

The most basic acknowledgement of the articulation of financial statements in the context of cash flow prediction research is already implicit in studies such as Barth et al. (2001) and Arthur et al. (2010), who suppose that the current cash flow from operations that is reported in the cash flow statement is equal to the difference between the earnings reported in the income statement and the accruals reflected in balance sheet changes.

Moreover, many studies attempt to highlight the importance of disaggregating the accruals, either in the prediction of future cash flow or future earnings, and all conclude that disaggregated models are superior to aggregated models. The main concern of this thesis is to further develop research design in this area, using an advanced level of financial statement articulation based on all line-items in the income statement, the cash flow statement and the opening and closing balance sheets, and by making explicit the most basic property of double entry book-keeping in the disaggregation of financial statement items in order to predict future cash flow and future earnings.

It was expected that the model that employed all the available disclosed variables with their double entry identity would be superior to other models that do not consider the double entry equilibrium. This thesis can be summarised into three major sections: the prediction of future cash flows, the prediction of future earnings and finally the reliability of accruals in the prediction of future earnings.

With respect to the first part of the study, dependent variables are operating, investing and financing cash flow that are created using the variables in the income statement and the changes in the balance sheets. The predictor variables in the cash flow forecasts are, accordingly, the revenue and cost items in the income statement, the balance sheets that comprise total accruals and the changes in the components of cash and shareholders' equity.

In the prediction of future cash flow three models have been compared with each other. The *OLS* model (1) applies the *OLS* regression and assumes that each of the linear equations describing the cash flow components (operating, investing and financing) is unrelated to the other two and consequently the equation residuals are uncorrelated, which is unrealistic. The next model is the *SUR* model (2), which applies the Seemingly Unrelated Regression (*SUR*) and assumes the three linear cash flow equations are related to one another as a set of simultaneous equations, through their correlated residuals. Finally, the *decSUR* model (3) also applies the *SUR* system, and the accounting relations between the predictor variables are imposed as a set of constraints, based on the double entry system.

Several issues have been compared between the three above-mentioned models, in order to find the superior model. The first and most important issue that is compared

between the models is the capability of the models in providing the coefficient signs consistent with the expected signs. For instance, it is expected that an increase in current receivables in the prediction of operating cash flows one year ahead would be consistent with an increase in cash from customers in the following year. An appropriate sign is expected to be seen for the coefficient of changes in receivables, which was explained in Chapter 6.

The *OLS* model (1) was unable to provide reasonable signs for the coefficients in the prediction of future cash flows. As there was a possibility that multicollinearity caused this problem, the multicollinearity issue was also tested for all three equations. The results reveal that no multicollinearity problems were found among the variables. In the *SUR* model (2), although all the variables were significant in the prediction of future cash flows, it was still not capable of providing the coefficient signs consistent with the predicted signs; however in this model it is assumed that all the equations are presented as simultaneous equations.

The only model that was capable of providing the coefficient signs consistent with the expected signs was the *decSUR* model (3). This is because, in addition to considering the equations as simultaneous equations and considering that there are correlations among the residuals of equations, a set of endogenous and exogenous constraints were also imposed onto the predictor variables, including the double entry constraint, which assumes that the sum of all the coefficients of predictor variables has to be equal to zero.

In order to check the robustness of our findings, all the regressions were run again, restricting the usable sample to the number of observations within each industry.

The results revealed that among almost all the industries the *decSUR* model (3) was capable of providing coefficient signs consistent with the expected signs, which other models were not able to show. However, some inconsistencies were observed through industries with a very small number of observations, such as Telecommunication. In the prediction of future earnings, results were demonstrated that show that even in very small industries, such as the Telecommunication industry, the findings are consistent.

One of the issues that has been tested among the models is the summation of the coefficients. In fact, in Chapter 4 this issue was explained in detail and it is assumed that, as all the available variables in all the financial statements have been employed in the regression equations, the sum of all the variables is set to zero and consequently the sum of their coefficients also has to be set to zero. Neither model (1) nor (2) is able to provide the zero value as the sum of the coefficients; however, the summation of the coefficients in the *decSUR* model (3) is set to zero. Moreover, to test this issue statistically, the *Wald linear coefficient test* was applied, and the result of the test strongly confirms that the summation of the coefficients is set to zero.

One of the outcomes of applying the *SUR* system instead of the *OLS* system is to derive smaller values of standard errors. In doing so, our findings reveal that the *SUR* system (models 2 and 3) provides smaller numbers for standard errors and consequently larger t-statistics in comparison to the *OLS* model (1). Although the *SUR* model (2) provides smaller standard errors and larger t-statistics in comparison to the *decSUR* model (3), as it is an unconstrained model, we cannot choose this model as the superior model to the *decSUR* model (3), as it was shown that this model was

completely unable to provide coefficient signs consistent with the expected signs, and the summation of the coefficients was not equal to zero.

As it is assumed that the residuals of the equations are correlated, the Breusch and Pagan Lagrange Multiplier test was carried out to test this issue. The results confirmed that the residuals of the equations are correlated at the significant level of 1%; however, this issue was completely ignored in the *OLS* regression, as it is assumed that residuals are uncorrelated. In fact, this test confirms the essential need for the *SUR* system.

The other important issue highlighting the superiority of the *SUR* system to the *OLS* system is the log of likelihood ratio. Our findings reveal that in terms of the log of likelihood ratio, *AIC* and *BIC* are better for the *SUR* system in comparison to the *OLS* system. Moreover, in a comparison between models (2) and (3), the likelihood ratio is better for the *SUR* Model (2) in comparison to the *decSUR* model (3), as the result of imposing a set of constraints onto the equations. In fact, we cannot choose the *SUR* model (2) just because it provides a higher likelihood ratio, as it was completely unable to provide coefficient signs consistent with the expected signs.

The other issue that was tested among the three models is whether financial statement items that incorporate both *cash and accrual* information make a higher contribution to prediction (i.e. they are more significant and larger) than those that are either just cash amounts or accrual amounts. For example, as the sales figure contains both cash and accrual information, sales may be expected to contribute more to cash flow prediction than receivables, which is based only on accrual information. The results indicate that all three models were capable of showing the mentioned difference. Indeed, this issue has no effect on the identification of the superior model.

In the end, the *decSUR* model (3), as the constraint model applied to the *SUR* system that observed the three equations as simultaneous equations, considered that the equations' residuals are correlated, and imposed a set of endogenous and exogenous constraints onto the predictor variables, has been chosen as the superior model, and in fact the rest of this thesis is based only on the *decSUR* model (3).

In the second part of this thesis, after the *decSUR* model (3) was identified as the superior cash flow prediction model, its demonstrated accounting validity should support the use of such a model in the prediction of future *earnings.* In this respect, therefore, the aim of the second part of this thesis was to compare the explanatory power of cash flows, accruals and Δ equities in the simultaneous prediction of future income, which, in this case, is disaggregated into future *operating income* and future *non-operating income.*

Indeed, the motivation for the second part of the thesis is to unravel the current confusion relating to the relative power of accruals versus cash flow predictors, where some of studies emphasise either the relative strength of accrual information (e.g. Chan et al. (2004)) or the greater predictive power of cash flow information (Arthur et al. (2010) and Sloan (1996)).

Our results solve this confusion, as all the cash flow variables in the prediction of future operating income are associated with larger significant coefficients in comparison to the accruals. Moreover, following this superiority, all the cash flow variables also provide smaller standard errors in comparison to the accruals, indicating a more accurate prediction with less prediction error. Following these two factors, the t-statistics are very large for all the cash flow variables in comparison to the accruals,

again indicating the higher predictive ability of these variables in the prediction of future operating income.

With respect to the prediction of future non-operating income, the results were not as consistent as the results for the prediction of future operating income, as some of the cash flow variables are not as capable as the accruals in explaining future non-operating income.

Both the prediction of future operating and non-operating income have been retested based on industries. When the usable sample is restricted to observations within each industry, the same results are derived, especially in the prediction of future operating income. Almost all the cash flow variables are capable of providing larger significant variables in comparison to the accruals within all the industries. However, in the prediction of non-operating income the same inconsistency can be observed, as some of the cash flow variables are not able to maintain their superiority over the accruals.

There are also possibilities to impose other cross-equation constraints for recovering same level of coefficients with respect to time-specific fixed effects or macroeconomic-specific fixed effects in addition to the industry effects, as they may have different impacts in explaining the expected variation of future cash flows and future earnings as exogenous variables; however, as our results in Chapter 5 show the superiority of the *decSUR* model (3) that considered a set of constraints including the industry effects to other models, we have not checked the sensitivity of our models to these sorts of exogenous variables. In fact, this would be an area of interest for future

research to check this sensitivity for models in imposing a set of new constraints such as time-specific fixed effects or macroeconomic-specific fixed effects.

Finally, at the end of this thesis, using the results of the prediction of future earnings, a qualitative assessment was carried out of *low and high reliability* accruals in order to assess their predictive ability with respect to future earnings. Prior work by Lev and Sougiannis (1996) and Richardson et al. (2005) suggests that accruals that are less *reliable* in terms of the greater managerial discretion they afford are expected to be more *relevant*, and accruals that are more *reliable* in terms of the less managerial discretion that they afford are expected to be less *relevant*. However, in contradiction to the work of Richardson et al. (2005), the results of this thesis show that, by using the modelling framework that was built up in the first and second parts of this thesis, the final findings provide evidence that low reliability accruals make a stronger contribution to explaining future income, and high reliability accruals make a less powerful contribution to explaining future income, which offers further insights into the trade-off between the reliability and relevance of accruals.

Our results revealed that all the accruals that are associated with the greater managerial discretion they afford are accruals with a low level of reliability, and consequently they are related to a higher degree of relevance, which is translated into the larger explanatory power that is shown by their larger significant coefficients. Conversely, it was found that other accruals that are associated with the less managerial discretion they afford are the accruals with a high level of reliability, and consequently they are related with a lower degree of relevance, which is translated into the weaker explanatory power that is shown by their smaller significant coefficients. Moreover, the

difference between the coefficients of the accruals with low levels of reliability and high levels of reliability has been confirmed by very large *f-statistics*.

As the results indicate, the constraint model is not only capable of showing the superiority of the cash flow variables in comparison to accruals, which has been a controversial issue among researchers, but it is also capable of differentiating between accruals with low levels of reliability and high levels of reliability.

Our results are of relevance to investors and analysts in estimating earnings forecasts and cash flows forecasts, managers and regulators in making decisions and recommendations relating to accruals information and cash flow information, in particular decisions regarding the level of voluntary and mandatory disclosure, and also to the Financial Accounting Standards Board (*FASB*) and the International Accounting Standards Board (*IASB*), in demonstrating the importance of financial statement articulation, which is not widely observable at present and can result in a disadvantage in terms of prediction.

Finally, as the key conclusion, the results of this thesis show that even with the current problem regarding the articulation between the financial statements, as investors are not able to see the link between the balance sheet, the income statement and the cash flow statements, we can reach unbiased results (presented by model 3). Indeed, this thesis is a very important motivation for both *FASB* and *IASB* in order to move more quickly toward preparing a more transparent articulated framework for financial statements that could help investors to see the articulation between the financial statements, and enable analysts to make unbiased estimations of the future performance of a company.

REFERENCES

ALEXANDER, D., BRITTON, A. & JORISSEN, A. 2011. *International Financial Reporting and Analysis*.

ARTHUR, N., CHENG, M. & CZERNKOWSKI, R. 2010. Cash flow disaggregation and the prediction of future earnings. *Accounting & Finance,* 50**,** 1-30.

ARYA, A., FELLINGHAM, J. C., GLOVER, J. C., SCHROEDER, D. A. & STRANG, G. 2000. Inferring Transactions from Financial Statements. *Contemporary Accounting Research,* 17**,** 365-385.

BAHNSON, P. R., MILLER, P. B. W. & BUDGE, B. P. 1996. Nonarticulation in Cash Flow Statements and Implications for Education, Research and Practice. *Accounting Horizons,* 10**,** 1-15.

BALL, R. A. Y. & WATTS, R. 1972. SOME TIME SERIES PROPERTIES OF ACCOUNTING INCOME. *Journal of Finance,* 27**,** 663-681.

BARONE, G. J. & MAGILKE, M. J. 2009. An Examination of the Effects of Investor Sophistication on the Pricing of Accruals and Cash Flows. *Journal of Accounting, Auditing & Finance,* 24**,** 385-414.

BARTH, M. E., BEAVER, W. H., HAND, J. R. M. & LANDSMAN, W. R. 1999. Accruals, Cash Flows, and Equity Values. *Review of Accounting Studies,* 4**,** 205-229.

BARTH, M. E., CRAM, D. P. & NELSON, K. K. 2001. Accruals and the Prediction of Future Cash Flows. *The Accounting Review,* 76**,** 27-58.

BARTH, M. E., CRAM, D. P. & NELSON, K. K. 2002. Accruals and the prediction of future cash flows. *Finance India, Indian Institute of Finance Business School,* 16.

BASU, S. 1999. Discussion of International Differences in the Timeliness, Conservatism, and Classification of Earnings. *Journal of Accounting Research,* 37**,** 89-99.

BAUM, C. F. 2003. *An Introduction to Modern Econometrics Using Stata*, Stata Press, StataCorp.

BAUWHEDE, H. V. & WILLEKENS, M. 2003. Earnings management in Belgium: a review of the empirical evidence. *Tijdschrift voor economie en management,* 43**,** 199-218.

BINKLEY, J. K. 1982. The Effect of Variable Correlation on the Efficiency of Seemingly Unrelated Regression in a Two-Equation Model. *Journal of the American Statistical Association,* 77**,** 890-895.

BOLLEN, K. A. & JACKMA, R. N. 1985. Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. *Sociological Methods and Research,* 13**,** 510-42.

CARLSON, R. L. 1978. Seemingly Unrelated Regression and the Demand for Automobiles of Different Sizes, 1965-75: A Disaggregate Approach. *The Journal of Business,* 51**,** 243-262.

CFA, I. 2005. A Comprehensive Business Reporting Model: Financial Reporting for Investors. CFA Centre for Financial Market Integrity.

CHAN, K., JEGADEESH, N. & SOUGIANNIS, T. 2004. The Accrual Effect on Future Earnings. *Review of Quantitative Finance and Accounting,* 22**,** 97-121.

CHARNES, A., COOPER, W. W. & IJIRI, Y. 1963. Breakeven Budgeting and Programming to Goals. *Journal of Accounting Research,* 1**,** 16-43.

CHATTERJEE, S., HADI A, S. & PRICE, B. 2000. Regression Analysis by Example. *John Wiley & Sons,* New York.

CHENG, C. & HOLLIE, D. 2008. Do core and non-core cash flows from operations persist differentially in predicting future cash flows? *Review of Quantitative Finance and Accounting,* 31**,** 29-53.

CHRISTODOULOU, D. & MCLEAY, S. 2010. The double entry constraint and the unbiased estimation of the equilibrium level of accruals. *Texas A&M University.* Texas, United States

COLLINS, D. W., GONG, G. & HRIBAR, P. 2003. Investor Sophistication and the Mispricing of Accruals. *Review of Accounting Studies,* 8**,** 251-276.

DECHOW, P., GE, W. & SCHRAND, C. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics,* 50**,** 344-401.

DECHOW, P. M. & DICHEV, I. D. 2002. The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. *The Accounting Review,* 77**,** 35-59.

DECHOW, P. M., KOTHARI, S. P. & L. WATTS, R. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics,* 25**,** 133-168.

DECHOW, P. M., RICHARDSON, S. A. & SLOAN, R. G. 2008. The Persistence and Pricing of the Cash Component of Earnings. *Journal of Accounting Research,* 46**,** 537-566.

DEMSKI, J. S., FITZGERALD, S. A., IJIRI, Y., IJIRI, Y. & LIN, H. 2006. Quantum information and accounting information: Their salient features and conceptual applications. *Journal of Accounting and Public Policy,* 25**,** 435-464.

DESAI, H., RAJGOPAL, S. & VENKATACHALAM, M. 2004. Value-Glamour and Accruals Mispricing: One Anomaly or Two? *Accounting Review,* 79**,** 355-385.

EBAID, I. E.-S. 2011. Persistence of earnings and earnings components: Evidence from the emerging capital market of Egypt. *International Journal of Disclosure and Governance***,** 174–193.

ELLERMAN, D. P. 1985. The Mathematics of Double Entry Bookkeeping. *Mathematics Magazine,* 58**,** 226-233.

FISCHER, M., GORDON, T. P. & KHUMAWALA, S. B. 2008. Tax-Exempt Organizations and Nonarticulation: Estimates Are No Substitute for Disclosure of Cash Provided by Operations. *Accounting Horizons,* 22**,** 133-158.

GILES, D. E. A. 2002. Calculating a Standard Error for the Gini Coefficient: Some Further Results. *Working Paper EWP0202,* Department of Economics University of Victoria.

HABIB, A. 2010. Prediction of Operating Cash Flows: Further Evidence from Australia. *Australian Accounting Review,* 20**,** 134-143.

HEALY, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics,* 7**,** 85-107.

HIRSHLEIFER, D., HOU, K. & TEOH, S. H. 2009. Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics,* 91**,** 389-406.

HOLTHAUSEN, R. W. 1981. Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switch-back. *Journal of Accounting and Economics,* 3**,** 73-109.

HRIBAR, P. & COLLINS, D. W. 2002. Errors in Estimating Accruals: Implications for Empirical Research. *Journal of Accounting Research,* 40**,** 105-134.

JONES, J. J. 1991. Earnings Management During Import Relief Investigations. *Journal of Accounting Research,* 29**,** 193-228.

JORDAN, C. E. & WALDRON, M. A. 2001. Predicting Cash Flow From Operations: Evidence On The Comparative Abilities For A Continuum Of Measures. *Journal of Applied Business Research,* 17**,** 87.

KHANSALAR, E. 2010. The reliability of accruals and the prediction of future cash flow. *European Accounting Association and British Accounting Association.* Istanbul: EAA and BAA.

KIM, M. & KROSS, W. 2005. The Ability of Earnings to Predict Future Operating Cash Flows Has Been Increasing—Not Decreasing. *Journal of Accounting Research,* 43**,** 753-780.

KMENTA, J. 1997. The elements of econometrics. *Macmillan Co*.

LARGAY, J. & WEAVER, S. 2009. The proposed financial reporting overhaul and business valuation: proper disclosure of financial information can be critical to a company's health. *Management Accounting Quarterly,* 10.

LEV, B., LI, S. & SOUGIANNIS, T. 2010. The usefulness of accounting estimates for predicting cash flows and earnings. *Review of Accounting Studies,* 15**,** 779-807.

LEV, B. & SOUGIANNIS, T. 1996. The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics,* 21**,** 107-138.

MANN, H. 1984. A Worksheet for Demonstrating the Articulation of Financial Statements. *The Accounting Review,* 59**,** 669-673.

NUSBAUM, E. E. & WEISS, J. 1993. AICPA/Special committee on financial reporting. *Journal of Corporate Accounting & Finance,* 5**,** 289-292.

OHLSON, J. A. 1999. On Transitory Earnings. *Review of Accounting Studies,* 4**,** 145-162

OZUNA, T., JR. & GOMEZ, I. A. 1994. Estimating a System of Recreation Demand Functions Using a Seemingly Unrelated Poisson Regression Approach. *The Review of Economics and Statistics,* 76**,** 356-360.

PFEIFFER, R. J., JR. & ELGERS, P. T. 1999. Controlling for Lagged Stock Price Responses in Pricing Regressions: An Application to the Pricing of Cash Flows and Accruals. *Journal of Accounting Research,* 37**,** 239-247.

REVANKAR, M. S. 1974. Some Finite Sample Results in the Context of Two Seemingly Unrelated Regression Equations. *Journal of the American Statistical Association,* 69**,** 187-190.

RICHARDSON, S. A., SLOAN, R. G., SOLIMAN, M. T. & TUNA, I. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics,* 39**,** 437-485.

SHAH, Z. A., BUTT, S. A. & HASAN, A. 2009. Corporate governance and earnings management an empirical evidence form pakistani listed companies. *European Journal of Scientific Research,* 26**,** 624-638.

SLOAN, R. G. 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *Accounting Review,* 71**,** 289-315.

SRIVASTAVA, V. & GILES, D. 1987. Seemingly Unrelated Reression Equations Models: Estimation and Inference. *Statistics:textbooks and Monographs,* 80.

WALDRON, M. & JORDAN, C. 2010. Comparative Predictive Abilities Of Accrual Earnings And Cash Flows In Periods Of Economic Turbulence:The Case Of The IT Bubble. *Journal of Applied Business Research,* 26**,** 85-97.

WATTS, R. L. 2003. Conservatism in accounting part I: explanations and implications. *Accounting Horizons,* 17**,** 207-221.

WU, C. & XU, B. 2008. Deflator selection and generalized linear modelling in market-based regression analyses. *Applied Financial Economics,* 18**,** 1739-1753.

YANAGAWA, R. 2003. Accruals Components and Predictive Future Cash Flows. (In Japanese. With English summary.) *Osaka Economic Papers,* 53**,** 409.429.

ZELLNER, A. 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association,* 57**,** 348-368.