The Design of Dynamic and Nonlinear Models in Cash Flow Prediction

Yang Pang

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Adam Smith Business School

College of Social Sciences

University of Glasgow

October 2015

Abstract

This thesis is concerned with the development of more accurate models for cash flow prediction. Cash flow and earnings are both important indicators of a firm's profit. The extant literature has reported various models for cash flow prediction, but they have been unable to cope with dynamics or nonlinearity in cash flow modelling.

This thesis proposes a grey-box model to capture the nonlinearity and dynamics of cash flow models. They retain a clear-box model structure but their parameters are modelled as a black box with a Padé approximant as a functional form. Two exogenous variables are used as input variables that are considered to have explanatory power for the parameter process. In addition, the thesis also employs a Bayesian forecasting model in an attempt to capture parameter dynamics of the cash flow modelling process. The Bayesian model offers an advantage of applicability in the case of a limited number of observations. Compared with the grey-box model, the Bayesian model places linear restriction on the parameter dynamics. The prior is required for the implementation of the Bayesian model and this thesis only needs to use a random parameter model as the prior.

Further, panel data estimation methods are also applied to test whether they could outperform pooled regression that is widely applied in the extant literature. There are four datasets employed in the thesis for the examination of the performance of various models in predicting cash flow. All datasets are in a panel form. This work studies the pattern of net operating cash flow (or the cash flow to asset ratio) along with time for different datasets. Out-of-sample comparison is conducted among the applied models and two measures of performance are selected to compare the practical predictive power of the models.

The grey-box model developed is verified to offer promising and encouraging performance in all the datasets, especially for U.S. listed firms. However, the Bayesian model does not appear to be superior compared to the simple benchmark models in making practical predictions. Similarly, the panel data models also cannot outperform pooled regression.

In this thesis, the traditional discounted cash flow model for equity valuation is employed to take account of the cash flow prediction models that have been developed to obtain the theoretical value of equities based on the cash flows predicted by the various models developed in this thesis. The results show that simpler models such as the random walk model is closer to market expectation of future cash flows because it leads to a better fitness for the market share prices using the new discounting model. The results show that the new valuation models developed in this thesis could have investment value.

This thesis has made contributions in both theoretical and practical aspects. Through the derivations of various models, it is confirmed that there exist nonlinear and dynamic features in cash flow prediction. Therefore, it is crucial to capture the nonlinearity using applicable tools. In addition, this thesis builds up a framework to analyse problems of similar kinds, such as panel data predictions. The models are derived from theoretical level and then applied to empirical data analyses. The results suggest that the models developed in this work could provide useful guidance for decision-making practitioners.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviations	*Definitions*
In the main body:	
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
BCN	Barth et al. (2001)
CECF	Certainty Equivalent Cash Flow
DCF	Discounted Cash Flow
DGP	Date Generating Process
DKW	Dechow et al. (1998)
DLM	Dynamic Linear Model
DSCF	Discounted Simulated Cash Flow
FASB	Financial Accounting Standards Board
GMM	General Method of Moments
IV	Instrumental Variable
LS	Least Square
ML	Maximum Likelihood
MSE	Mean Squared Errors
NPV	Net Present Value
OLS	Ordinary Least Square
P/E	Price/Earnings
RRA	Relative Risk Aversion
RW	Random Walk
SFAS	Statement of Financial Accounting Standards
TV	Time-varying
VAR	Vector Autoregressive
In the equations:	
AMORT	Amortisation
AP	Accounts Payable
AR	Accounts Receivable
CF	Cash flow
DA	Depreciation and Amortisation
DEP	Depreciation
EARN	Earnings
INV	Inventory
WC	Working Capital
In the tables and figures:	
NOB	Number of Observations

Acknowledgement

I would like to thank my family for their encouragement and support during my Ph.D. study, including my supervisors --- to me, you are already my family.

My first supervisor, Prof. Kwaku Opong, has put great effort in supporting my study. I have benefited very much from his guidance, especially have been inspired by his knowledge and critical thinking. Prof. Yun Li, my second supervisor, has been doing his best to help me become better. His attitude towards working has been always motivating me. My innovation in this thesis could never be brought out without his guidance. My third supervisor, Prof. Luiz Moutinho, has been providing his help as long as I need, even though he is very busy mostly. I am very grateful for his consideration. All of them have been caring me and I feel very lucky to have them.

Great thanks is devoted to my mother, Xia Zeng, who loves me more than her life and who is always proud of me for any piece of achievement I have ever made. She gives me confidence, without which this study could never be accomplished. I would also like to thank my father, Guoliang Pang, for his love, support, expectation and faith.

My special appreciation is given to my girlfriend, Mingshang Liu, who has been encouraging me all the time. I want to thank her for her faith in me and her prospects for our future.

My colleagues and friends are also of great help in my Ph.D. study, without whom there would be much less joy in my Ph.D. study. Thank them for friendship, especially for Yang Zhao’s and Yuping Huang’s help and encouragement! I have put all of them in my heart and wish them every success in the future paths.

Finally it is my honor to give thanks to Kelvin Smith scholarship provided by University of Glasgow. This financial support for inter-disciplinary studies has made the work possible.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature:

Printed name: Yang Pang

Chapter 1 Introduction

1.1 Background and Motivation

The most important financial information about a firm is usually contained in its annual financial statements. In financial analysis, analysts extract information from financial statements of firms and make decisions such as equity valuation, bond rating etc., based on such obtained information. From the accounting numbers that are disclosed and made available to the public, readers could form their own views about a firm's prospects. Profitability, the foundation upon which a firm exists (firms that are not for profit, e.g. charity, are not considered in this thesis), is the most important property that is of interest to financial analysts and other users who may have reasonable rights to the financial information of the firm. Financial analysts, for instance, use all their knowledge and proprietary methods in attempt to make 'correct' judgement about the financial information published by the firm.

If the value of a firm is considered as the aggregation of discounted cash flows it will receive in its lifetime, investors will price the firm and trade its equity in the stock market based on their expectation regarding the firm's future profits, which will be reflect in the share price. For commonly used valuation methods such as multiples and discounted cash flow method, the primary input is the estimated cash flow. There are then two questions to be answered by analysts:

Q1. How much is the current profit that the firm is making?

Q2. How much profit can the firm make in the future?

The first question is about the recognition of a firm's income. There are different measures for incomes, e.g. sales and earnings, of which the former is a gross quantity and the latter is a net amount considering costs and expenses. Usually the net income is of more concern because it shows whether a firm is making profits or losses, which is not conveyed by the gross income. However, net income or earnings that are disclosed in the income statement also have drawbacks. There could be doubts in the minds of readers because the earnings numbers are the results of aggregation of financial items by accountants and managers and thus may be less transparent and reliable. Whether these numbers truly reflect the profitability of firms is debatable.

Cash flow is an alternative measure for profitability and has its own advantages and disadvantages in fulfilling such purposes. Chapter 2 presents the literature review, focussing on academic studies in accounting incomes, including cash flow. The relative importance of earnings and cash flow are discussed in Chapter 2 and various methods that have been developed to forecast the two variables are also considered and analysed. The literature review sorts out the logic of why there is need for disclosure of both earnings and cash flows. The relationship between the two accounting numbers is of supplement rather than substitute. To obtain a better understanding of firms' prospects, the predictions for both earnings and cash flows are demanded.

Question two focuses on the predictive aspect of accounting variables that are associated with firms' profitability. Prediction is a more practical subject. Financial analysts provide their prediction of firms' future profits to their clients. Managers of firms also make predictions. These predictions will assist them in further decision making by managers and investors. Apart from practitioners, academic researchers also have interest in predicting a firm's profit. The disclosure of earnings has a longer history than cash flow and thus there are more academic studies on earnings in the literature. The conclusion drawn for earnings and the methods developed to forecast earnings are not necessarily applied to cash flow. Cash flow deserves more comprehensive study, including its information content compared to earnings and its predictive ability compared to earnings.

The main aim of this thesis is to develop methods that can better predict a firms' future cash flow. This topic has practical value. First, an accurate cash flow prediction will aid investors and analysts in forming rational judgements and making decisions. Secondly, a better cash flow prediction model could be explored further to provide guidance in the stock market. For instance, superior predictive power of cash flow may help investors to find stocks that are undervalued and thus profit opportunities can be created. There are differences in the prediction of earnings and cash flows, which have been developed in the extant literature. In general, the prediction of cash flow uses more information than earnings and more variables are involved. Apart from the variables in use, the shortcoming of the methods in cash flow prediction will be discussed during the review as well.

1.2 Main Contributions

1.2.1 Cash Flow Prediction Models

Linear model is widely used in accounting and finance field as it is easy to implement and analysis. The studies that will be reviewed in Chapter 2 cover a range of topics including accounting prediction, stock market and so on, most of which apply linear model as the main method. Linear regression method is popular in studies that aim to discover interactive relationships between observed variables and the regression is usually conducted using cross-sectional data. For predictive purposes, time-series models are more favoured, for instance the Box-Jenkins (1970) framework, which is also linear. Linear model is parsimonious and can provide promising performance for predictive applications. Relations between variables that can be captured by linear model are easier to understand and explain.

This thesis mainly focuses on developing superior models in predicting cash flows. Therefore, the practical performance of the models is of more concern. Linear models may not be adequate for such applications, even though they have favourable features as stated above. This study will emphasise on the modelling issue, designing novel models that could be fill the gap between simple linear models and complicated application, such as cash flow prediction. The model development is especially difficult for cash flow because the data availability is limited. With limited data and perhaps lack of theory, the dynamics of cash flow is not easy to capture.

However, this thesis has made progress in the task. This work starts from analysing the drawbacks of the models applied in previous studies and makes attempts to address the problems. Out of the procedure, a breakthrough is made in the design of a novel model that could capture the potential complexity of cash flow dynamics. The model could take account of nonlinearity and dynamics in cash flow models, which have not been addressed by previous studies. The nonlinearity is captured numerically by a black-box model. Moreover, the fundamental model for cash flow prediction can keep the simple linear form, which is easily understood and is consistent with prior literature. Therefore, the new model has the advantages of both linear and nonlinear models. The linear model in this thesis is seen as a clear-box model. The new model combining both clear-box and black-box models is thus a grey-box model. In addition, the new model could partially explain the heterogeneity across firms in making cash flow predictions.

1.2.2 Cash Flow Prediction and Stock Valuation

Apart from cash flow prediction models, this work also makes a contribution in augmenting the discounted cash flow model, making it applicable with cash flows predicted by rigorous models. The new discounting model uses the predicted distribution of cash flows by the statistical models applied in this thesis as input and produces theoretical equity value as the output. The new model incorporates Monte Carlo simulation and expected utility theory into the discounting framework and thus can be applied in more general cases. The new model has the potential to be applied in conjunction with cash flow prediction models.

Chapter 3 is devoted to explaining the problems in the existing methods that have been applied in cash flow prediction and suggest ways to solve or smooth them. There are also methods introduced in chapter 3, which are developed or designed to deal with other problems and have not been applied to cash flow prediction. There are features in those methods that make significant contribution to the extant literature on cash flow prediction, and its application. These features will be discussed and the modelling procedures outlined. More importantly, the cash flow process is shown to be dynamic and complicated. Therefore, it can be argued that the existing methods that have been applied directly in cash flow prediction may not be good enough to capture the dynamics. The novel cash flow prediction model, named the grey-box model, and the enhanced cash flow discounting model will be fully explained in the chapter.

1.2.3 Empirical Analysis

The performance of the grey-box model and other candidate models will be examined in the third part of this thesis that involves the application of the models on empirical datasets.

The empirical chapters examine the performance of the models developed in chapter 3, including cash flow prediction models and discounting model. The examination is conducted on four datasets, including stock markets and listed firms of developed countries (U.S. and U.K.) and China as a developing country and a major economy. The three datasets are analysed with results shown in Chapter 4, 5 and 7 respectively. In Chapter 4, the models will be shown with procedures and results in detail, while Chapters 5 and 7 will not list as many details as Chapter 4 to avoid repetition. The empirical part of the thesis also include a datasets of U.K. unlisted firms, which is analysed in chapter 6. This is among the first studies of unlisted firms' cash flow prediction. It is expected that the analysis on unlisted data can lead to different results from listed data and the potential difference can be informative for practitioners when making decisions. Unlike the other empirical chapters, Chapter 6 will only apply cash flow prediction models and not apply the new DCF model.

As the main innovation of this thesis, the grey-box model has shown remarkable performance in cash flow prediction. The results imply that the parameters in the cash flow model are dynamic and nonlinear, which could be partially captured by the growth rate of sales and firm age. The grey-box model hence performs very well according to the two measures that are adopted specifically for panel data application, i.e. mean squared error and average rank. The grey-box model is especially impressive in the U.S. dataset, where it does not only have the lowest mean squared error but also the lowest average rank in the out-of-sample test. In the other datasets, the grey-box model is also ranked among the best ones by these two criteria. Therefore, the nonlinear model and the selected two additional variables have been proven successful in this application. In the U.S. sample, the superior power of the grey-box model in predicting operating cash flow is utilised to see whether it could lead to encouraging investment performance in the stock market. Despite the simple and mechanical strategy applied to construct portfolios of stocks that are expected to increase in value, the performance of the portfolios is very impressive compared with contemporary market index. The results further confirm that the grey-box model could generate significant economic value.

Apart from the cash flow prediction models, this thesis has for the first time visualised the market risk aversion as a by-product of the augmented stock valuation model. Lower risk aversion implies that people are willing to pay more for the same asset, and vice versa.

Therefore, the risk aversion coefficient can be used by investors to judge to what extent the movement of the stock prices is due to the fundamental economic situation and the risk attitude of the market. For instance, in the empirical analysis of China market data, the market risk aversion until the first quarter of 2015 has reached a locally low point since the 2007 global financial crisis, along with a sharp rise of market index. The extremely low risk aversion thus can be seen as a signal for danger. By the date of submission of this thesis, the China market has been experiencing a dramatic decline, which could evidently support the assertion made according to the model.

However, not all the models applied in this thesis are successful. First, the Bayesian model's performance in the panel data setting is not superior to the benchmark models despite its dynamic and heterogeneous features. There are occasions where the Bayesian model might provide encouraging results but further studies are still demanded. Secondly, the panel data estimation methods cannot beat the pooled regression method in out-of-sample prediction, which is counterintuitive because the former has taken account of individual effects that are represented by the intercept term. Therefore, the pooled regression, as a benchmark model in this thesis, can provide strong predictive power in the panel data setting even though it makes the assumption that all individual observations are homogeneous.

1.3 Limitations

As discussed above, this thesis has made contributions to the literature by designing a nonlinear and dynamic grey-box model in predicting cash flow. For different applications, the goodness of models can be defined differently. During the study, the performance of the models will be measured by selected criteria. The criteria are selected to demonstrate limited aspects of the models' power. It should be noted that in particular situations, the criterion for selecting a good model may be different from those used in this thesis.

In addition, the comparison of the cash flow prediction models is based on the assumption that the information that is available and relevant is limited to that adopted in this thesis. All the models are based on the same set of information. In this way, the comparison focuses on the gain by designing more complicated model forms. The benchmark models are accordingly selected as they are commonly used in academia. In practice, there are much more possible models that can be applied to solve this problem. Simultaneously, there is more publicly available information than adopted by this thesis. This thesis does not have the intention to take account of as many possibilities as possible, but aims to focus on the implementation of the proposed models and discuss the underlying logic. Inevitably, there will be issues that are closely related to the topic but are not addressed or discussed in this thesis.

1.4 Organisation of Thesis

This thesis can be seen as a good opening, which can lead to many further studies and practical applications in the area of asset pricing and investment for instance. The structure of this thesis is as follows. Chapter 2 will be the literature review, discussing the relative importance of accounting variables that are used to measure incomes and listing the methods in the literature that are applied for predicting them. Following reviewing the models in the literature, Chapter 3 focuses on the model development that aims to fill the gap in the prior literature. The models designed and discussed in Chapter 3 include various cash flow prediction models and a cash flow discounting model. These models will be empirically applied to real data analysis, the results of which are described and discussed in Chapter 4 to 7, each showing the analysis of one individual dataset. Chapter 8 summarises and compares the results of different datasets, which is followed by the final conclusive Chapter 9.

Chapter 2 Literature Review

This chapter discusses the general framework of cash flow prediction and reviews the extant literature in the area. The literature reviewed will shed some light on the question of what advantages earnings and cash flow have in measuring profit and will discuss the attempts previous studies have made to predict them. In the extant literature, there are debates and comparisons of the information content of relevant accounting variables, especially earnings and cash flow. The literature will provide explanation for why there is need for cash flow prediction, especially when there are already numerous studies on earnings predictions. Studies regarding their predictive power for a firm's future income are also discussed. The difference in the prediction models of earnings and cash flows, mainly reflected by different predictive variables, will be raised. Specific models developed to forecast earnings and cash flow are explained, compared and contrasted in this session.

2.1 Measure of Firms' Profitability

Actual cash is the true income of a business transaction. Any accounting design will not alter this fact. Firms need cash to pay their debts and expenses so without cash incomes they will fail no matter how good the earnings number looks. Earnings are not necessarily treated as equivalent to cash flows. Due to the existence of inventory and trade credit, earnings are often un-matched with cash income. In addition, non-cash items such as depreciation also impact on earnings. If cash is genuinely the measure of profit, accounting earnings do not accurately reflect it.

2.1.1 Advantages of Cash Flow and Earnings

Cash is of especially great concern when it comes to evaluating a firm's solvency. Lack of cash flow usually indicates insolvency and thus difficulties in continuing business. Therefore, in judging a firm's solvency, cash flow will be more looked into than earnings as the latter can be manipulated more easily. The difference between earnings and cash income are accruals. Accrual terms are frequently seen to be manipulated by managers for such purposes as boosting or suppressing the earnings numbers. Healy (1985) has studied the connection between managers' bonus plans and their accounting choices of discretionary accruals and found that the accruals level is related to management bonus plans that are based on earnings. The study indicates that to maximise their rewards managers would attempt to manipulate accruals. Holthausen et al. (1995) have further extended Healy's study, and have emphasised that earnings tend to be managed downwards when managers' bonuses are at their maximum. DeFond and Jiambalvo (1994) have concluded that managers are likely to increase accounting income when their firms are close to debt covenant violation, providing evidence about another motivation for earnings management. Other objectives of managing earnings include smoothing income streams (Fudenburg and Tirole, 1995; Defond and Park, 1997) as managers have concerns about losing jobs as a result of volatile earnings and maintaining a growing pattern of earnings (Burgstahler and Dichev, 1997) because share price tends to be punished once an increasing pattern of earnings is broken.

Taking account of such motivations, financial statements may not fully reflect the true economic situation of a firm if there is potential benefit in managing earnings. Even if there did not exist earnings management, errors in income measurement may still occur when analysis is based on income statement and balance sheet. For example, Drtina and Largay (1985) have argued about flaws in calculating cash flow from operations using indirect method and therefore have argued for the need for cash flow disclosure. Ketz and Largay (1987) have exposed several problems in determining whether certain transactions should be categorised as operating or else financing and investing, and have therefore proposed the necessity of designing some statement that incorporates both income and cash flow. Hribar and Collins (2002) have indicated that mergers and acquisitions, divestitures and foreign currency transactions induce bias in estimating accruals, and such bias may result in false detection of earnings management. Therefore, cash flow information is crucial while the earnings are 'flexible' due to the possibility of earnings management. Cash flow, although also likely to be manipulated, e.g. categorising borrowed cash as earned cash, is more firmly and clearly measured. Cash flow disclosure hence could help reduce the difficulty in making judgements for analysts and other users of financial statements.

People do use earnings to measure profit and there are good reasons for this. For one thing, earnings are embedded with the idea that they are going to eventually turn into cash.

People who care about cash could thus estimate cash flow by deducting from earnings the unmatching proportion between earnings and cash flow. The difference is referred to as accrual terms including changes in working capital and non-cash items and so on. This is the indirect method to estimate cash flow. Earnings, besides providing an indication of the current cash flow estimate, also serve as forecast of a firm's future cash flow. By incorporating accrual components investors could grasp additional information regarding future cash occurrence. Take accounts receivable for instance, it means that a transaction has been made and the products or services may have been delivered to customers who instead of paying for them immediately are allowed to pay later. This foreseeable amount is reflected in earnings but not in cash flow for the current period.

A firm's value depends on the cash flows it generates during its whole life and these cash flows, unlike loan repayment, are not necessarily of fixed amount. Following this logic, users of financial reports ought to form their panoramic views regarding the cash flows over the target firm's entire life rather than those of merely one or two periods. Current cash income, compared to accounting earnings, may play a better role in presenting a picture of a firm's ability to run its business and thus provide a starting point for the readers of financial statements to make a forecast. However, cash flow may not be as important as earnings in changing people's expectation about a firm's future profits. Earnings may not reflect the actual cash income at that single period over which it is measured, however over the firm's whole life, the sum of each year's earnings and the total amount of cash income should approximately be equal and thus earnings can be seen as a rearranging amount of cash flow reporting. Such rearrangement is designed not to mislead financial report readers, but rather to help them better form an overview regarding cash flow patterns.

Summarised from the above discussion, both cash flow and earnings are important measures of firms' profitability. The information they provide suits different purposes. There are two main subjects users of financial information need to consider, i.e. judging and forecasting, which form the basis upon which users of the financial information make decisions and act accordingly. Users need to make such judgements when they read financial statements like: whether a firm is in a healthy situation and whether the reported earnings are the results of manipulation and so on. Cash flow plays a relatively more important role in such situations because cash is more straightforward and easier to understand. It does not mean that earnings could be completely ignored. Forecasting is even more important. In the short term, managers need to make predictions for cash income in order to decide, for example, if further debt should be brought in. Cash will be used to pay expenses and debts and managers should make sure there is sufficient cash in holding to meet the day to day needs of the firm. Long term forecasts may reflect one's opinion of a firm's prospects, which is crucial in equity valuation. Cash flows are, in general, more volatile than earnings and therefore earnings are expected to fulfil the predictive purpose better.

2.1.2 Information Contents of Cash Flow and Earnings

It is very actively studied in academia as to whether earnings are more informative than cash flow. If the comparison is based on the practice of predicting future earnings or cash flows, the results favour earnings more. Bowen et al. (1986) have studied the performance of earnings in predicting future cash flows. Their conclusion goes against the statement made by the Financial Accounting Standards Board (FASB) that earnings predict future cash flows better than current cash flow. However, more recent studies have provided evidence that earnings can make better predictions of future cash flow. Bernard and Noel (1991) have concluded that inventory disclosure can assist in predicting future sales and earnings, which implies that cash flow does not wholly capture the necessary information to predict a firm's future performance. Information provided by accrual terms still helps in making prediction. Dechow et al. (1998) have argued that earnings are a better predictor for future cash flow because of the function of accrual terms. Their study supported FASB assertions while their conclusion is against that of Bowen et al. (1986). The disagreement could probably be due to the different choice of criteria for their studies. Bowen et al. (1986) drew their conclusions based on the median of absolute prediction errors and ranks of each predictive method while Dechow et al. (1998) conclusions were based on mean absolute errors. The data availability is also different in the two studies. However, the two studies share a similarity in that both studies use earnings that are the aggregation of cash flow and accrual terms without further looking into its components.

Barth et al. (2001) extended the model of Dechow et al. (1998) by disaggregating earnings. They examine the predictive power of accrual terms for future cash and conclude that accruals have explanatory power for future cash flows. On the contrary, evidence against the superiority of earnings in predicting cash flow has recently been reported by Lev et al. (2010), where the predictive power of earnings components for cash flow was tested in an out-of-sample setting. Their results suggest that earnings do not predict cash flow better than cash flow. However, it is noteworthy that their individual predictions were based on the cross-sectional estimated model parameters. There may exist heterogeneity across individual firms, which should be taken into account when drawing conclusions.

To summarise, earnings include information of accrual terms while cash flow does not. Hence earnings are more forward looking in this sense and may offer more predictive power than current cash flow. The superiority of earnings in practice will be more apparent if cash components and accrual terms are treated separately.

More commonly, the information content of accounting variables are examined by relating them to the stock market performance of relevant companies. For one thing, equity value is considered to be equal to the sum of discounted future cash flows so stock price could be a proxy for future cash flow prediction. Moreover, stock price determines the performance of the investors' portfolio and relates more directly to their benefit and welfare. Therefore more studies try to test what accounting variables could better explain stock market performance. It is worth noting that whether earnings or cash flows are more closely related to market return will imply that one variable's influence on market expectation of future cash flow is greater than the other variable. It does not necessarily lead to the conclusion as to which variable produces more accurate prediction of future cash flows unless the market participants are rational and that their expectation is consistent with the empirical predictive results of earnings and cash flow. Earnings generally have a higher association than cash flow with stock return (see Ball and Brown, 1968, Beaver and Dukes, 1972, and Dechow, 1994 among others), which implies that earnings reflect the investor's profitability expectation better than cash flow. Dechow (1994) further indicated that for firms in a steady state and firms with a short operating cash cycle, the difference between earnings and cash flow is insignificant, which implied that the effect of accruals in predicting future cash flow is smaller for firms with these two conditions. Earnings are the sum of cash flow and accrual terms. When people look at earnings alone, cash flow cannot be differentiated from accrual terms. Because they may provide different information contents, it would be better to analyse accrual terms and cash flow income separately.

Although earnings seem to explain the stock price better, it does not indicate that cash flow provides no additional information. Cash flow could rather be incrementally informative beyond earnings number due to the differential effect of cash flow and accrual terms. Rayburn (1986) studied the association of operating cash flow and accruals with abnormal stock returns. His study suggested that both components were associated with abnormal stock return and hence both components were informative. Bowen et al. (1987) examined the incremental information of earnings, cash flow and accruals beyond the others in explaining unexpected return. They found that cash flow has incremental information conditional on earnings and it is also true the other way around. Sloan (1996) suggested that cash and accrual components of earnings should be analysed separately when predicting future earnings. He argued that the cash flow components had higher persistence and accruals had lower persistence. He constructed portfolios of stocks with long position in stocks of relatively low level of accruals and short position in those of higher level of accruals. Sloan suggested that the resulting portfolio could produce an abnormal returns if the market did not reflect the influence of the different persistence of cash flow and accrual terms, otherwise the portfolio should not have an abnormal returns. The empirical results show that an abnormal return exists. Similarly, DeFond and Park (2001) studied the impact of abnormal accruals on stock return. They predicted that abnormal accruals cannot persist and thus they were due to reverse in the future period. If this should be anticipated by the market, the share prices' reaction to a positive earnings shock would be less when abnormal accruals is positive (because the anticipated reverse will reduce future profits) than those with negative abnormal accruals, and vice versa. Therefore, looking at both earnings and cash flows, or equivalently, treating cash flow and accrual terms separately, will help us better understand the relationship between share price and financial statement.

In summary, cash flow can more accurately reflect the contemporary income of firms. Earnings numbers do not indicate that the firms have really received such an amount of money or even have the ability to generate such a level of profit. Cash, as the fundamental payment method, is as crucial to firms as lifeblood to a human. On the other hand, financial statements not only serve to show the firms' operation for the current year, but are also designed to help readers predict the prospects in terms of the firms' development. Earnings are able to better predict future cash flow than cash flow. Earnings could be seen as aggregation of cash flow and accrual terms, therefore joint analysis of cash flow and accrual terms is helpful in obtaining a sounder prediction of firms' future operation and thus, also in better understanding share price. Share price reflect investors' expectation of firms' income during their lifetime.

Both earnings and cash flow are important measures of a firm's profitability. Hence, efforts are made by managers, practitioners and academicians to predict these two variables. Ability to successfully predict a firm's future profitability can aid decision makers. The methods adopted in the extant literature to predict earnings and cash flows will be discussed in the next section.

2.2 Cash Flow and Earnings Forecast

2.2.1 Motivations of Predicting Cash Flow

Financial decisions are not based merely on the disclosed accounting numbers, but also, on the forecast of future financial status of relevant firms. In the long term, knowledge of how much a firm earns in the current period may not be as important as clues of how much this firm can earn in the coming years, especially when earnings are less persistent and/or more volatile. Financial analysts' forecasts include sales, earnings and also cash flow. A comprehensive study by Defond and Hung (2003) indicated that their database of cash flow forecasts provided by analysts increased from 1% in 1993 to 15% in 1999. They investigated the conditions under which analysts are more likely to provide cash flow forecast for companies that they follow. These conditions include large accruals, more heterogeneous accounting choices relative to industry peers, high earnings volatility, high capital intensity and poor financial health. Financial analysts tend to provide cash flow forecasts for firms with these problems or conditions so that investors who have access to their forecasts can be assisted to make better judgement and decisions than by relying on earnings numbers alone. Their study is consistent with the notion that cash flow includes extra information apart from earnings, and lists the cases as mentioned above where such extra information is highly important for decision making. In Defond and Hung's opinion, cash flow forecasts are useful where interpreting earnings and assessing firm viability is relatively difficult. As previously discussed, accruals are more easily manipulated for earnings management purposes. Therefore, a cash flow forecast plays a role in increasing earnings transparency (McInnis and Collins, 2011). Cash flow forecasts thus functions as a monitoring tool.

In addition, a cash flow forecast is fundamental for a firm's evaluation. A project's theoretical value is equal to all the cash flows it generates in its whole life discounted at particular rates. Thus financial analysts provide forecasts of target firms' future cash flows and investors rely on these forecasts to value such firms. Such valuation methods do not merely focus on the forecast for one period but require forecasts into the future. Kaplan and Ruback (1995) empirically investigated the discounted cash flow method in asset valuation. The assets they analysed are highly leveraged transactions (HLTs) because these transactions provide cash flow forecasts data for up to 10 years (or at least 4 years). They found the discounted cash flow method performed as well as the comparable method.

2.2.2 The Starting Point of Predictions

Financial analysts cover firms which they expect to be of interest to their clients and collect as much information as possible apart from the financial statements. In contrast, academic studies tend to cover a large number of firms as samples so that they can reach general conclusions. As a result, most researchers study the effect of public information disclosed by firms without bothering to collect private information for the sample firms individually and hence academic studies tend to fit models to a large sample assuming the model would suit all firms equally. In practice, however, analysts would study the specialities of their target firms individually and assign the most appropriate model as they consider accordingly. The aims of both parties diverge. Analysts who work for clients strive to make as accurate a prediction as they can for the target firms, no matter what models work and why. Researchers usually care more about the underlying reason why a specific model works.

It is common to use time-series models in making predictions. The principle is to find the evolving pattern of the target variable and describe the pattern with a parsimonious model. The simplest prediction one can make for a random variable is a constant number. If one can forecast a firm's profit for instance, the constant number assumption indicates that this firm deals with fixed amount of business without ever expanding or shrinking. Any difference between each year's incomes will be purely due to random events. Firms that could fit this description, if any, would interest investors less, although such firms may perform better than declining or even failed businesses. Investors are more likely to talk about 'growth' rather than 'constant' performance. Time series models are developed to model random processes of various types, stationary or nonstationary. The constant number process described above is a stationary process. An example of a nonstationary process is random walk. In predicting profit, constant and random walks are different. The former suggests the existence of a level from which the income of each year will not deviate far and that the incomes between consequent periods are not related. The prediction made thus does not take account of current status. A random walk process describes a firm that expands or reduces business randomly and the effect is permanent. Therefore, the prediction made for this firm's future incomes will be based on (equal to) the current income. There are other time series models beyond these two simple ones that have common applications in finance.

The challenging part in predicting incomes is the lack of theory regarding which model describes the income processes best. For example, Dechow et al. (1998) derived a cash flow forecast model based on the random walk assumption of sales. Because Dechow et al. assumed that working capital and earnings are constant proportions of sales, earnings are also implied to be random walk. Cash flow is calculated as the difference between earnings and accruals. Under these assumptions, their model leads to the statement that earnings predict future cash flows better than cash flow because the latter excludes information from accruals. Beyond one period, the prediction for earnings and cash flows will be identical, with accruals expected to become zero. Dechow et al.'s study provided an explanation of the greater predictive power of earnings for cash flow. However in forecasting, using current earnings as the prediction for future cash flow is not sufficiently good. First, as a company experiences different stages in its life (e.g. growth or maturity), its income tends to show varying patterns. A random walk assumption that expects the future to be no different from the present only describes one possible case therefore the assumption is very restrictive. Secondly, the assumption of accrual terms as constant proportions of sales shock also ignore the managerial aspect of these variables. These accounting variables partially reflect the manager's incentives and behaviours and their predictive power is thus affected. Therefore, to forecast a firm's income, it would be better to first study the time-series properties of relevant variables.

2.2.3 Time-series Properties of Earnings

As there is rarely a theory describing the process of accounting incomes, usually, knowledge is gained through looking at empirical data. Several early studies attempt to view properties of accounting incomes using statistical techniques. Little (1962) studied the growth rate defined as the percentage change of earnings for U.K. firms. Analysis of empirical data suggested that there was lack of dependence between consequent growth rates. Therefore past growth rates seem to be of no help in predicting future earnings. Beaver (1970) demonstrated that earnings had features of mean reverting but he also suggested that accounting earnings might be the product of income smoothing and thus the first difference would be a better subject for investigation. Ball and Watts (1972) did a more comprehensive analysis in an attempt to identify the possible data generating processes of incomes. Their study includes net income (original and deflated by total assets), earnings per share and sales for U.S. firms. They studied the changes of accounting incomes that were also found to be independent. To further investigate the income processes, they compared the actual data (levels instead of changes) with simulated data using four processes and observed that these income measures did not resemble processes that were deterministic functions of time. Instead they were better described by martingale or submartingale processes.

Dopuch and Watts (1972) applied the Box-Jenkins (1970) method for modelling net income of 11 steel companies. Although their study was not focused on the time-series property of incomes, they did imply that accounting changes such as transferring from straight-line depreciation to accelerated depreciation would affect the income generating process. Lookabill (1976) adjusted the methods applied by Beaver (1970) in explaining the reason why the income process is better described as a Moving Average (MA) process. Salamon and Smith (1977) adopted the same method by Ball and Watts (1972) but argued that their results did not quite support the income process being submartingale. They also indicated potential diversity for individual firms. Disagreeing with Salamon and Smith, Watts and Leftwich (1977) believed that annual earnings were well described by random walk process (with a trend) despite the statement they made that statistically with Box-Jenkins (1970) method, many individual firms' earnings seemed not to follow a random walk. They indicated that the test results should not be heavily relied on. The random walk does as well as the alternatively identified time-series models (ARMA) in prediction. Albrecht et al. (1977) expressed a similar idea. These findings have substantial impact on the interpretation of the earnings numbers because these processes would result in huge differences in people's expectations regarding future earnings, which thus need to be distinguished. The descriptive processes proposed in these work are fundamental processes but they could serve as a starting point and could be extended to add more complexity.

It is hard though to convince all market participants that earnings cannot be predicted. Rather than the random walk, which is suggested by statistical evidence, people such as financial analysts would prefer to believe that there are ways that earnings could be predicted otherwise there would be lots of job shedding. Fuller et al. (1992) argued that earnings growth was predictable other than being purely random. The rationale is that if earnings growth were truly random, there would be no difference in the expectation for all companies and the multiple of stock price to earnings would be identical for all (with risk adjusted). However the fact is that price/earnings (PE) ratio of stocks do differ across the whole market, which implies that people make heterogeneous predictions for individual firms' future earnings. Fuller et al. sorted companies into quintiles ordered by their E/P ratio. Firms of lower E/P ratio are those predicted to have higher future growth, vice versa. They then looked at the earnings growth rates of these samples in the following periods, which were consistent with the prediction. Therefore they argued that people do have ability to predict future earnings' growth, therefore random walk was not appropriate to describe the earnings process. Actually, the study by Fuller et al. was no contradiction to those discussed in the previous paragraph. Studies by e.g. Ball and Watts (1972) focused on the statistical properties of earnings growth whereas Fuller et al. (1992) looked at this question from a practitioner's point of view. Statistically, earnings growth is no different to a random process and there seems to be no correlation between consequent growth rates, which, however does not suggest the earnings' growth rate cannot be predicted by other information beyond past growth rates. Stock price is considered to incorporate all available information not limited to past earnings growths and stock price per se reflects market expectation of future earnings.

A recent study by Kryzanowski and Mohsni (2013) however argued that growth rates of earnings and cash flows, other than being random, were mean-reverting. Their study covered a much longer period of data than those discussed so far. They conducted regression of growth rates on past growth rates and some other exogenous variables such as size, age and so on. The coefficients on the past growth rates were both statistically and economically significant, suggesting that past growth rates should have predictive power for future growth rates. As time passes, one will learn by observing new data, and such results that are against studies of older times are to be expected. On the other hand, in spite of the encouraging results, Kryzanowski and Mohsni did not examine how their models actually perform in a prediction test, so it is still not determined whether the random walk process fails.

Many studies have been conducted on quarterly earnings and the results show a different pattern from annual earnings. For instance, Lorek et al. (1976) applied the Box-Jenkins methodology to model 40 U.S. firms' quarterly earnings. The predictions made by this method were compared to management forecasts, where the time-series techniques shows slightly superior power overall. The superiority is dominant in the groups for which management forecasts are relatively poor. Due to the existence of a seasonal component, a submartingale process does not capture the dynamics well (see Griffin (1977) and Foster (1977)). Griffin (1977) adopted an Autoregressive Integrated Moving Average (ARIMA) model within the Box-Jenkins framework to fit the data. Three different structures of the model were tested in Griffin's study, which suggested that changes of quarterly earnings, unlike annual ones, were not independent. However, the predictive power of the models was not tested. Foster (1977) recognised the potential problem of a limited amount of data, which might result in the diversity of identified models even if all firms' data were generated by the same process. In addition, due to finite sample, an incorrect model could outperform the true underlying process in prediction by some criteria. Apart from the sample problem, Foster also realised that structural change may bring in additional problems to model estimation. Six models were compared in Foster (1977) with a dataset including sales and expenses as well as earnings. In the prediction test, average rank and error based measures were listed as criteria for comparing these models. The results suggested the best one among them was the AR (1) model for the fourth difference (or quarter-by-quarter difference), which suggested differently from Griffin (1977) who selected a MA (1) model for the first difference of the fourth difference. These two distinct models were not compared directly because the two studies picked different models to analyse but Foster did list a candidate model that is individual estimation for each firm i.e. no restrictions on the parameters. Therefore we expect that Foster had covered a larger range of models than Griffin. It is noteworthy that in the prediction test, the parsimonious model beats the model without restriction on parameters even though the latter would be optimal in estimation, which also illustrate Foster's concerns regarding finite samples. Brown and Rozeff (1979) did a comprehensive comparison of models preferred by Foster and Griffin as well as an ARIMA model of alternative orders. Their test suggested that the Griffin model was superior to the Foster model. In addition to one-period forecast, they also tested the predictive power of five- and nine-quarter ahead prediction made by these three models. The model proposed by Brown and Rozeff seemed to outperform the other two. A more detailed comparison of quarterly earnings models can be found in Lorek and Willinger (2007). They provided further evidence on the predictive performance of the models that have been proposed above. In their comparison, they categorises their sample firms, such as high-tech, regulated and so on. According to different samples their empirical results suggested that the best model might differ, which evidently supported the diversity issue raised in Salamon and Smith (1977). Similarly, Lorek and Willinger (2008) did a comparison of the models between seasonal firms and non-seasonal firms, concluding that random walk model could be better for non-seasonal firms.

Despite its parsimony there were efforts made to enrich the ARIMA model. Manegold (1981) argued that beyond univariate time series models, it would be better to use multivariate models to predict earnings, based on the rationale that earnings as accounting numbers may have interactions with other variables in the accounting system. He added polynomial components of variables such as depreciation, debts and so on to predict earnings changes. The empirical results suggested little gain in making prediction. It is interesting however that market prices of the firms are more strongly associated with their predicted model. Lee and Chen (1990) introduced methods to detect structural changes of earnings series in addition to a simple ARIMA model. They found significant evidence that structural change is very common in accounting earnings, at least in the utility industries. This finding is against that of Beaver (1970) who did not think there were significant structural changes in accounting earnings.

To summarise the methods adopted by those studies, first it seems that there does not exist a universally optimal model, especially when people attempt to use a single model to describe all sample firms' data by assuming that the income generating process is identical or homogeneous for all firms. However there are studies that have noted the problem of this restrictive assumption (see Salamon and Smith (1977), Lorek and Willinger (2007 and 2008)). The income generating process could be more complicated than the linear models discussed above suggest. Therefore, with different data, the results may differ too. Secondly, the decision regarding whether a candidate model might be the true income generating process requires a criterion to be made. The criterion could be in-sample data fitness of the models. However there is an over-fitting problem if this criterion alone is relied on. More rigorously, an out-of-sample prediction test is often adopted to further check if the best model in the fitting stage still stands out for predictive application. The prediction tests of these studies are mostly of one period ahead. Nonetheless, the market forms the expectation of firms' future income, which is not limited to one-period ahead but for the rest of the firms' entire life. Therefore, a model that predicts one period ahead income well, although very useful in many applications, is not necessarily the best model that is consistent with the market expectation or even the true income generating mechanism. For this reason, longer term forecasting should be given more attention. This had been noted in Brown and Rozeff (1979) but appears to be neglected in most other studies. Alternatively, a model could be jointly examined by its prediction and by the performance in terms of exploiting such a model in the stock market. A model that could capture more market reaction, or results in higher abnormal return, would be considered to better fit the market expectation. Studies that address this issue include Foster (1977) and Manegold (1981) among others. However it should be noted that market expectation is not directly observed and we do not know whether it is consistent with the true underlying process of accounting incomes that is also unobservable. What we observe are the measured accounting numbers and stock prices.

2.2.4 Attempts in Cash Flow Prediction

In the U.S, a statement of cash flow has been a compulsory part of financial reports since 1987 when Statement of Financial Accounting Standards (SFAS) No. 95 was published. Before the U.S, Canada was the first country which required cash flow disclosure in 1985.

Thereafter, the time series property of cash flow started to be studied along with earnings. Before then, cash flow was indirectly estimated by deducting accrual terms and non-cash items from earnings, but there exist measurement errors in this procedure. For instance, Drtina and Largay (1985) listed several cases and examples where the indirect method cash flow may not equal the actual cash flow. Studies on accounting earnings, as discussed above, do not agree on which process should be used to describe and predict the earnings. Earnings are generally smoother than cash flows due to the smoothing effects of accruals. Therefore, it is expected that cash flow is even harder to model. Greenberg et al. (1986) studied the prediction of cash flow using earnings and cash flow (both lagged) as two candidate predictors. Not surprisingly, Greenberg et al. suggested that earnings could better predict cash flow than past cash flow, both for one-year ahead and multi-year (2 and 3 years) ahead prediction. There was a multiple regression model developed in Wilson (1987) for predicting cash flow, incorporating many accounting variables such as lagged earnings, cash flow, accruals, and capital expenditures and so on as independent variables. The model was again adopted by Bernard and Stober (1989) in an attempt to verify Wilson's study. They did not provide much discussion on the performance of the model because their focus was on the extent that unexpected cash flow may affect stock price changes rather than predicting cash flow. The model was just a by-product of their study.

Later, Hopwood and McKeown (1992) and Lorek et al. (1993) studied the time series property of quarterly cash flow. Hopwood and McKeown examined the Box-Jenkins model, the Griffin model and the Brown and Rozeff model that have been applied in predicting earnings in cash flow prediction. They found quarterly cash flow to be more difficult to forecast than quarterly earnings. Among the three models, individually identified model by the Box-Jenkins method showed slightly better performance than the other two. Lorek et al. (1993) not only tested the performance of univariate time series models as the Hopwood and Mckeown's work but again put the model of Wilson (1987) in the prediction test. The univariate time series models perform better than the multiple regression model, but Lorek et al. realised that this counterintuitive result might be explained by the fact that the estimation of the Wilson model was based on cross sectional data, which put extra restriction on the model parameters compared to the 'freer' individually estimated time series model. Lorek and Willinger (1996) proposed a multivariate time series model for quarterly cash flow prediction. The predictors include cash flow from operations, operating income before depreciation, accounts receivable, inventory and accounts payable. This model was examined in another paper by Lorek and Willinger (2008) for quarterly cash flow prediction and by comparison underperformed the Brown-Rozeff (1979) ARIMA model.

Finger (1994) examined the incremental predictive power of earnings to cash flow. A crucial contribution made by her was conducting unit root test on earnings and cash flows of the sample firms, which previous works did not have available techniques to do. The test is used to tell statistically whether a series is stationary or not. The results suggested that earnings and cash flows of different firms may present different properties. If earnings and cash flow series are nonstationary, i.e. they are better described by a random walk process, the first difference should be taken to the income series that will be stationary. To proceed with the construction of time series models, this step needs to be completed first. Finger's study showed that for about 25% of the firms their earnings and cash flow series could be treated as stationary, for which the levels instead of differences are used in predictive study. Her study indicated that there exists heterogeneity among different firms about the time series properties of their incomes, of which the majority show nonstationary income series.

Dechow et al. (1998) also studied the correlations between changes in earnings, cash flows and accrual terms and their autocorrelation as well. They started by assuming the random walk as the process for sales, according to which the autocorrelation of changes in earnings would have theoretical value of 0 and that of changes in cash flow would be negative. Empirical data showed results (on average) close to the theoretical prediction but the autocorrelation of changes in earnings appeared to be significantly negative. The negative autocorrelation of changes in cash flow is mainly due to the negative autocorrelation of changes in accrual terms. Therefore, cash flow computation which is made more complicated than earnings due to the effects of accruals would not be sufficiently captured by a random walk process. Instead, cash flow could be deemed as a combination of a random walk (earnings) and a random and stationary process (accrual terms). Dechow et al.'s study states importantly that unlike earnings, a univariate time series model may not be sufficient to model cash flow because there are predicable components in cash flows brought on by accrual terms. It would be better for cash flow prediction to include more variables in addition to lagged cash flow. The theoretical model in Dechow et al. (1998)

(DKW model from now on) simply states that future cash flow is expected to equal current earnings. Earnings do, as reported in their study, outperform cash flow in predicting cash flow. Their empirical study also included regression of cash flow on past earnings and cash flow, which was similar to and had consistent results with Greenberg et al. (1986).

Apart from the random walk assumption, the DKW model also assumes that earnings and working capital components are constant proportions of sales. Derived from these assumptions, earnings would be the optimal prediction of future cash flow. In a practical sense, this constancy assumption might easily collapse. These working capital components are important tools for financial management, and managers would place more effort in controlling them than simply adjusting them proportionally to sales changes. Such managerial behaviours would be reflected in the predictive power of these components for future cash flow. Barth et al. (2001) (BCN model from now on) therefore adjusted DKW model by disaggregating earnings into cash flow and accrual term components as multiple predictors other than DKW's use of aggregated number of earnings. In this way, the BCN model allowed these components of earnings to have different parameters in the predicting equation, which made the predictive model better fit empirical data. BCN estimate their model using pooled regression, implicitly assuming their sample firms are homogeneous in their profit generating process, whereas the DKW model was applied individually to the sample firms. Therefore BCN did not attempt to examine their model in an actual prediction test. Lorek and Willinger (2010) tested DKW and BCN models, comparing them in time-series and cross-sectional forms. Their empirical results are more supportive for time-series estimation, i.e. estimating the model individually for different firms would generate more accurate predictions.

Operating cash flow is usually reported from the statement of cash flow as a net amount. Net operating cash flow is the cash income minus the cash paid out for operative purposes but the two gross cash amounts are usually undisclosed in the statement. Cash flow can be reported in either direct method or indirect method formats, of which the latter is more commonly adopted. Direct method specifies the details of cash inflows and outflows while indirect method calculates cash flow from earnings deducting accruals and non-cash items. Direct method cash flow disclosure is considered to provide more information than the indirect method (see Hales and Orpurt, 2013 for a comprehensive discussion). Cheng and Hollie (1996) and Cheng and Hollie (2008) realised that, apart from components of accrual term, components of cash flow might also have different persistence. They then disaggregated cash flow into core and non-core components. Core cash flow components are defined as cash flows (either inflow or outflow) generated from: sales, cost of goods sold, and operating and administrative expenses. The non-core cash flow components are interest, taxes, and others. The core components were expected to have higher persistence than the non-core components, which was supported by the empirical results. Cheng and Hollie applied BCN's pool regression method as well and did not test the model in out-of-sample prediction either. From the in-sample fit, the model with disaggregated cash components improved cash flow prediction compared with the BCN model. Orpurt and Zang (2009) confirmed that cash flow disaggregation provided incremental information, and further argued that direct method cash flow statement were more informative, compared with the indirect method cash disclosure. Their study benefits from a database that includes firms with direct method cash flow statements. Therefore they could examine whether directly disclosed cash flow components would be additionally informative compared with the cash flow components calculated using the indirect method (as in Cheng and Hollie, 2008). Their study could be seen as an extension of Cheng and Hollie's because they did not develop new models but demonstrated how the disaggregated cash flow model could be enhanced with extra data availability. Farshadfar and Monem (2013 and 2013b) verified the results in an Australian setting.

Differing from the BCN model that disaggregate accruals into fundamental elements, Farshadfar and Monem (2011) used a different method to separate total accruals as discretionary accruals and non-discretionary accruals, expecting the two to have different a contribution in cash flow prediction. The discretionary part of the total accruals are distinguished by the methods originally developed by Jones (1991) and further adjusted by Dechow et al. (2003). However Dechow et al.'s method is a forward-looking model, which means that Farshadfar and Monem exploited the information only available in the future to make the prediction. Their results show that discretionary accruals have higher persistence and they suggest that Australian companies may use discretionary accounting to enhance earnings' predictive power for future cash flow, and hence a need for further verification.

The cash flow models are summarised in Table 2.1. These models show that researchers use accounting variables that they believe to be useful as predictors for cash flow prediction. The cash flow process is complicated, which is partially reflected in the fact that one cannot explain some variables predictive power in cash flow modelling, e.g. depreciation that is a non-cash item. It should be a reasonable assertion that there are other variables from the financial statements which have not been included in prior studies which can actually predict cash flow. Apart from the effect of exogenous variables, there may also exist complexity in the cash flow dynamics. So far, these discussed models are estimated in ways implicitly assuming that the models are persistent over time, no matter whether they are estimated in time series or cross section. Recall that in the discussion on earnings, Lee and Chen (1990) realised the possibility of structural changes in the earnings series. Similarly for cash flow, the time series property does not necessarily remain unchanged. A study by Kim and Kross (2005) found that cash flow became easier to predict more recently than during the 1970s as the association between current earnings and future cash flow were getting stronger over time. The models they used to predict cash flow selected lagged earnings and cash flow as predictors, as appeared in Dechow et al. (1998). The parameters of the two predictors tend to get higher with time. Therefore, cash flow might be better described by a dynamic process, which brings modelling challenges.

2.3 Accrual's Effects on Cash Flow Prediction

The cash flow forecast studies cited above applied regression of cash flow on past values of cash flow (or components) and other selected accounting variables. As stated in previous paragraph, the assumption that the persistence of the predictors will remain steady across time might be relaxed in future studies. We also need to recognise that there is diversity or heterogeneity among individual series. Rather than being exogenous and fixed, cash flow's persistence has shown dependence on levels of accrual terms. Dechow and Ge (2006) studied in detail how the persistence of earnings and cash flow differ according to accrual level. They partitioned the observations into deciles and found that the persistence of earnings and cash flow show a parabolic shape, which is to say that the observations with extreme accruals (on both high and low ends) have lower persistence for both earnings and cash flows. Their persistence increases as the accruals' level moves to the middle deciles. Such level effect of accruals might induce a biased impression of cash flow's persistence if we do not adjust for it when estimating relevant models by linear regression. Apart from the level effect, accrual quality is also of great concern to accounting researchers. Earnings management by discretionary accruals reduces the reliability of accounting information and therefore can affect people's prediction of firms' future cash flow. Brochet, Nam, and Ronen (2008) attempted to explore factors that influence the contribution of accruals to cash flow prediction. They compared the absolute prediction error of a model using cash flow alone as the predictor and a model using cash flow and accrual components (the BCN model) as predictors and they considered the difference between the absolute prediction error of these two models as a measure of accruals' contribution to cash flow forecast. This measure was regressed on several factors, including absolute value of discretionary/ non-discretionary accruals and so on, in order to examine which variables influence the additional contribution accruals would make for predicting cash flow. The method developed by Jones (1991) was the tool to disaggregate accruals into discretionary and non-discretionary accruals. The absolute value of discretionary accruals components will be a measure of accounting quality. They found that observations with positive accruals experienced greater improvement of cash flow prediction by the addition of accrual terms. Similarly, higher volatility of cash flow also indicates a greater contribution of accruals in cash flow forecasting. Finally, for accrual quality, larger magnitudes of discretionary accruals, as expected, indicates lower accruals' contribution. Therefore, treating discretionary and non-discretionary accruals separately in cash flow prediction is nontrivial due to their different behaviors. However, whether cash flow's persistence is affected by accrual quality has not been examined by these studies. Moreover, neither do the studies mentioned above show whether accruals level and quality affect the relationship of accrual terms and future cash flow.

2.4 Summary and Discussion

The literature review has discussed the advantages and disadvantages of modelling two accounting variables, i.e. earnings and cash flows. In annual reports, earnings and cash flows are not necessarily equal to each other by construction, but accumulated earnings and cash flows throughout the life of a firm are considered to converge in the long run. Firms in a steady stage or with a short operating cash cycle tend not to have a substantial deviation between earnings and cash flow. In general, earnings with information of accrual terms provide a better indication of future income than cash flow whereas the cash flow measure has the advantage of less manipulation than earnings. Therefore, the financial analysis procedure that takes both items into account should be more informative than looking into either one of them separately.

As measures of profit, prediction of earnings and cash flow is of great concern to financial market participants. Short term prediction of cash flow assists in judging a firm's ability to carry on its day-to-day business and long-term cash flow prediction is necessary to evaluate a firm's survival and equity value. For predicting earnings or cash flow, there have been studies in the time series properties of both measures. Earnings are likely to follow a submartingale process and changes in earnings seem to be independently random. Models of other ARIMA forms (submartingale could be a case of them) have also been applied to model and predict earnings. Cash flow is different from earnings because of the effect of accruals. Univariate models as applied to earnings are not sufficient for cash flow, since they ignore additional information embedded in other accounting variables. Cash flow is better predicted using past cash flow and accrual terms. Accrual level and accrual quality both affect the performance of the cash flow prediction model, which implies that the traditional linear regression model is of limited value in practice.

The studies reviewed here reported mixed results. The criteria for model comparison differ according to the researchers' preference. A comprehensive study should not only look at the in-sample performance of models, but also should include out-of-sample prediction tests to reach a sounder conclusion. In the out-of-sample test, the prediction period should not be limited to only one period ahead because longer term prediction is of more value, both to academic researchers who aim to explore the data generating process of accounting incomes and to practitioners who rely on the prediction to make real world decisions. Moreover, accounting income prediction models ought to be examined in the stock market setting. Stock prices are considered to reflect the market expectation (prediction) of future incomes of relevant firms. Combinative study of both pure predictive test and market price association would serve the purpose of describing the underlying mechanism of market expectation that is unobservable.

Cash flow is a more complicated process than earnings. Due to the limitation of data, it is difficult to find the true cash flow generating process of firms. Few attempts are made to explore the time series property of cash flow and there is no clear and general conclusion about it. The linear models developed to predict cash flow have at least three limitations, where breakthrough might start, and no study has shed light on them all:

1. Diversity in model identifications;

Although there are common factors that firms tend to share, heterogeneity inevitably exists. It cannot be expected that firms within different sectors managed by different teams will have very close operational styles and activities.

2. Possible dynamic feature of cash flow process;

   As firms grow into different stages in their life cycle, the income behaviour would be significantly different across the various stages. One static model would not be able to make an accurate prediction of the variable of interest, i.e. cash flow.

3. Criteria of comparing models;

   Apart from the development of more advanced model, there is potential problem that different criteria may suit different models, which are optimal under their respective criteria. As the loss function depends on the specific economic application, it is hard to finally pick out a model that is universally optimal.

The main contributions of this thesis are threefold. First, drawbacks of the models used in prior studies will be explained and how the drawbacks could be eliminated will be suggested. The models in this thesis are built on the DKW and BCN models, the latter of which could be seen as an extension of the former. Therefore, this thesis will start from the derivation of DKW model and suggest potential improvements. In contrast with DKW and BCN models that use linear forms to capture the pattern of cash flow, it will be shown that there are likely to be dynamic and nonlinear components in the model that have been overlooked.

Nonlinear models will have more complication in their structures than linear models and sometimes brings in computational burdens for modellers, which is not helpful if the nonlinear model does not provide better performance. After all linear model in many cases may perform sufficiently well to capture a large variation in the empirical datasets. Cash flow prediction is a practical problem, and the features of empirical data have an impact on the estimation of the linear model. However previous research did not place much emphasis on the practical aspects because the studies were mainly focused on finding out the factors, or predictors, that are able to enhance the predictability of cash flow. Most studies of cash flow prediction use panel data. There are econometric methods that have been developed to deal with the panel data issues that are not resolvable in simple linear regression studies. Therefore, this thesis will introduce the panel models and explain why they are superior to simple regression in cash flow prediction studies. It will be a bridge leading from the BCN model to more advanced practical applications without changing the model's structure.

Secondly, the main breakthrough of this thesis is the implementation and application of dynamic and nonlinear forms of the cash flow prediction model. The parameters in the original static model will be treated as time-varying (TV) state variables and a Bayesian updating method is applied to calculate the optimal prediction of these parameters and then cash flow given some certain prior distribution. The models are still within the range of linear form as the state variables are assumed to follow linear forms such as a random walk (RW) or autoregressive (AR) process. This limitation could also be relaxed and there will be no linear restrictions placed on the TV parameters. The nonlinearity of the parameter variables will be approximated by a tool known as a Padé approximant.

To make this thesis a more comprehensive study in cash flow prediction, two more issues that have not been emphasised much in previous studies will be discussed, i.e. long-term cash flow prediction and criteria adopted to compare models' performance. The idea of vector autoregressive (VAR) model will be analogously extended to a nonlinear form, in order to adapt the developed cash flow models to the multiple-period setting. Besides, taking account of features of panel data, multiple criteria will be applied to compare the performance of models. One model could be better in one attribute but worse in another. The models that will be applied in the thesis will attempt to reach a balance point in terms of the criteria.

The third main contribution of this thesis is linking the results of these cash flow prediction models to an application in pricing equity. The cash flow prediction models will produce both predicted level of future cash flows and the uncertainty about the prediction. The cash flows of firms are associated with their equity price by the framework of the discounted cash flow (DCF) model that was mathematically formalised by Williams (1938). The DCF model is very theoretical and there are drawbacks for it in practice. Therefore, the model is usually used with simple assumptions on the cash flow process and the corresponding discount rates. For instance, in the Gordon growth model (Gordon and Shapiro, 1956 and Gordon, 1959), cash flows are assumed to grow by a constant rate beyond some period in the future. Even if all market participants agree identically on the same cash flow prediction, the rates to discount them might be determined individually with respect to subjective opinions, and thus the same stock would worth differently to them, which then leads to trading activities. The method developed in this thesis will distinguish itself with the previously designed models in the following two aspects. First, the cash flows used as the discounted amounts will be the output by statistical prediction models instead of a subjectively assumed process. The first issue, i.e. predicting cash flow into the infinite future, could be achieved by implementing the models discussed in this thesis. Secondly, the uncertainty of cash flow prediction has not been an issue of major concerned in the literature. It could be measured by the dispersion of the predicted distribution of future cash flows, which could be a crucial model input to decide the discount rates for firms. This thesis will place the uncertainty of cash flow prediction into the framework of expected utility theory, where the concept of risk aversion will be introduced and will contribute to the discount rates calculation. As an analytical form of solution may not be available, numerical solutions will be obtained through Monte Carlo simulation methods. Therefore, the process of implementing the model benefits from the development of computing power. The proposed method provides an alternative way to the well-known capital asset pricing model (see e.g. Sharpe, 1964) in the determination of discount rates. This is the first time the uncertainty of cash flow prediction is used to price equities.

Chapter 3 Models and Methods

This chapter analyses and improves the cash flow models and forecasting methods reviewed in Chapter 2. Each model or method has its advantages and disadvantages; there is no single method so far that can be claimed to outperform others in all attributes. For instance, a gain in accuracy usually leads to higher complexity. Conversely, a more complex model is harder to explain and implement than a simple one. This chapter is focused on prediction models that are applied in this thesis in an ascending order of complexity.

3.1 Selection of Structures and Variables of a Cash Flow Model

The DKW model suggests that a univariate model, e.g. the Autoregressive (AR) or the random walk (RW) model, ignores important information of accrual terms for cash flow prediction. Changes in working capital will result in predictable cash flow for a later period. Therefore, the DKW model indicates that earnings, as a naive predictor, works better than cash flow itself. That is to say, even if the sales of a firm follow a random walk, cash flow does not because of the delayed cash proportions due to accruals, and the prediction of cash flow should be:

$$E_t[CF_{t+k}] = EARN_t, k = 1, 2, ..., n \tag{3.1}$$

where CF denotes net operating cash flow and $EARN$ denotes earnings. Model (3.1) is a form of a naive predictor. With sampled data, one could regress cash flow on past earnings for better data fitting and probably better out-of-sample prediction. In connection with the DKW model, a firm-specific regression:

$$CF_{i,t+k} = \gamma_{i,0} + \gamma_{i,1}CF_{i,t} + \gamma_{i,2}EARN_{i,t} + \varepsilon_{i,t} \tag{3.2}$$

shows that earnings are highly significant and incrementally useful in predicting cash flow whereas the AR term is not always statistically significant. The DKW analysis is based on a time series and the conclusion is drawn from the averaged results of predictions for individual firms.

The DKW model made several assumptions that can be summarised by two major aspects. The first is that sales follow a random walk, and the second is that working capital components, i.e. the accounts receivable, accounts payable and inventory (actually, inventory was assumed to follow a slightly more complicated procedure in their paper but here we simplify it a little), and costs are constant proportions of sales. These conditions lead to the naive prediction (3.1) using earnings instead of cash flow.

To fit empirical data better, the regression model (3.2) was applied in case the random walk did not empirically describe sales well. The regression model implies that the parameters such as $\gamma_{i,0}$, $\gamma_{i,1}$ and $\gamma_{i,2}$ are constant over time, only differing between individual firms as denoted by the subscript *i*. There is a lack of a link between the theoretical model (3.1) and the empirical model (3.2) as to why the parameters should have those features. Based on the simple DKW model, this thesis intends to explore possible features of the parameters in cash flow prediction models. Therefore, introducing a new variable r_t to denote the growth rate of sales, the DKW model can be re-derived as follows.

Net operating cash flow is the difference of cash received and cash paid out, which is represented as in the DKW paper:

$$\begin{aligned} CF_t &= (SALES_t - \Delta AR_t) - (PURCHASE_t - \Delta AP_t) \\ &= (SALES_t - \Delta AR_t) - (COST_t + \Delta INV_t - \Delta AP_t) \\ &= EARN_t - \Delta WC_t \end{aligned} \tag{3.3}$$

where the definitions of the terms are as follow:

CF : Net operating cash flow

$SALES$: Sales

ΔAR : Changes in account receivable

$PURCHASE$: Purchase that is equal to the sum of cost and changes in inventory

ΔAP : Changes in account payable

$COST$: Cost

ΔINV : Changes in inventory

$EARN$: Earnings that is the difference of sales and cost

ΔWC : Changes in working capital that is equal to $\Delta AR_t + \Delta INV_t - \Delta AP_t$

Assume that the cost, accounts receivable, accounts payable and inventory are constantly proportional to the sales. Then earnings and working capitals are also constantly proportional to the sales. Assign two constants α and β such that:

$$EARN_t = \alpha SALES_t \tag{3.4}$$

$$WC_t = \beta SALES_t \tag{3.5}$$

Define r to be the growth rate of sales and the following relation holds:

$$SALES_t = (1+r_t)SALES_{t-1} \tag{3.6}$$

With some manipulations, recursive relationships for earnings and working capital can be derived as:

$$\begin{aligned} EARN_t &= \alpha SALES_t \\ &= \alpha(1+r_t)SALES_{t-1} \\ &= (1+r_t)EARN_{t-1} \end{aligned} \tag{3.7}$$

$$
\begin{aligned}
\Delta WC_t &= \beta \Delta SALES_t \\
&= \beta r_t SALES_{t-1} \\
&= \beta r_t (1+r_{t-1}) SALES_{t-2} \\
&= (\frac{r_t}{r_{t-1}} + r_t) \Delta WC_{t-1}
\end{aligned} \tag{3.8}
$$

Therefore, (3.3) can be rewritten as:

$$
\begin{aligned}
CF_t &= EARN_t - \Delta WC_t \\
&= (1+r_t) EARN_{t-1} - (\frac{r_t}{r_{t-1}} + r_t) \Delta WC_{t-1} \\
&= (1+r_t) EARN_{t-1} - (1+r_t) \Delta WC_{t-1} + (\frac{r_{t-1} - r_t}{r_{t-1}}) \Delta WC_{t-1} \\
&= (1+r_t) CF_{t-1} + (\frac{r_{t-1} - r_t}{r_{t-1}}) \Delta WC_{t-1}
\end{aligned} \tag{3.9}
$$

Combining (3.8) and (3.9) gives:

$$
E_{t-1}\left(\begin{bmatrix} CF_t \\ \Delta WC_t \end{bmatrix}\right) = E_{t-1}\left(\begin{bmatrix} 1+r_t & \dfrac{r_{t-1} - r_t}{r_{t-1}} \\ 0 & \dfrac{r_t}{r_{t-1}}(1+r_{t-1}) \end{bmatrix}\right) \times \begin{bmatrix} CF_{t-1} \\ \Delta WC_{t-1} \end{bmatrix} \tag{3.10}
$$

This is consistent with the original DKW model (3.1) as the prediction of r_t will be zero if sales are assumed to be random walk, in which case model (3.10) reduces to model (3.1). More importantly, model (3.10) does not support the implication of model (3.2) where relationships between consequential cash flows and accrual terms remain constant over time as long as the sales do not grow at a constant rate. If sales were empirically predictable by some statistical technique rather than a random walk, the parameters in model (3.2) would be more likely to show a time-varying feature other than staying constant all the time.

Model (3.10) further implies that the heterogeneity of individual firms in predicting cash flows, i.e. parameters estimated for individual firms might be different, mainly comes from the growing states they are in. As model (3.10) predicts, in the early stages of firms, when the growth rate is usually high, there will be a large gap between the parameters of lagged cash flow and lagged accruals. Only a small portion of cash flow is predicted by the accruals of previous periods. As firms become more mature, their growth slows down and the explanatory power of lagged cash flow and accruals will gradually converge. Therefore, for different firms, cash flow prediction accuracy should turn out different as the firms are in various stages and have diverse paths of sales when data are sampled for regression, e.g. model (3.2).

The potential dynamics of the cash flow model parameters have been overlooked in previous research, but existing literature has extended the DKW model in a different way. According to (3.2), the components of accrual terms (earnings minus cash flow) are of the same parameters. The BCN model suggests that these accrual terms could have different effects in the model and therefore it extends model (3.2) to:

$$\begin{aligned} CF_{t+1} &= \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t \\ &\quad + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1} \end{aligned} \tag{3.11}$$

where *DEP* denotes depreciation, *AMORT* denotes amortisation and *OTHER* denotes other accruals. Clearly, model (3.11) has more details than model (3.2) and thus requires more data points to estimate the many parameters. Cheng and Hollie (2008) made a further extension on the BCN model by disaggregating cash flow into components, which involves even more regressors. As there are more parameters, the range of firms that allow individual estimation becomes narrower --- only those with adequate observations meet the demand. This could be a reason why BCN and Cheng and Hollie used cross-sectional regression rather than complying with the DKW framework that estimates model parameters for each individual firm. Empirically, cross-sectional regression is not appropriate for individual prediction due to the features of panel data, which will be described later. However, cross-sectional regression has the advantage in that it is a general model for all sample firms and thus it is a good tool in academic research that studies a large sample of firms without looking for individual characteristics. For instance, Hou et al. (2012) used the results of an earnings predictive model, which is estimated by cross-sectional regression, aiming to calculate the implied discount rate for equity valuation. In such cases, observations of all firms are pooled together as one sample and here individual prediction is trivial. If individual effects are considered, panel data models are more appropriate than cross-sectional regression due to the issues discussed in the next section. The predictive variables in the BCN model, i.e. model (3.11), are adopted to build all the proposed models in this thesis because model (3.11) has more details that appear to be necessary for cash flow prediction than the DKW model (3.1) and (3.2) and on the other hand is more parsimonious than Cheng and Hollie model with comparable data fitting ability.

Nonetheless, model (3.11) is still a static model. A simple, parsimonious static model can be used as a benchmark model, against which dynamic and nonlinear models developed in this thesis are tested. The next section discusses static-parameter model estimation using panel data methods.

3.2 Panel Data Models

In panel data analysis, we face the problem of dealing with data of two dimensions that combine a time series and cross-sectional data. There are a number of firms, each having their own time-series observations. This study mainly focuses on accounting variables at firm level. Therefore, the panel is relatively short because of the low frequency of a firms' financial information disclosure but broad as there are many firms in the market. In the DKW paper, model (3.2) is estimated for each firm individually and independently and thus panel data models are irrelevant. Individual estimation in this thesis is difficult to conduct because DKW uses a relatively long dataset and fewer parameters to estimate. In many applications, data of different groups are analysed together in order to draw a common and general conclusion, especially when we have a short panel that does not allow individual estimations.

A question naturally arises that there may exist heterogeneity across different groups and the methods of model estimation should adapt to this situation accordingly. From the simplest case where all groups are homogeneous, i.e., for different groups all variables follow the same data generating process (DGP), same distribution and moreover the relationships between the studies variables are also identical to the most complicated or heterogeneous case where every group is different from each other, the statistical tools used to analyse the data are carefully developed to meet the assumptions made on different levels of heterogeneity. For instance, fixed effect and random effect models are linear estimators used in the situation where we assume there is heterogeneity only in the intercept term among the groups while the factor loadings on the predictive variables remain homogeneous. Model (3.11) then becomes:

$$\begin{aligned} CF_{i,t+1} &= \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} \\ &\quad + \beta_5 DEP_{i,t} + \beta_6 AMORT_{i,t} + \beta_7 OTHER_{i,t} + \varepsilon_{i,t+1} \end{aligned} \tag{3.11a}$$

Note that subscript i for individual effects only shows in the intercept $\beta_{i,0}$. If the coefficients for the input variables are allowed to differ but they still follow some certain distribution (usually normal), we could use a random parameter model that has structure:

$$\begin{aligned} CF_{i,t+1} &= \beta_{i,0} + \beta_{i,1} CF_{i,t} + \beta_{i,2} \Delta INV_{i,t} + \beta_{i,3} \Delta AP_{i,t} + \beta_{i,4} \Delta AR_{i,t} \\ &\quad + \beta_{i,5} DEP_{i,t} + \beta_{i,6} AMORT_{i,t} + \beta_{i,7} OTHER_{i,t} + \varepsilon_{i,t+1} \\ \boldsymbol{\beta}_{\mathbf{i}} &\sim N(\boldsymbol{\beta}, \mathbf{V}) \end{aligned} \tag{3.11b}$$

Further, if the firms are completely heterogeneous, there will be no gain in analysing the data in panel. It would be optimal to undertake individual estimation. In reality it would be too ideal and optimistic to expect total homogeneity across firms. There will be some differences between two firms even if they sell the same product. First the sizes and/or business scales of different companies vary, which often imply that their financial variables are not likely to be in the same statistical distribution. Besides, firms of different sizes also have vulnerability of different degree to business risks. On the other hand, firms at least share some similarities. The management style cannot be innovative for every firm. Managers try to learn from successful companies (leaders) no matter what industries they specialise in. There is the tendency for people to study firms in the same sectors as these firms may share some common rules. The degree of heterogeneity is an empirical question that is hard, if possible, to measure. Every modelling method will have its applicability.

3.2.1 Estimation of Model with Individual Effects

The fact that firms are different in size potentially causes the problem that the unexplained parts of the dependent variable by the predictors for different firms are not identically distributed. In model (3.11a), this problem is reflected in two aspects: the first is that the intercept may differ in level and the second is that the error term may have different variances. Difference in intercept will bias the parameter estimation by cross-sectional regression if the individual effects correlate with the independent variables. This case is named fixed effect in panel data econometrics. In predictive applications, the individual effects of each firm have to be considered and distinguished, otherwise the prediction will be significantly biased. The individual intercepts could be resolved by adding dummy variables, such as in Barth, et al. (2001), which is applicable but results in too many parameters to be estimated. Alternatively, the intercept term of individual effect can be eliminated in the following two routines --- demean and first difference.

The different variances of the error terms also have an impact on the estimation procedure. Once estimation is based on raw data without transformation, the observations are naturally weighted by the firms' sizes. A regression based estimation procedure aims to minimise the sum of squared errors or another error norm. Therefore, firms which have a large variance in the error term tend to dominate the results, especially when the degree of heterogeneity is high. A routine applied in the BCN paper is to deflate all the variables by a size-related variable, such as the total assets and shares outstanding. The same deflation method will be followed in this thesis with average total assets of each firm as the deflator.

The two routines for estimating a panel data model with individual effects are well developed in panel data econometrics (for example, see the textbooks by Cameron and Trivedi, 2005). This thesis first attempts to apply the panel data estimation method in a predictive application (including out-of-sample performance). Therefore, whether the parameters are consistent is not the major concern of this thesis as long as the estimation methods discussed below could provide satisfactory performance in prediction.

Demean

If (3.11a) holds, for each individual firm *i* the following holds too:

$$\overline{CF}_i = \bar{\beta}_{i,0} + \beta_1 \overline{CF}_i + \beta_2 \overline{\Delta INV}_i + \beta_3 \overline{\Delta AP}_i + \beta_4 \overline{\Delta AR}_i$$
$$+ \beta_5 \overline{DEP}_i + \beta_6 \overline{AMORT}_i + \beta_7 \overline{OTHER}_i + \bar{\varepsilon}_i \tag{3.12}$$

Deduct (3.12) from (3.11a), and we have:

$$CF_{i,t+1} - \overline{CF}_i = \beta_{i,0} - \bar{\beta}_{i,0} + \beta_1 (CF_{i,t} - \overline{CF}_i) + \beta_2 (\Delta INV_{i,t} - \overline{\Delta INV}_i)$$
$$+ \beta_3 (\Delta AP_{i,t} - \overline{\Delta AP}_i) + \beta_4 (\Delta AR_{i,t} - \overline{\Delta AR}_i) + \beta_5 (DEP_{i,t} - \overline{DEP}_i)$$
$$+ \beta_6 (AMORT_{i,t} - \overline{AMORT}_i) + \beta_7 (OTHER_{i,t} - \overline{OTHER}_i) + \varepsilon_{i,t+1} - \bar{\varepsilon}_i \tag{3.13}$$

Therefore, removing the group mean of each variable will eliminate the individual effect by the intercept term while keeping the other parameters unaffected (actually the demeaning have an impact bringing endogeneity into the model, which will be further discussed later). The individual $\beta_{i,0}$ could be calculated by manipulating equation (3.13) as:

$$\hat{\beta}_{i,0} = \overline{CF}_{i,t} - (\beta_1 \overline{CF}_{i,t-1} + \beta_2 \overline{\Delta INV}_{i,t-1} + \beta_3 \overline{\Delta AP}_{i,t-1}$$
$$+ \beta_4 \overline{\Delta AR}_{i,t-1} + \beta_5 \overline{DEP}_{i,t-1} + \beta_6 \overline{AMORT}_{i,t-1} + \beta_7 \overline{OTHER}_{i,t-1}) \tag{3.14}$$

First Difference

An alternative to demeaning is to take the first difference for all the variables, and then the model becomes:

$$\Delta CF_{i,t+1} = \beta_1 \Delta CF_{i,t} + \beta_2 \Delta^2 INV_{i,t} + \beta_3 \Delta^2 AP_{i,t} + \beta_4 \Delta^2 AR_{i,t}$$
$$+ \beta_5 \Delta DEP_{i,t} + \beta_6 \Delta AMORT_{i,t} + \beta_7 \Delta OTHER_{i,t} + \Delta \varepsilon_{i,t+1} \tag{3.15}$$

where Δ^2 denotes the second order difference i.e. Δ_t-Δ_{t-1}. By differencing, the intercept, which is assumed to differ across individuals but to be constant over time, has been eliminated. The difference model is a bit easier in practice than the demeaned model as it does not bother to calculate the individual group mean.

The least squares (LS) method to estimate the parameters in (3.13) and (3.15) is unable to solve the endogeneity problem because of the inclusion of AR term (for details see e.g.

Cameron and Trivedi, 2005). Endogeneity causes inconsistency in the estimation of parameters. Arellano and Bond (1991) apply the general method of moments (GMM) in order to obtain theoretically consistent parameters. The GMM estimator is applied to the differenced model i.e. (3.15), and it takes the form of:

$$\widehat{\beta}_{AB} = \left[\left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i'\tilde{\mathbf{X}}_i\right)\right]^{-1}\left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_i'\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i'\tilde{\mathbf{y}}_i\right), \tag{3.16}$$

where $\tilde{\mathbf{X}}_i$ is a $(T-2)\times(K+1)$ matrix with all regressors including the lagged dependent variable for the *i*th company and $\tilde{\mathbf{y}}_i$ is a $(T-2)\times 1$ vector of the *i*th company's dependent variable. $\mathbf{Z}_i$ is a $(T-2)\times r$ matrix of instrumental variables (IV). $\mathbf{W}_N$ is weighting matrix. The GMM estimator is used for the first time in applications of cash flow prediction, which will be compared empirically with the demean method to see whether this theoretically more consistent model could provide better prediction as well.

3.2.2 Estimation of Random Parameter Model

Comparing with (3.11a), the random parameter model (3.11b) assumes the parameters of (3.11), instead of being strictly identical, could differ from firm to firm as long as they follow a known distribution. The results of estimating (3.11b) will be the distribution of the parameters among those sample firms, including the vector of the parameter means and their covariance matrix. Therefore, the LS method is not applicable for this requirement. The distribution of the parameter vector will be estimated by the maximum likelihood (ML) method.

Given a known distribution, the likelihood function of an observation is known and could be calculated. Because the likelihood function is calculated with parameters, the principle of ML method is to find the parameters that lead to highest likelihood for the observed data. Assume parameter vector $\boldsymbol{\beta}_\mathbf{i}$ is randomly distributed as in (3.11b), and the probability density function for $\boldsymbol{\beta}_\mathbf{i}$ is denoted by $p(\boldsymbol{\beta}_\mathbf{i} \mid \boldsymbol{\beta}, \mathbf{V})$. Given T observed cash flows within a group *i*, the joint conditional likelihood of these observations is:

$$f(\mathbf{CF}_i \mid \mathbf{X}_i, \boldsymbol{\beta}_i, \sigma^2) = \prod_{t=1}^{T} f(CF_{i,t} \mid \mathbf{X}_{i,t}, \boldsymbol{\beta}_i, \sigma^2) \tag{3.17}$$

where $\mathbf{X}_{i,t}$ denotes the vector of the predictors and σ^2 denotes the variance of cash flow prediction residual. To obtain the unconditional likelihood, the function above needs to be integrated over $\boldsymbol{\beta}_\mathbf{i}$:

$$f(\mathbf{CF}_i \mid \mathbf{X}_i) = \int_{\boldsymbol{\beta}_i} f(\mathbf{CF}_i \mid \mathbf{X}_{i,t}, \boldsymbol{\beta}_i, \sigma^2)\, p(\boldsymbol{\beta}_i \mid \boldsymbol{\beta}, \mathbf{V}) d\boldsymbol{\beta}_i \tag{3.18}$$

The log-likelihood function of all the observations across all sample firms will be:

$$\begin{aligned} \ln L &= \sum_{i=1}^{n} \ln[f(\mathbf{CF}_i \mid \mathbf{X}_i)] \\ &= \sum_{i=1}^{n} \ln\left\{ \int_{\boldsymbol{\beta}_i} f(\mathbf{CF}_i \mid \mathbf{X}_i, \boldsymbol{\beta}_i, \sigma^2)\, p(\boldsymbol{\beta}_i \mid \boldsymbol{\beta}, \mathbf{V}) d\boldsymbol{\beta}_i \right\} \\ &= \sum_{i=1}^{n} \ln\left\{ \int_{\boldsymbol{\beta}_i} \left[\prod_{t=1}^{T} f(CF_{i,t} \mid \mathbf{X}_{i,t}, \boldsymbol{\beta}_i, \sigma^2) \right] p(\boldsymbol{\beta}_i \mid \boldsymbol{\beta}, \mathbf{V}) d\boldsymbol{\beta}_i \right\} \end{aligned} \tag{3.19}$$

which is to be maximised to obtain the optimal estimation of $\boldsymbol{\beta}$, $\mathbf{V}$ and σ^2. The estimation of random parameter models in this thesis follows the method introduced in Pinheiro and Bates (2006) because the estimation procedure is implemented by MATLAB software, which is used in analysing the empirical datasets.

The random parameter model does not only provide the point estimate of the parameters as is done by the methods discussed in section 3.2.1 but also calculates the variance-covariance matrix of the parameters. Therefore, this method illustrates the distribution of parameters among different firms. Models (3.11a) and (3.11b) do not provide individual estimation of the parameters, therefore the distribution of parameters is not directly used in practical prediction. However, the variance-covariance matrix is useful to give the modellers an impression regarding the sample firms and also could be treated as an important input to the model that will be described in next section. In this model, the parameter vectors are allowed to be heterogeneous across different firms in addition to being recursively dynamic.

3.3 Dynamic Linear Model

The section above has demonstrated the estimation procedure of model (3.11) assuming the parameters are static, i.e. they do not change over time. Equation (3.10) proposed a dynamic form for the cash flow process and from here on methods for implementing dynamic models will be attempted. Rewriting model (3.11) in dynamic form (by dynamic, it means time-varying parameter) gives:

$$\begin{aligned} CF_{i,t+1} = \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t} \\ +\beta_{i,t,5}DEP_{i,t} + \beta_{i,t,6}AMORT_{i,t} + \beta_{i,t,7}OTHER_{i,t} + \varepsilon_{i,t+1} \end{aligned} \tag{3.20}$$

Obtaining efficient estimate for those parameters is extremely challenging. They change value all the time. If the question is about a sufficiently long time series, a rolling window would be a good choice for estimating time-varying parameters. However we are unlucky in that the length of data series is merely enough for static model estimation for most firms. The rolling window is a 'free' way of parameter estimation as it does not place any restrictions on or assumptions about the parameters' processes. But in our case, we need to do so for compromise. In this thesis, two methods for TV parameter models are implemented. The first applies a Bayesian method assuming the parameters to follow a linear process, i.e., AR or random walk. The second applies non-parametric (black box) model, in an attempt to fit the data using nonlinear parameter processes. These two methods will be thoroughly explained separately in two sections.

The rationale of adopting the Bayesian framework is based on its two major advantages. First Bayesian model treats subjective probability as a crucial input, which opens a lot of room for business applications. Secondly but even more importantly, the method shown below is capable of dealing with a limited amount of observations. Moreover, a Bayesian model could also provide a routine to construct dynamic models. The model that will be introduced in this section is named dynamic linear model (DLM) as in West and Harrison (1997).In the business disciplines, people need to make various decisions. Most important decisions will be based on careful analysis and reasoning, which are believed to be optimal under certain circumstances. Business environments are complex as a large number of factors interact with each other, which also results in a large amount of information for analysts and decision makers to process. Decisions on one hand are made with reference to collected data and on the other hand also reflect people's subjective views. The process that people's judgement changes and adjusts when they observe can be seen as learning, and we adopt the Bayesian method to model this process.

The DLM method (West and Harrison, 1997) is not developed particularly for panel data application but the topic of this thesis benefits much from this method. In practice, not like studies in dynamics of stock prices or returns of which observations are made available at high frequency by trading activities, the low frequency of accounting disclosure (annual reports) by firms has caused the sample of individual firms to be short for most firms. This feature limits the possibility of individual estimation and forecasting and thus panel data methods are applied as described in section 3.3 with the compromise of stricter assumptions. The DLM model, on the contrary, could deal with a limited amount of data even if the number of observations is less than the number of parameters. Therefore, the by-product of DLM is that heterogeneity in firms could simultaneously be taken into account along with the parameter dynamics. The DLM could be used as an extension to the static panel models, especially the random parameter model. The Bayesian method has an important input that is the prior. Other than assigning the prior arbitrarily or randomly, it would more make sense to assign some statistically estimated results as the prior distribution, from which the Bayesian rule could be applied to update.

3.3.1 Model Structure

Bayes' law describes how a belief is adjusted or updated after one observation arrives. Considering proposition A and event B, one's belief in A is reflected in the prior probability $P(A)$ and after the event B occurring such belief will change accordingly. Denote the conditional probability of event B on A as $P(B|A)$ and event B's unconditional probability as $P(B)$ and Bayes' theorem could be presented in the form:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.21}$$

where $P(A|B)$ is the posterior probability of A after observing B. This is the conditional probability of A based on the condition of observing B. The formula describes the probabilistic updating procedure. The posterior belief will become the new prior when next observation is due. Bayesian analysis is based on this theorem and has many applications. The idea is attractive as by the theorem subjective opinions are brought together mathematically with observed data that is considered as 'objective'. This is extremely important and useful for business applications where decisions are made in complicated ways and subjective intervention plays a crucial role. Extending the Bayesian rule to the cash flow model (3.21) we have:

$$\begin{aligned} CF_{i,t} &= \mathbf{x}'_{i,t-1}\boldsymbol{\beta}_{i,t} + \varepsilon_{i,t} \\ \boldsymbol{\beta}_{i,t} &= \boldsymbol{\beta}_{i,t-1} + \mathbf{v}_{i,t} \end{aligned} \tag{3.22}$$

where $\mathbf{x}_{i,t-1}$ denotes the vector of independent variables in (3.20) including lagged cash flow and $\boldsymbol{\beta}_{i,t}$ is the parameters vector. Differently from static models, parameters in this case are assumed to evolve over time and, for simplicity, follow a random walk process. In this model the vector of state variables $\boldsymbol{\beta}$ is allow to be time-varying, with an innovation vector $\mathbf{v}$. The disturbance terms of the cash flow process and the parameters are assumed to follow the distributions of:

$$\begin{aligned} \varepsilon_{i,t} &\sim N(0, W_i) \\ \mathbf{v}_{i,t} &\sim T_{n_{t-1}}(\mathbf{0}, \mathbf{V}_{i,t}) \end{aligned} \tag{3.23}$$

where N denotes normal distribution and $T_{n_{t-1}}$ denotes t distribution with n_{t-1} degrees of freedom. In this linear system, cash flow and the predictors can be actually observed while vector $\boldsymbol{\beta}_{i,t}$ is latent. Before observing any data for a particular firm, there is some information available to help us form a prior distribution of the initial value of the parameters $\boldsymbol{\beta}_{i,0}$. Such prior estimates could come from the analysis, if available, on other individuals or groups that are believed to possess similar features to the target firm. The prior is quite subjective as no data is available to support it so the prior is not necessarily true but it will reflect what the decision makers believe to be true according to their knowledge, experience and relevant information. Apart from $\boldsymbol{\beta}_{i,0}$, the uncertainty of ε_t, measured by its variance *W*, is unknown too and thus also requires a prior approximation. Generally, the Bayesian model can be summarised in three steps which will be listed below. The subscript *i* will be dropped for simplicity.

3.3.2 Procedure of Bayesian Updates

Step one: Quantify the initial estimates at time 0

Based on information at time 0, D_0, make a prior about $\boldsymbol{\beta}_0$ and *W*. In the Bayesian framework, people use another concept called precision ϕ instead of dealing with *W*:

$$\phi = \frac{1}{W} \tag{3.24}$$

which makes the analysis easier as a gamma distribution can be used to describe ϕ. The prior distributions will be:

$$\begin{aligned} &(\boldsymbol{\beta_0} \mid D_0) \sim T_{n_0}(\mathbf{b_0}, \mathbf{VB_0}) \\ &(\phi \mid D_0) \sim G(\frac{n_0}{2}, \frac{n_0 S_0}{2}) \end{aligned} \tag{3.25}$$

The following variables thus need initial guesses:

n_0: Degrees of freedom; large n makes the distribution converge to normal and small n implies lower confidence in the prior belief;

S_0 : Initial estimate of *W*;

$\mathbf{b_0}$: Initial estimate of parameter vector;

$\mathbf{VB_0}$: Initial estimate of parameters' variance-covariance matrix; this could be approximated depending on the confidence interval of the parameters;

Step two: Make prior prediction based on previous estimate

When the data of explanatory variables is available, a reasonable prediction could then be made. First the dynamic parameter vector is predicted, and then the prediction of cash flow follows:

$$\begin{aligned} E_{t-1}(\boldsymbol{\beta}_\mathbf{t}) &= \mathbf{b_{t-1}} \\ E_{t-1}(CF_t) &= E_{t-1}(\boldsymbol{\beta}_\mathbf{t})\mathbf{x}_{t-1} \end{aligned} \tag{3.26}$$

As the parameters are assumed to be a random walk, their prediction made on prior information is simple and naive. The prior for the variances of $\boldsymbol{\beta}_\mathbf{t}$ and CF_t made at time *t*-1 are:

$$\begin{aligned} \mathbf{V}(\boldsymbol{\beta}_\mathbf{t}) &= \mathbf{VB_{t-1}} + \mathbf{V_t} \\ V(CF_t) &= \mathbf{x}'_{t-1}\mathbf{V}(\boldsymbol{\beta}_\mathbf{t})\mathbf{x}_{t-1} + S_{t-1} \end{aligned} \tag{3.27}$$

Details about how to assign values to $\mathbf{V_t}$ will be discussed later. The prior distributions of $\boldsymbol{\beta}_\mathbf{t}$ and CF_t for the forthcoming period based on the information of previous period are:

$$\begin{aligned} (\boldsymbol{\beta}_\mathbf{t} \mid D_{t-1}) &\sim T_{n_{t-1}}(E_{t-1}(\boldsymbol{\beta}_\mathbf{t}), \mathbf{V}(\boldsymbol{\beta}_\mathbf{t})) \\ (CF_t \mid D_{t-1}) &\sim T_{n_{t-1}}(E_{t-1}(CF_t), V(CF_t)) \end{aligned} \tag{3.28}$$

Step three: on observed y_t, update the posterior estimate of the model

When the new cash flow data is observed at time t, it will be compared with the prediction made *a prior*. The prediction error will be calculated as:

$$e_t = CF_t - E_{t-1}(CF_t) \tag{3.29}$$

The error term will be used to update the posterior distribution of the parameters as the way shown below:

Posterior estimate of $\boldsymbol{\beta}_{\mathbf{t}}$:

$$\mathbf{b_t} = E_{t-1}(\boldsymbol{\beta}_{\mathbf{t}}) + \mathbf{A_t} e_t \tag{3.30}$$

$$\mathbf{A_t} = \frac{\mathbf{V}(\boldsymbol{\beta}_{\mathbf{t}})\mathbf{x}_{t-1}}{V(CF_t)} \tag{3.31}$$

Degrees of freedom *n* is increased by 1:

$$n_t = n_{t-1} + 1 \tag{3.32}$$

S, the estimate of *W*, is updated as well:

$$S_t = S_{t-1} + \frac{S_{t-1}}{n_t}\left(\frac{e_t^2}{V(CF_t)} - 1\right) \tag{3.33}$$

$\mathbf{VB}$, the variance-covariance matrix of the dynamic parameter vector $\boldsymbol{\beta}_{\mathbf{t}}$:

$$\mathbf{VB_t} = \frac{S_t}{S_{t-1}}(\mathbf{V}(\boldsymbol{\beta}_{\mathbf{t}}) - \mathbf{A_t}\mathbf{A_t^{'}} V(CF_t)) \tag{3.34}$$

As a result, after the observation of CF_t , the difference between the prior belief and the reality will emerge and then the updating procedure operates on the prior estimates, resulting in the posterior distribution:

$$(\boldsymbol{\beta}_\mathbf{t} \mid D_t) \sim T_{n_{t-1}}(\mathbf{b}_\mathbf{t}, \mathbf{VB}_\mathbf{t})$$
$$(\phi \mid D_t) \sim G(\frac{n_t}{2}, \frac{n_t S_t}{2}) \tag{3.35}$$

The posterior estimates of the linear model combine the prior belief and the new data, being updated each time a new observation shows up. A large prediction error e_t normally suggests that the prior used to make prediction seems unreliable, so the greater e_t is the more $\mathbf{b}_\mathbf{t}$ is adjusted accordingly, vice versa. The newly updated posterior estimates will be applied in predicting future targets. In consequence, step two (prediction) and three (updating) will form a loop as new observations keeps entering the system.

3.3.3 Variance Discounting

One thing that remains unexplained is the covariance matrix $\mathbf{V}_\mathbf{t}$ of state equation shock term, which is usually unknown. $\mathbf{V}_\mathbf{t}$ is used in predicting the covariance matrix of the parameter vector $\boldsymbol{\beta}_\mathbf{t}$. A simple approach dealing with it is to assign $\mathbf{V}_\mathbf{t}$ as a proportion of the variance of the state variables:

$$\mathbf{V}_\mathbf{t} = \frac{\mathbf{VB}_\mathbf{t\text{-}1}}{\delta_v}, \tag{3.36}$$

so that predicted state variance becomes:

$$\mathbf{V}(\boldsymbol{\beta}_\mathbf{t}) = \mathbf{VB}_\mathbf{t\text{-}1} + \mathbf{V}_\mathbf{t} = \frac{(\delta_v + 1)\mathbf{VB}_\mathbf{t\text{-}1}}{\delta_v} = \frac{\mathbf{VB}_\mathbf{t\text{-}1}}{\delta_B}. \tag{3.37}$$

The discount factors δ_v and δ_B are chosen subjectively and it is an empirical question what the optimal value should be in particular applications.

3.3.4 Monitoring

Apart from the three steps described, the Bayesian model could potentially detect deteriorating performance based on the prediction errors. The model would be considered to work well if the errors lie within a reasonable range that is induced by the model's estimates. However, occurrences of errors of large magnitudes would indicate that the model may start to fail and an alternatively specified version is needed. This is usually due to the changes in the circumstances, or structural break, e.g. changes of management. The previously estimated model that is applicable for the past may not work as well for the current time and the future. In the Bayesian model design, there is a monitoring mechanism that could track the possibility of models' failure or occurrence of large outliers. Once either possibility becomes high, the model will experience an adjustment before regular updating. The monitoring procedure is based on the calculation of Bayes' factor defined as:

$$H_t = p_0(Y_t \mid I_{t-1}) / p_1(Y_t \mid I_{t-1}), \tag{3.38}$$

where p_0 denotes the probability of observation Y_t under the monitored model M_0 and p_1 denotes the probability of Y_t under alternatively specified model M_1. The Bayes' factor for the sequence of k consecutive observations is defined as:

$$H_t(k) = \prod_{r=t-k+1}^{t} H_r = \frac{p_0(Y_t, Y_{t-1}, \ldots, Y_{t-k+1} \mid I_{t-k})}{p_1(Y_t, Y_{t-1}, \ldots, Y_{t-k+1} \mid I_{t-k})}. \tag{3.39}$$

One option for M_1 is the scale shift model. Assume under M_0

$$e_t \mid I_{t-1} \sim N(0,1), \tag{3.40}$$

then

$$p_0(e_t \mid I_{t-1}) = (2\pi)^{-0.5} \exp(-0.5 e_t^2). \tag{3.41}$$

M_1 will suggest that e_t has a standard deviation k rather than 1, so

$$p_1(e_t \mid I_{t-1}) = (2\pi k^2)^{-0.5} \exp(-0.5(e_t / k)^2)\,. \tag{3.42}$$

Hence the Bayes' factor at time t is:

$$H_t = k \exp(-0.5 e_t^2 (1 - k^{-2}))\,. \tag{3.43}$$

Define a quantity:

$$L_t = \min_{1 \le k \le t} H_t(k)\,, \tag{3.44}$$

and the run-length l_t takes the value $(1+l_{t-1})$ if $L_t<1$ and 1 if $L_{t-1}>=1$.

Using these quantities, the monitoring procedure is described below:

Step one:

Calculate H_t, and compare it with a threshold value τ. If $H_t<$ τ, then observation Y_t is treated as a potential outlier. In this case, omit this observation and proceed to t+1 for next observation, and reinitialise the monitor by setting $l_t=0$ and $L_t=1$. Further increase uncertainty $\mathbf{V}(\boldsymbol{\beta}_\mathbf{t})$ in equation (3.37) by discounting:

$$\mathbf{V}(\boldsymbol{\beta}_\mathbf{t}) = \frac{\mathbf{VB_{t-1}}}{\delta_B \delta_C}\,. \tag{3.45}$$

Calculate Bayes' factor H_t for the new observation. Once $H_t>=$ τ, proceed to step 2.

Step two:

Calculate L_t and corresponding run-length l_t. If $L_t>$ τ and $l_t<=4$, M_0 is favoured so proceed to regular updating. If either $L_t<$ τ or $l_t>4$, the monitor issues a signal of possible break down of model M_0. In this case, increase the uncertainty by (3.45) and reinitialise the monitor by setting $l_t=0$ and $L_t=1$. Then proceed to regular updating procedure. For the next new observation, start step 1 again.

3.4 Nonlinear Dynamic Model

The derivation of the Bayesian forecasting model is based on the assumption that the parameters follow a linear process. The section introduces a nonlinear dynamic model which suggests that the parameters' dynamics are nonlinear and unknown. In section 3.1, model (3.10) provides some clue regarding one potential explanation for the nonlinearity. Empirically, the nonlinearity is approximated using data with a more general form of function than (3.10). Assume that for the nonlinear dynamic model, each parameter from model (3.20) is assumed to be controlled by the process:

$$\beta_{i,t} = F(\mathbf{z}_t) \tag{3.46}$$

where $F(\mathbf{z}_t)$ is a nonlinear function of some variables $\mathbf{z}$. According to (3.10), the nonlinearity might closely relate to the growth of sales. r_{t-1} and r_t are likely to have effects in the nonlinear function. If there is predictability in sales' growth rates, the rule dominating the dynamics of the parameters could be predicted by statistical models. The empirically nonlinear form of the parameters series F is unknown so an approximation function is required to model the nonlinear relation. The described procedure is an application of a black box model, which is the antonym of a clear box model. The latter appears mostly in physics and engineering where physical laws are definitely known and applied without any uncertainty. In social science, however, the laws, if any, are unknown and the social environment is full of uncertainty, structural changes, chaos etc. Therefore clear forms of function describing the relationship and interaction between variables are not available. With the assistance of data, the relationship could be approximated using some functional forms with certain accuracy. There are several options for such functions. For instance, a neural network is considered as a universal approximator that is able to approximate any function (Cybenko, 1989). Similarly, Taylor series and Fourier series are two more examples that can approximate functions with any degree of accuracy. This thesis adopts the Padé approximant (Tan and Li, 2002) for the nonlinear function as the method is efficiently accurate with only a few coefficients to determine. In the spirit of Tan and Li (2002), the joint model of (3.20) and (3.46) has comprised of a grey-box system as model (3.20) could be considered as a clear-box in this application.

Say $F(r_{t-1}, r_t)$ is function of r_{t-1} and r_t .For the predictive application r_t is not available at time t-1. Therefore before determination of F for cash flow prediction, the prediction of r_t is required as an input along with r_{t-1}. In the simplest case, assume r_t is predicted by some function of r_{t-1} only, then the nonlinear function F is reduced to a function only with r_{t-1} as input variable, which could be approximated by a Padé approximant of order 2/2:

$$F(r_{t-1}, r_t) = F^*(r_{t-1}) \approx \frac{a_0 + a_1 r_{t-1} + a_2 r_{t-1}^2}{1 + b_1 r_{t-1} + b_2 r_{t-1}^2} \tag{3.47}$$

where a_i and b_i are the coefficients to be determined. In this model form, the dynamics of parameters depends totally on the predictability of exogenous variables. If r_t is purely random and thus could not be predicted by r_{t-1} , the parameters, especially the AR parameter $\beta_{i,t,1}$ in model (3.20), will tend to be insignificantly affected by r_{t-1}, and thus it could be expected that the expression of (3.47) would reduce to a constant, i.e. the model will be no different from a static one. Other variables that can be used as input to empirically predict r_t could also be included in the nonlinear model.

Clearly the grey-box model places a heavier burden on the amount of calculation, which is not a major problem with the rapid development of computing power. The coefficients could be estimated with conventional statistical method such as LS and ML by defining various forms of loss function.

3.5 Long-Term Prediction

In previous sections, the models that will be applied in the thesis are listed and described thoroughly. Either static or dynamic, linear or nonlinear, they are all based on the original one-period BCN model. In practice, long term cash flow prediction is of concern for investors, without which it will be inefficient to price the assets that generate the cash flows. A model that better predicts in the long term will help people to evaluate the target asset or project. A natural extension of the one-period BCN model is to increase the lag length for multi-period ahead predictions. The main drawback of this option is that the maximum period ahead that can be predicted is critically limited by data availability. Share value is considered as the aggregation of all cash flows that will be received in the future discounted back to the current time. In principle, the cash flows need to be predicted to infinite future, which is impossible to realise by a regression model.

The difficulty is mainly brought about by the fact that the cash flow model includes multiple predictive variables. In univariate time series models, it is simple and straightforward to derive a multiple-period model by the recursive results. In multi-variable models, however, the effects of the other variables need to be considered as well. For simplicity, those variables could be treated as exogenous. Alternatively, long term forecast can also be implemented by estimating predictive models of all those variables and recursively forecast them into any period in the future. In linear models, the Vector AutoRegressive (VAR) model (Sims, 1980) is the corner stone and has the form:

$$\mathbf{y}_{i,t} = \boldsymbol{\beta}\mathbf{y}_{i,t-1} + \mathbf{v}_{i,t} \tag{3.48}$$

Where $\mathbf{y}_{i,t}$ is the vector of all relevant variables in the model; $\boldsymbol{\beta}$ is the parameter matrix identifying the predictive system; $\mathbf{v}_{i,t}$ is the disturbance vector for all variables. Once $\boldsymbol{\beta}$ is estimated, all variables could be predicted by the following relationship:

$$\widehat{\mathbf{y}}_{i,t+k} = \boldsymbol{\beta}^{k}\mathbf{y}_{i,t} \tag{3.49}$$

Where k denotes the number of periods ahead required to be predicted. VAR is not limited by the length of data samples and hence is a more flexible tool for the thesis, especially for the application that will be introduced in section 3.7. Dynamic models will be in the same spirit but the mathematical form is different from that of static parameters. The dynamic models, either linear or nonlinear, take the form below instead:

$$\mathbf{y}_{i,t} = \boldsymbol{\beta}_t\mathbf{y}_{i,t-1} + \mathbf{v}_{i,t} \tag{3.50}$$

Hence the prediction will take the form of:

$$\widehat{\mathbf{y}}_{i,t+k} = \prod_{n=1}^{k} \boldsymbol{\beta}_{t+n} \mathbf{y}_{i,t} \tag{3.51}$$

3.6 Model Performance Evaluation

Sections 3.2 to 3.5 have discussed various possible ways of predicting cash flow using the foundation BCN model. It is not trivial to select one criterion or multiple criteria to evaluate these models, especially in the panel data setting. In problems of individual time-series prediction, it is simple and straightforward to compare the performance of different models. The only effort is to pick a loss function as the criterion of comparison. For instance, there are error-based criteria that are most commonly used, such as mean square errors (MSE) or mean absolute error (MAE) etc. A better model is supposed to generate smaller prediction errors than others. Alternatively, there are non-parametric measures of performance in prediction. Sometimes, people's decisions are not based on the exact predicted value but on the predicted directions. For instance, an investment decision could be made as long as the return is predicted to be positive no matter whether it is high or low. Multiple criteria could be jointly raised for a comprehensive comparison.

In panel data, because there are many individual firms, the comparison procedure becomes contradictive. In practice, the results of specific individuals are more concerned than the aggregate group, and firm-based specific prediction is of more value. Therefore, the models, even if estimated with all firms, need to reach predictions for individual firms. Usually, it is difficult to judge two models according to the aggregated measure, e.g. SSE of all firms. It is highly likely that one model that produces a smaller SSE perform worse for half of the sample firms. Therefore, apart from the usually used measures in single series problems, the two-dimensional feature of panel data has posed another requirement on predictive models in practice, i.e. generality. A good model is not limited to fulfilment of aggregated accuracy but also need to show superior power for as many sample individuals as possible. Studies in the early periods, e.g. Ball and Watts (1972), calculate the average rank of each model in fitting each observation or each group in general as a measure to evaluate the models. This thesis will also adopt the rank measure as a criterion for judging models' performance.

3.7 Equity Pricing by Discounting the Predicted Cash Flows

This section is an application of the outputs of the cash flow prediction models introduced in previous sections. A direct application of cash flow prediction is the DCF method to stock valuation. In financial theory, the value of a stock is equal to the cash incomes the firm will generate discounted by some rates. For most cases, the cash flows are unknown beforehand so they need to be approximated or predicted with uncertainty. Usually the discounted cash flow (DCF) method is applied to individual stocks: first predict the cash flow for the firm and secondly assign a discount rate for it. There is a problem in the DCF model that is individuals might have different predictions and discounting factors regarding the same firm, therefore the worth of the firm may differ among investors. In this case, one cannot decide whose price is 'correct'. This section will demonstrate a framework that takes all available stocks in the market into consideration and develop a general rule to discount the predicted cash flows to reach a relatively fair price for stocks.

To apply DCF, both the prediction of cash flow and determination of discount rates are challenging tasks. In the appendix, it is shown that by expected utility theory, all risky cash flow could be transformed into a certainty equivalent cash flow (CECF). The resulting CECF, as the name suggests, could be discounted by risk free rate to reach a present value.

3.7.1 Stock Valuation Using the DCF Method

Step 1: Define the utility function

In the Appendix A, it is derived that, with some assumptions, the utility function could take the form of:

$$u(c) = \frac{c^{1-\rho}}{1-\rho} \tag{3.52}$$

where ρ is a constant measuring the relative risk aversion (RRA). The larger it is, the more risk averse are the people. Assume that the utility function represents the whole market and ρ is the parameter to be calibrated. This utility function has the attribute to restrict the net present value (NPV) by DCF to be divisible. One share of a company entitles the receipt of one share of its cash flows, and hence the price of the share should be the NPV divided by the number of shares. This conclusion is derived with the assumption of normal distribution for the cash flow. However, empirically cash flows distribution is skewed. Therefore, the utility function could not guarantee exactly the divisibility of the NPV, but it still could work as a proxy and good starting point due to its mathematically simple form and calculative viability.

Step2: Predict cash flow distribution using Monte Carlo simulation

Because the distribution of cash flow is not normal, there will be no analytical form solution available as equation (A11) in the Appendix A. Monte Carlo simulation provides a possible path to model cash flow distribution numerically. For simplicity, assume the cash flow follows a random walk with normally distributed noise. For such a process, possible paths of cash flows would look like those depicted in Figure 3.1a. There are 100 simulations for a period of 100 with starting value of 1 and a random innovation term of which the standard deviation is 0.1. The chart looks roughly symmetric, which should be the case ideally for random walk. However in business world, cash flow income would not exactly look like this. If a firm incurs losses continuously, the firm would no longer exist. Therefore, some condition to control for the mechanism should be added to the simulation process: for instance, if there are 3 negative cash flows in consequence, all future simulated cash flow for the firm will be 0 (meaning the firm broke down then). The newly simulated paths look like Figure 3.1b. The paths are no longer symmetric and there is no clear mathematical form to describe them, therefore the calculation from now on gives numerical solutions rather than analytical form solution.

Step 3: Transform the simulated cash flow distribution to certainty equivalent cash flow.

For each period in the simulation, the certainty equivalent cash flow c^* is calculated by the relationship:

$$u(c^*) = E(u(c)) \tag{3.53}$$

In this procedure, negative cash flow will be defined as 0 as the utility function (3.52) could not deal with negative number all the time.

Step 4: Discount the certainty equivalent cash flow series by risk free rate.

A risk free rate is not always available for every year required to match the simulated cash flow series. Therefore the method by Nelson and Siegel (1987) will be applied to interpolate and extrapolate the yield curve.

As long as the simulated series are long enough, the present value of the cash flows i.e. the theoretical price of the firm, will converge to a finite number (the certainty equivalent cash flow will get smaller as the uncertainty increases and the discount factor gets larger as time passes).

3.7.2 Calibrate the Constant Relative Risk Aversion Coefficient

Denote the constant relative risk aversion coefficient by ρ, which is assumed to represent the risk aversion of the whole market. The risk aversion coefficient should be of identical value for all stocks. Therefore, stocks of which the prices imply different values of ρ will be considered as mispriced. Denote the market share price for firm i by P_i and the theoretical value by the DCF method by PV_i. Define expected return er_i of share i as:

$$er_i = \frac{PV_i - P_i}{P_i} \tag{3.54}$$

which measures the rate of mispricing (not related to time). In the range of the whole market, the aggregated mispriced proportion of all shares should go to zero (otherwise arbitrage opportunity emerges), therefore ρ is calibrated by setting:

$$E(er_i) = \frac{\sum_i \frac{PV_i - P_i}{P_i}}{n} = 0 \tag{3.55}$$

where n denotes the number of firms in the market. ρ for a particular time point can be solved by:

$$\rho = \arg\min[E(er_i)^2] \tag{3.56}$$

3.8 Summary

In summary, this chapter proposes and derives models that can be used to predict cash flows or evaluate stocks according to cash flow forecast. The main focus of this thesis is on making cash flow prediction. Therefore, this chapter starts from the derivation of a cash flow prediction model with some assumptions that are adopted by prior studies. It is noted that there is potentially nonlinearity in the cash flow model. This thesis develops a grey-box model to capture both the fundamental linear form of previously developed cash flow models and the nonlinearity. To be specific, the nonlinearity is modelled by the Padé approximant.

Along with the grey-box model, several other models are also proposed to compare with each other, including panel data models and the Bayesian model. The principles and estimation methods are discussed in this chapter. This chapter also argues that it is necessary to select two criteria, i.e. MSE and average rank, for measuring each model's performance.

In addition, an equity pricing model is developed in this chapter in order to relate the cash flow prediction methods to potential applications in stock markets. This novel model applies Monte Carlo simulation method. The cash flow prediction models and this valuation model will be examined in the following empirical chapters.

Chapter 4 Empirical Study I: U.S. Market

The first empirical study of this thesis is on annual data of U.S. firms. As the world's largest economic entity and the most developed country, evidence from analysing U.S. data would be important and could be treated as a benchmark. The variables required for this study are all disclosed financial information, which are directly available from firms' annual reports. All accounting variables are obtained from the WRDS Compustat database. The data in this chapter covers all listed firms in the U.S. spanning the period from 1957 to 2013. Apart from the firm-level data, risk free rates are also required to calculate discounted cash flows. From the Datastream database, the yields on the constant maturity one-year, two-year, three-year, five-year, seven-year, ten-year, twenty-year and thirty-year bonds are obtained as risk free rates[1]. The yields data is of monthly frequency.

Before proceeding to the application of cash flow prediction models, it seems helpful to have a general impression about the cash flow process first. There is no doubt that each sample firm will have their particular cash flow paths. Larger firms tend to generate higher cash flows and vice versa, which makes the cash flows of different firms uncomparable. Therefore, the first step in comparing them is to normalise the cash flow series of each firm which is accomplished by deflating each firm's cash flows by their initial positive cash flow observation. Negative observations, if they appear at the beginning of any firm's cash flow series, are excluded. In such way, every firm will grow from the same starting point, i.e. cash flow of exactly unity, no matter what time they start from. This process will be less influenced by specific year effects because the time when each firm started to enter the sample is diversified. The indicator for time will then not be the absolute years, e.g. 1987, 1990 or 2015 etc. but is denoted as the number of years ahead of each firms' beginning time. Thus, firms of not only various sizes, but also, different decades could be compared in parallel. The results should therefore be more general than alternatively bringing firms' observations of the same particular period together.

[1] The Datastream symbols for the Treasury bond yields are USTRCN1, USTRCN2, USTRCN3, USTRCN5, USTRCN7, USTRCN10, USTRCN20 and USTYCO30R.

Cash flow disclosure was not compulsory until 1987. The DKW paper attempts to estimate cash flow indirectly from balance sheet and income statement in order to increase the available sample with tolerable measurement and calculation errors. In this study, cash flows before 1987 will be estimated indirectly following the DKW approach so as to increase the number of observations available. Operating cash flow is estimated using the general formula: net income + depreciation and amortisation – changes in non-cash current assets + changes in current liability. Using the data after 1987, the correlation of actually disclosed cash flow and indirectly estimated cash flow by above formula is calculated to be 0.82. Therefore, it seems the level of estimation error is quite tolerable.

4.1 Trend of Cash Flow Levels

There are 21,905 firms with at least one available cash flow observation that is either disclosed or indirectly estimated and 239,835 firm-year cash flow observations in total from 1957 to 2013 for the entire sample. The distribution of available observations of each firms are shown in the histogram of Figure 4.1. The firm with the most observations is IBM that has 56 observations. Most sample firms have less than 30, even 20, observations. Both the upper and lower extremes 1 percent of the normalised cash flows are excluded. For each leading period, i.e. period ahead of the starting time, there will be a distribution of cash flow levels. The mean, median and 95% range (from 2.5 to 97.5 percentiles) are shown in Figure 4.2. The figure only shows for 42 periods ahead of starting time as the number of sample firms available reduces along with the x-axis. The results beyond 43 periods ahead are based on very few observations (no more than 16) hence they cannot be seen general and meaningful enough. In the chart, there is an obvious and almost monotonic upward trend for the mean and median, the latter being lower, cash flow series. The mean level of cash flow in period 42 is nearly 52, which implies an annual growth rate of roughly 10%. Similarly, the median level in period 42 is 26, implying an annual growth rate of 8%.

However, is it directly conclusive that U.S. firms in general have very high cash flow growth rates as shown in Figure 4.2? The answer is not that straightforward and the conclusion needs to be treated carefully because of survivorship bias. From Figure 4.2, it is clearly seen that the cash flow distributions of each period are asymmetric. The asymmetry increases with periods. In 42 years, there are firms whose cash flow could reach 250 times of their initial number but there is no firm that can possibly have cash flow of -250. There seems to be a lower bound for cash flows. Companies are managed by people who eliminate the possibility of symmetric cash flow distributions. Firms that incur losses one year after another will not be allowed to lose permanently. There would be hard decisions made, e.g. the management might be replaced, go into bankruptcy or be taken over. If the firms could not recover, there would be no motivation for them to exist in business any longer. Therefore, there is that chance that firms may grow positive cash flows permanently to a very high level but little chance that firms' negative cash flows' decline will continue permanently.

In this study, a simulated illustration to show the effect of survivorship bias was conducted. Assume cash flow follows a random walk process with the noise term following normal distribution of zero mean and variance of 16:

$$\begin{aligned} &CF_t = CF_{t-1} + \varepsilon_t \\ &CF_0 = 1 \\ &\varepsilon_t \sim N(0,16) \end{aligned} \tag{4.1}$$

Theoretically, this process will have an expectation of 1, i.e. the initial value, in any period. However, a quit rule to the series is set. That is, if there are 5 series of negative numbers in a row the series stops. The maximum length for the simulation is 42. The simulation is run for 10,000 times the number of sample firms. The distribution of the simulated sample is illustrated in Figure 4.3. There are some similar features between Figure 4.2 and 4.3 visually. The simulated data also shows an increasing pattern for the mean level of cash flow series despite the fact that they are generated from a random walk process. Therefore it is suggested that survivorship bias will have an upward effect on the general conclusion drawn from the sample. The true expectation of cash flows trend may not be as high as shown in Figure 4.2.

To take the survivorship bias into account, the survival rates of the firms are calculated. For each period from 1 to 42, the survival rate is calculated by dividing the number of firms that have the available number of observations of the specified length by the number of firms that appear early enough in the sample to provide the required number of observations. For example, there are 17043 firms providing observations of one year ahead of their initial time. To examine the 1-year survival rate, the denominator is the number of firms of which the first observations appeared before (including) year 2011, which is 18257. Thus the 1-year survival rate is 93.35%. To calculate the 42-year survival rate, the numerator is the number of firms that provides observations of 42 years ahead of their initial time, i.e. 294 and the denominator will be the number of firms that first appeared before (including) 1970, i.e. 2015 firms. Firms that are younger cannot provide the required length of cash flow series anyway and thus are not counted. The rates indicate the proportion of firms that are listed at any time in the market for a certain length of period, and this is depicted in Figure 4.4. The survival rates suggest that in 10 years less than half of the U.S. firms remain in the market; in 42 years, less than 15%. Therefore, it should be realised that when calculating the mean of cash flow series with the surviving sample firms, a large proportion of bad cases are not contained in the sample. To make an adjustment, the mean of cash flow series is multiplied by the survival rates, which should be a more accurate way to describe the true unconditional expectation of the cash flow pattern. The adjusted mean cash flow series are depicted along with the originally calculated cash flow means in Figure 4.5. The adjusted mean suggests that it is more appropriate to expect, in general, that a firm's cash flow could grow to 7.58 times, rather than 52 times, of its original value in 42 years, which implies an annual growth rate of roughly 5%.

4.2 Parameter Estimation Using Different Methods

From this section on, the actual disclosed cash flow data is used in the modelling of cash flow. Data before 1987 when cash flow disclosure became compulsory are not included. The criteria in the BCN paper are followed and exclude observations if they belong to any of the following categories:

- Financial services firms (SIC codes 6000-6999);
- Sales less than $10 million;
- Share price less than $1;
- Earnings or cash flow in the extreme upper and lower 1 percent of their respective distributions.

This provides a sample of 99,845 firm-year observations. Table 4.1 provides the descriptive statistics for: cash flow, depreciation and amortisation, changes in account receivable, changes in account payable, changes in inventory and other accruals, all variables deflated by average total assets of each firm. On average, cash flow deflated by average total assets is about 0.07 for the sample firms, but the dispersion is very large with a standard deviation of 0.13. There are also special cases in the sample where the minimum and maximum cash flow observations are greater than the firms' average total assets in magnitude.

The model to be estimated is:

$$\begin{aligned} CF_{i,t+1} = {} & \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} \\ & + \beta_5 DA_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon_{i,t+1} \end{aligned} \tag{4.2}$$

where DA denotes depreciation and amortisation. Note the intercept term is assumed to be identical for all firms in the pooled regression. First, pooled regression is applied for data between 1987 and 1996 to compare with the results in the BCN paper which were based on this time period. For the rest of the thesis, the whole sample is partitioned into two subsamples: data from 1987 to 2005 is used for in-sample estimation, and data from 2006 to 2013 is used for out-of-sample prediction performance comparison. Parameters in Equation 4.2 are then estimated using four different methods: pooled regression, demean, first difference and Arellano-Bond estimator, all of which have been explained in the previous chapter.

The estimated results are summarised in Table 4.2. Numbers in parentheses are *t* statistics based on heteroskedastic robust standard error. The results in Table 4.2 are obtained using the Stata package. The second column shows the results for the period between 1987 and 1996. The number of sample observations is 27,630. The results are very close to that of the BCN paper. All selected variables are both statistically and economically significant and the signs of the parameters are consistent with those reported in the BCN paper.

For the rest of the table, estimation period is from 1987 to 2005. Column 3 lists the results of pooled regression, which do not deviate much from that using the shorter period of data. The fourth, fifth and sixth columns are the estimators considering individual effects, where the intercept terms vary across firms but do not stay constant over time. Therefore, the intercept terms are not shown in the table. Column 4 gives the estimation results applying demean method. There is a major difference in the AR parameter, i.e. β_1, between the results by this method and pooled regression. Pooled regression, which ignores individual effect, tends to bias parameters upwards, therefore the demean method shows that AR parameter is 0.392, much lower than 0.61 in the second column and 0.69 in the third column, both of which are estimated by pooled regression. Column 5 provides the results estimated using the first difference of the variables. This method results in a negative autocorrelation in the AR term. The AR parameter shown in column 5 is negative and statistically significant at 0.01 level. The negative parameter is due to endogeneity caused by the taking of first differences and thus it is not consistent. Conclusions that can be drawn from the other parameters are not too different from previous 3 columns except for the depreciation and amortisation term (β_5). This term is no long statistically significant when using the first difference estimator. Results reported in column 6 using the Arellano-Bond estimator, also support the insignificance of depreciation and amortisation. The Arellano-Bond estimator applies the GMM method, which is implemented in this thesis by assigning all independent variables as instrumental variables (IV). It is considered that the Arellano-Bond estimator would take account of the endogeneity brought on by taking the first difference for the variables. Therefore, the AR parameter in column 6 is positive in contrast with column 5.

In conclusion, the parameters in Equation 4.2 could be quite diverse by the use of different methods and it is therefore hard to make a sound conclusion out of the results. However, the results in Table 4.2 suggest that pooled regression, which is widely used in the extant literature, biases the AR parameter upwards. Thus applying pooled regression in such applications would give readers the false impression that the cash flows are very persistent, which could cause severe problems for decision makers. In addition there is no clear conclusion as to whether depreciation and amortisation is actually significant for the cash flow model.

4.3 Random Parameter Model and Bayesian Model

Although Equation 4.2 is estimated by different methods, they all assume the parameters to be identical across firms (except the intercept). This assumption is very restrictive and questionable in practice. To verify the applicability of this assumption, one needs to see how the parameters are distributed across different firms.

In order to obtain reliable results when estimating the cash flow model parameters for individual firms, the firms should have large enough sample sizes. The results of individual estimations are mainly used for illustration, which will not be used out-of-sample, therefore the whole sample is used to make individual estimation of model (4.2) instead of using the in-sample period. Sample firms that have no less than 20 observations are used to make individual estimations and 960 firms are available for this purpose.

The distributions of the parameters could be visualised by plotting histograms. After calculating the regression results for each available firms, the distributions of parameters, adjusted R squared and the residual variance of each firms are plotted in Figure 4.6. The parameters seem to be distributed quite symmetrically from a visual impression. The dispersion of each parameter's distribution is so large that it is hard to convince readers that the parameters are identical across firms. Therefore, it is doubtful whether a general model that assumes homogeneity such as model (4.2) is appropriate for practical use. The distribution of the adjusted *R* squared values can be seen as a demonstration of the extent that the sampled firms' cash flows are predictable. The situation varies substantially across firms. There are firms whose data is fitted to an accuracy of over 90% while there are also firms whose cash flows are so unpredictable that the adjusted R squared values turn out to be negative. The covariance and correlation matrix of the 7 parameters are calculated and they are merged into Table 4.3. The variances of all variables are very large, especially for changes in inventory and depreciation and amortisation. The parameters are also highly correlated with each other.

As shown by the individual estimation results, it seems inappropriate and inflexible to apply a simple, general model with uniform parameters to all firms. However, the sample size is a crucial restriction for the majority of the sample firms to conducting individual estimation. A potential solution for the problem is the Bayesian model that was introduced and explained thoroughly in section 3.3. The Bayesian model relies heavily on prior information and thus can deal with samples of smaller sizes. In the most extreme case where there is no observed data at all, the predictions will fully depend on the prior. The choice of prior would subjectively depend on decision makers' initial information and belief. For instance, the individual estimation results just obtained provides the distribution of the parameters, therefore they could be used as prior for Bayesian updating. Moreover, to exploit as much information as provided by the data, the random parameter model can be estimated.

The random parameter model assumes that the parameters of model (4.2) are randomly distributed rather than remaining identical across firms. Rewriting (4.2) adding subscript *i* that indicates different firms gives:

$$\begin{aligned} CF_{i,t+1} = {} & \beta_{i,0} + \beta_{i,1} CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} \\ & + \beta_{i,5} DA_{i,t} + \beta_{i,6} OTHER_{i,t} + \varepsilon_{i,t+1} \end{aligned} \tag{4.3}$$

The point estimate of the parameters will be the mean of their distribution and they are therefore not directly applicable in making individual predictions. To make individual predictions, one needs to identify the specific values of parameters that apply to the particular target firms. Therefore, the random parameters model could lead to the knowledge of the distribution the parameters might follow, which would be a good option for the Bayesian model prior. Unlike the individual estimation procedure, the random parameter model uses all the data without excluding firms with few observations. This might therefore be more informative than individual estimation results. The random parameter model is estimated using the maximum likelihood (ML) method, assuming the parameters are normally distributed across firms and the variances of prediction error of each firm are identical. The results for cash flow prediction are summarised in Tables 4.4 and 4.5.

Table 4.4 provides the comparison of mean values of parameter distributions between the random parameter model and the results averaged for individual estimations. There are similarities for the seven parameters for the two estimation procedures. In the last row the variances of residuals are also listed. The value on the left, column 2 , (individual estimation column) is the average value for individual firms while the right column, column 3, (random parameter model) assumes this quantity to be identical for all firms. Table 4.5 provides the variance-covariance matrix of the seven parameters by the random parameter model, similar to Table 4.3 that is obtained by individual estimation. In general, the variances of each parameter that are estimated using the random parameter model are smaller than those summarised from the individual estimation results, which might be explained by the occurrence of outliers in the individual estimation procedure. In this aspect, it seems that the results by the random parameter model should be more appropriate for the Bayesian model prior. For instance, the variance of β_2, i.e. the parameter on changes in inventory, is calculated to be over 13 by the individual estimation results, which does not seem to make sense as the mean value is only 0.2; on the contrary, the random parameter model indicates the variance of β_2 to be 0.4 that is much smaller and is within a more reasonable range. Probably due to the lower magnitude of variance estimation for the parameters' distribution, the random parameter model in general suggests higher correlation (either positive or negative) between the parameters than that implied by individual estimations. With the distribution of the parameters as prior, it could proceed to the Bayesian forecasting model.

The Bayesian model is used to explore the heterogeneity of parameters across firms. It will take the form of:

$$\begin{aligned} CF_{i,t+1} &= \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t} \\ &\quad + \beta_{i,t,5}DA_{i,t} + \beta_{i,t,6}OTHER_{i,t} + \varepsilon_{i,t+1} \\ \boldsymbol{\beta}_{\mathbf{i,t}} &= \boldsymbol{\beta}_{\mathbf{i,t-1}} + \mathbf{V}_{\mathbf{i,t}} \end{aligned} \tag{4.4}$$

where the second equation means that the parameter vector follows a random walk process rather than remaining constant; **V** denotes the shocks to the parameters. Although all firms share the same prior, the parameters vector will be updated individually by new observations. Therefore it is expected that the parameters will be different for individual firms after several updating. The results of the random parameter model is adopted to be the prior for the Bayesian model, which is estimated using the in-sample data (1987-2005) in order to be consistent with the other methods. Table 4.4 (rightmost column) and Table 4.5 list the mean values and covariance matrix of the seven parameters, which are denoted as $\mathbf{b_0}$ and $\mathbf{VB_0}$ respectively in equation (3.25). Besides, the prior also include n_0 and S_0. The parameter n_0 is set to be 5, which indicates little confidence in the prior. The parameter S_0 is set to be 0.0035.

Apart from the priors, there are several coefficients to be set before the updating procedure. According to West and Harrison (1997), the coefficient δ_B in (3.37) for variance discounting is usually set between 0.8 and 1 and thus it is set to be 0.9 in this study. For monitoring procedure, the coefficient k in (3.42) is set to be 2.5 and δ_C in (3.45) is set to be 0.9. τ in step one of monitoring procedure is set to be 0.2.

Example: AAR Corp

AAR Corp's Global Company Key is 1004, which is positioned in the first of the sample and thus is used for demonstration. The trend of AAR Corp's net operating cash flow until 2005 is depicted in Figure 4.7 (denoted as 'y'). There is no clear trend in this company's cash flow series. After 2000 the data looks more volatile and there are two large negative observations in 2001 and 2005. Table 4.6 provides the posterior estimate of the parameters in each period before 2005. With the monitoring mechanism, the two negative observations are detected to be outliers and therefore excluded from the updating procedure. As a result, the estimates of parameters in these two years are the same as their prior prediction. In 2005, the parameters are updated to differ from the initial prior. However the signs of the parameters remain unaltered. The estimates of residual variance are also listed in the rightmost column of Table 4.6, which show an almost monotonically declining pattern. The results imply that the AAR Corp's cash flow series is actually not as volatile as indicated by the prior and thus the model lowers the estimates of residual variance gradually with observed data. Thus, an individual estimate of parameters that implies heterogeneity has been made possible through the Bayesian updating mechanism. It is noteworthy that the updating, which starts at the beginning, is not limited to sample sizes. However, data availability is still crucial as the estimates should be more accurate with more observations. The predictions made in each period are also depicted in Figure 4.7 (denoted as 'yhat'). Two large prediction errors appear due to the negative cash flows as all predictions are made to be positive. Due to the monitoring mechanism, the two errors are treated as outliers and do not enter the updating procedure.

If the monitoring mechanism is not used, the two negative observations will affect the posterior estimates in that particular year. The evolvement of parameters vector is listed in Table 4.7, which could be compared with Table 4.6. In 2001 and 2005, affected by the two negative observations, the posterior estimate of parameters deviate considerately from the prior prediction, i.e. the posterior estimate of previous period. In 2001, the posterior estimation of parameters β_1 and β_4 decline from 0.5071 and 0.5149 to 0.1874 and 0.0085 respectively; β_2 and β_3 even change signs. Similarly, there are dramatic changes in the posterior estimates of the parameters in 2005 as well. Besides the parameters on predictors, the residual variances after 2001 are generally higher than those in Table 7, which implies that the outliers increase the model's estimate of the cash flows' shock. As a result, after 2001, the evolution of the parameter estimates without the monitoring mechanism follows a very different trajectory from the one with monitoring. The monitoring mechanism in this example tends to avoid the dramatic jump of the parameter estimates due to outliers. The prior prediction of cash flow is also plotted in Figure 4.7 (denoted as 'no monitor'), which moves away from the 'yhat' series after 2001 due to the effect of the outliers, but within a short period it is hard to reach a conclusion how the monitoring mechanism affects the prediction performance.

To examine the effect of the initial prior, a different set of priors is applied in the Bayesian model to compare with the results by adopting the estimates of the random parameter model. The prior for comparison is made in a naive way by using the DKW theoretical model, i.e. the prediction of cash flow is equal to current earnings. The parameters on the predictors are initially set as: $\beta_1=1$, $\beta_2=1$, $\beta_3=-1$, $\beta_4=1$ and the rest of the parameters are set to 0, and their variances are all set to be 0.5 and covariances 0. The parameters n_0 and S_0 are set to 1 and 0.01 respectively. The new prior indicates higher uncertainty about the cash flow process than the prior applied before. The other coefficients remain at their default values as before. The resulted predictions are plotted in Figure 4.7 (denoted as naive prior). The predicted series are more volatile than using the prior from the random parameter model estimates. There are large prediction errors in 2001 and 2002, and the monitoring activity is triggered in 2004 instead of 2001. The observation in 2001 is not recognised as an outlier but the model is considered to deteriorate since then and the 'bad' performance accumulates until 2004 when the model starts to intervene. After the intervention, the observation in 2005 is not treated as an outlier either. The posterior estimates of parameters are shown in Table 4.8. The new prior has led to different estimates of the parameter in comparison with the previously used prior. However, the residual variances estimated with both priors are within a similar range.

In summary, the Bayesian model shows its main advantage in updating parameter estimate by each new observation. In the example shown above, it is noted that the estimated results for the parameters have a high dependence on the initial prior. The prior that adopts the results of the random parameter model seems more informative than the naive prior and the former results in a smoother prediction than the latter. Therefore, selection of the prior is crucial for the practical application of the Bayesian model. The monitoring mechanism tends to make a significant difference as well. The prediction performance of the Bayesian model will be further discussed in later sections.

4.4 Grey-box Models

From equations (4.2) to (4.4), the estimation of parameters gradually gets more complicated from being identical for different firms to having specific values for individual firms and further to changing values with each observation. The grey-box model which differs from the linear dynamic model discussed in the last section is implemented. The grey-box model does not assume that the parameters follow a linear random process but attempt to capture the parameters' dynamics and heterogeneity by a deterministic function of some certain variable that is considered relevant. Therefore, the grey-box model will be written in the form of:

$$\begin{aligned} CF_{i,t+1} &= \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t} \\ &\quad + \beta_{i,t,5}DA_{i,t} + \beta_{i,t,6}OTHER_{i,t} + \varepsilon_{i,t+1} \\ \boldsymbol{\beta}_{\mathbf{i,t}} &= \mathbf{F}(\mathrm{z}_t) \end{aligned} \tag{4.4}$$

Each parameter will be assumed to be a function of variable z and the function will be captured using a Padé approximant. In chapter 3, it is shown that the growth rate of firms' sales might have a deterministic explanation for the parameters' dynamics, therefore, the lagged growth rates of sales will be a candidate for variable z:

$$\boldsymbol{\beta}_{\mathbf{i,t}} = \mathbf{F}(r_{t-1}) \tag{4.4a}$$

The coefficients in the Padé approximant are estimated by minimising the sum squared error of prediction errors in the cash flow model (4.4) using the in-sample data. For each parameter, there are 5 coefficients to be determined (see chapter 3). The coefficients themselves do not have particular meanings but they are used to numerically replicate the functional forms of **F** that is unknown.

The functional form of **F** cannot be clearly written out, but it could be shown in graphic form. Figure 4.8 plots how the 7 parameters in (4.4) vary along with the lagged sales growth rates. The growth rates take the values from -1 that means an extreme scenario of sales dropping to zero to 1 that means that sales double. Sales could actually grow more than double and have no upper limit in theory, which however does not have frequent occurrence and thus not included in the chart. All seven parameters show nonlinear patterns on *r*. There is an interesting phenomenon in the charts that as growth rates of sales get higher, the effects of the predictors tend to decline except the AR term and thus the distance between the AR parameter and the others get greater. When the growth rate approaches 0, the parameters on cash flow, changes in accounts payable and changes in accounts receivable gradually get closer in absolute value. This pattern could imply that for mature firms that have relatively low growth rates, the gain of disaggregating earnings into components to predict cash flow will become smaller than those firms in growing stages.

There is one drawback in selecting lagged sales growth rate to determine the parameters' value in the cash flow prediction model: when one uses the model to predict cash flow more than one period ahead, the first thing is to predict sales growth rate, which brings in more complexity in the model. Firm age could be an alternative option proxy for growth rates to be used as input variable to model (4.4):

$$\boldsymbol{\beta}_{\mathbf{i,t}} = \mathbf{F}(AGE_{t-1}) \tag{4.4b}$$

Firms tend to have declining growth rates as time goes by, therefore, there would be a negative relationship between firms' ages and their growth rates. This assertion is supported by empirical observation. Using all sample data, mean growth rate of sales is calculated for each age, which is plotted in Figure 4.9. Firm age is calculated as the number of years ahead of that firm's first observation in the sample because this variable is not provided in the database. This calculation procedure will inevitably bring in bias and error. There is a clear declining trend of mean sales growth rate along with the growing firm ages. The growth rates gradually drop at a faster speed in the first 10 years. It is shown that after 20 years, the growth rates could still remain above 5 percent. After 40 years, the mean rates become spiky probably because of the small sample size. After 60 years, the trend of growth is not known. Age therefore seems appropriate as a proxy for growth to enter the grey-box model. Moreover, predicting firm age is simpler and more natural than growth rates.

The functions of equation (4.4b) are fitted in the same way as using the lagged growth rates, and the results are plotted in Figure 4.10, up to age of 100. The parameters do not change monotonically with age; they tend to reach their maximum or minimum value in the early ages and seem to approach some fixed levels after between 20 and 30 years. It implies that for firms that are older than 20 years, the grey-box model will not make much difference compared with the simple pooled regression model in making cash flow prediction.

4.5 In-sample Fitness to Data of Different Models

Sections 4.2 to 4.4 have introduced the estimation results of various models that are applied in this thesis. Their practical power will be examined by comparing their data fitting ability, especially out-of-sample performance. This section briefly shows the in-sample results of each model. The in-sample fitness is also examined based on the two measures, i.e. mean squared error and average rank.

In general, more complicated models and/or models with more parameters are considered to better fit the in-sample data and there is also risk of over-fitting. The most complicated model of this study is the grey-box model. However, the linear panel models have more parameters than the grey-box model to take account of individual effect of firms. Nonetheless, grey-box model has the advantage in making predictions for firms whose individual effect is not easy to calculate. For instance, consider a firm with only one observation. The models for comparison are the random walk model (Model 1), the theoretical DKW model that says the prediction of future cash flow is current cash flow plus the changes in working capital terms (Model 2), the BCN model estimated by pooled regression (Model 3), the panel model that assumes homogeneous and constant parameters except the intercept term estimated using demean (Model 4), difference (Model 5) and Arellano-Bond estimators (Model 6), the Bayesian model with naive prior (Model 7) and the prior from random parameter model (Model 8), the grey-box model using sales growth rates (Model 9) and firm age (Model 10) as additional input variables. The Bayesian model which differ from others, is not estimated frequently. Therefore, the concept of in-sample fitness of the Bayesian model does not mean much. However, it is expected that Bayesian model works better as more observation enter the updating procedure. Therefore, the performance of Bayesian model at the end of the in-sample period, i.e. year 2005, would be examined in comparison with the others.

Figure 4.11 plots the fitted value (yhat) and the actual cash flow series (y) of the in-sample period using AAR Corp's data as an example. In this particular example, Model 5 seems to best fit the in-sample data, especially in the period after 2000 when the series becomes more volatile. Model 2 provides the most volatile predictions and Model 7 that uses Model 2 as prior has similar performance. Differences are not apparent for the remaining models. The MSE and average ranks are calculated for each model and summarised in Table 4.9. The results are based on 62927 firm-year observations. The second row lists the resulting mean squared errors (MSE) and the numbers in the third row are the average ranks of each model. For both measures, a smaller number indicates better performance. The panel models 4 and 6 have produced lowest MSE as they calculate individual effects for each firm. Model 5, however, has obtained a higher MSE than model 4 and 6. Recall that the AR parameter estimated by model 5 is negative due to the bias brought in by taking the first difference of variables, and the in-sample results suggest that such biased estimation may not properly make practical predictions. Despite that the panel models 4, and 6 have lower MSE, their average ranks are among the highest tier, only lower than model 5. Therefore, it can be said that the panel models are inferior in general. Another point is that model 6 using the Arellano-Bond estimator that is considered consistent does not outperform model 4. Model 3 assumes total homogeneity even for the intercept term and is estimated simply by pooled regression. Although it has higher MSE than the panel models, the lower average rank of model 3 indicates more general description of the cash flow process, which is inconsistent with the expectation from an econometric angle because the results estimated by pooled regression without considering individual effects are likely to be biased. Model 1 and 2 predict cash flow in a naive way and are thus selected to be benchmark models. Their MSE are large compared with the others but their average ranks are lower than models 3, 4, 5 and 6 that apply optimisation methods in their estimation procedure. It is noteworthy that model 2 has poorer in-sample data fitting than model 1 for both criteria, which is against the assertion in the DKW paper that including accrual terms will make better prediction than the random walk model. The random walk model has very low average rank, second only to model 10, demonstrating its informative summary of the cash flow process. The two grey-box models, 9 and 10, fit the data very well according to both measures. The MSE of the two models are comparable with that of pooled regression, lower than the models with individual effect. However, grey-box model with firm age as the black-box input has the lowest average rank of all 10 models. Model 9 has the third lowest average rank, only higher than model 1 and 10.

Bayesian models have a medium performance compared with the other models. Model 8 with prior from the random parameter model works better than model 7 that uses the uninformative naive prior. Model 8 is ranked 4th according to the average rank criterion, showing a better performance than the linear static parameter models. Its MSE is smaller than model 1 and 2 but higher than the others, given that the Bayesian model needs no optimisation procedure for parameter estimation. Further comparisons year by year are also conducted (results are not listed) to examine the hypothesis made previously that the Bayesian model will improve with more data observed. However, there is no sign of improvement in 2005 for Bayesian model compared with the periods before then.

In summary, the grey-box models generally are the best form of model to fit the data in-sample comparing with other options. Firm age as the black box input variable seems to work better than sales growth rates. The random walk model, though producing high prediction error, better describes the cash flow process than some parameterised models. To gain deeper knowledge of the models' performance, out-of-sample data need to be used, which will be the content of next section.

4.6 Out-of-sample Prediction Performance of Different Models

In the last section, the in-sample fitness of the 10 models are examined using two criteria, i.e. MSE and average rank. As Table 4.9 shows, the grey-box model provides encouraging performance. Not only the MSE of the two grey-box models are as low as that of pooled regression, their average ranks are also among the best models of all. For practical prediction, out-of-sample examination is more important. The models perform well in-sample may not necessary extend their superiority to the out-of-sample period. In this particular application, i.e. cash flow prediction, one-period-ahead and multi-period-ahead predictions are both important and useful, therefore, this section will test the two types of predictions separately.

4.6.1 One-period-ahead Prediction

The data after 2005 is used for out-of-sample test purpose. For panel models, i.e. model 4, 5 and 6, the application of them in the out-of-sample period required that the individual intercept values for the target firm are available, which will exclude observations of firms that do not appear before 2005. For Bayesian models, the parameters vector that is last updated for each firm in-sample, i.e. in 2005, are applied out-of-sample. Figure 4.12 plots the one-period-ahead predictions for AAR Corp from 2006 to 2011 by the 10 models. In this particular example, model 2 seems to outperform the other models. Panel models 4, 5 and 6, do not perform as well as model 3, although the latter does not calculate individual effect. Bayesian models 7 and 8 have also generated large prediction errors, only marginally better than model 5. Both grey-box models have made good prediction for this firm in this period, with comparable performance with the pooled regression model 3.

The MSE and average rank for the whole out-of-sample data are calculated and listed in Table 4.10. The calculations are based on 17965 firm-year observations. For each criterion, the best two models are labelled by bold numbers. Surprisingly, the panel models 4, 5 and 6, despite their consideration of individual effects have higher in-sample MSE. They also have the poorest performance of all 10 models in the one-period-ahead forecast based on average rank. Model 5 has the poorest prediction. Model 4 has the least MSE among the three panel models but shows the highest average rank of all ten models. Model 6 that applies Arellano-Bond estimator does not appear to be a good predictive model according to both measures. Grey-box models prove their power in this comparison, especially for model 10 which has both the lowest MSE and the lowest average rank. Model 9 has the second lowest MSE and its average rank is in the middle position of the ten models. The model that has the second best average rank is the benchmark model 1, which has the same level of MSE as model 2. Model 3 and the two Bayesian models have provided medium performance, better than the panel models but not as good as the grey-box models. In conclusion, Grey-box model 10 is the best model in both criteria for the one-period out-of-sample competition.

4.6.2 Multi-period-ahead Prediction

Predictions beyond one period are made recursively by extending the predictions of cash flow to all the predictive variables and thus use the predicted variables to make further periods' predictions. Therefore, the work load of making multi-period-ahead prediction is over 6 times, because there are 6 predictors for the one-period-ahead case. For a given period in the future, there will be predicted series for each independent variable.

In this study, the other variables apart from cash flow are predicted using the same model, i.e. the same independent variables, model structure and estimation methods. Model 1 and 2 are exceptions as their multi-period-ahead predictions will simply be the last in-sample observation of cash flow or cash flow plus changes in working capital terms. Grey-box model 9 is not suitable for the multi-period task because the input variable, i.e. sales growth rates, requires more prediction, which adds extra complexity to the model. Therefore, only grey-box 10 can be utilised. Besides, Bayesian model 7 was used to show in contrast how an informative prior is useful for Bayesian prediction. For multi-period tasks, only Bayesian model 8 with the random parameter model results as prior will be compared with other models. Hence, there will be 8 models in total to compete in multi-period-ahead prediction using the out-of-sample data.

The parameters used for prediction are the in-sample estimated results and the predictor variables take their value in year 2005. Therefore, the out-of-sample prediction for each firm-year observation requires that that firm has available data in 2005. Figure 4.13 plots the comparison of actually observed cash flow for the AAR Corp from 2006 to 2011 and the cash flow predictions all of which are made using information up to 2005. It is interesting to see that model 3 and model 10 have produced an upward trended prediction for this firm's future cash flows, which made them better fit the firm's actual observations. The panel models 4, 5 and 6 have rather flat prediction pattern and the Bayesian model makes a predicted trend of irregular shape. Table 4.11 provides the results of the models' performance in the multi-period-ahead setting. The predictions are compared with the sample data from 2006 to 2012, where the models' performance in the predictions of one to up to seven years ahead could be examined. The results favor the grey-box model as it outperforms all other models in both criteria. The model with the worst performance is the Bayesian model, which underperforms all other models in both criteria. The panel models also generate higher MSE than simpler models 1, 2 and 3 and the more complicated grey-box model. It thus seems that considering individual effects does not help in making prediction in practice, which is a counterintuitive conclusion. Model 3 has the second lowest MSE and model 1 has the second lowest average rank, both of which are simple enough in estimation but provide encouraging results in real application.

To be strict, predictions for cash flows in year 2006 do not belong to the category of multi-period and thus another comparison is made by excluding the observations in 2006, which are listed in the 3rd and 4th rows in Table 4.11. Grey-box model still performs best in both criteria. The conclusions in general do not change much for the sub-period only except that model 2 outperform model 1 when data in 2006 is excluded, which suggests that the DKW assertion that earnings make better prediction for cash flow than cash flow *per se* is more descriptive in the long run.

4.7 Equity Valuation with Cash Flow Prediction

Sections 4.5 and 4.6 have shown the superior performance of grey-box model in predicting future cash flow from different angles. The grey-box model fits the in-sample data well and moreover does not appear to have over-fitting problem as its out-of-sample prediction performances of both one and multiple periods ahead are evidently better than other competitive models including simple benchmark models. To further exploit the economic value of cash flow prediction, this section will link the predictions of cash flow and the corresponding firms' equity value by a novel way of applying the discounted cash flow (DCF) method as introduced in section 3.7.

Although the grey-box model outperforms other candidate models selected in this thesis in predicting future cash flow of firms, it may not be necessarily true that the cash flow predictions made by grey-box model could result in better equity pricing. The equity price reflects market participants' expectation of the firms' future (cash flow) income and how the market expects is actually an unknown and unanswerable question. It is however expected that a better cash flow prediction model will be closer to the market expectation as long as the market is sufficiently efficient. In this section, the theoretical DKW model for cash flow, i.e. model 2 in previous two sections, will serve as the benchmark model, with which the Bayesian model and the grey-box model, i.e. model 8 and model 10 respectively in previous two sections, are compared.

The grey-box model will apply the in-sample estimated coefficients and thus the equity prices after (including) 2005 are examined. The benchmark model, Bayesian model and the grey-box model are not limited to available observations to make point predictions, which makes them easier to apply to a wider range of sample firms than the individual effects models. Besides, the Bayesian model provides posterior estimates of shock terms' variance during the updating, which will be used in the DCF procedure, specifically in the simulation stage. However, the benchmark model and grey-box model do not have this feature and their estimation accuracy of the shock variance will depend much on the number of observations. For the two models, the variance of prediction error will be simply calculated as:

$$S_{i,t} = \frac{\sum_{n=1}^{t} \varepsilon_{i,n}^2}{t-1} \tag{4.5}$$

where S denotes the estimate of variance of prediction errors ε. t indicates the time of making the estimation and it is shown in (4.5) that when only one observation is available for a firm, the estimation of prediction uncertainty will be unavailable. Bayesian model, provided with prior information, is not restricted in this situation.

4.7.1 Simulation of Cash Flow Series

Simulation is made by generating a large amount of random numbers that follow a specified distribution. At time t, we generate random numbers of normal distribution of zero mean and variance $S_{i,t}$. Again take AAR Corp as an example, until 2005 there are 17 observations. According to equation (4.5), the prediction error by model 2 has a variance of 0.00336 while for model 10 it is 0.00144. The Bayesian model provides the posterior estimate of 0.00128 for the shock term variance in 2005. These numbers are calculated on the average-total-assets-deflated data and for the purpose of stock pricing, the numbers need to be adjusted to per share values in this section. The variance calculated above is therefore adjusted by:

$$S^{*}_{i,t} = S_{i,t}(\frac{ATA_i}{SO_{i,t}})^2 \tag{4.6}$$

where ATA_i denotes average total asset of firm i and $SO_{i,t}$ denotes the number of shares outstanding of firm i at t. After the adjustment, the variances of prediction errors are 1.66 by model 2, 0.63 by model 8 and 0.71 by model 10 and the cash flow predictions the three models make for 2006 are 0.46 per share, 0.44 per share and 0.74 per share respectively.

For the DKW model, the simulation of cash flow series is easier because there is no need to predict the independent variables of the cash flow model. All future cash flow predictions are equal to the current cash flow plus changes in working capital terms and thus, after the first period, the prediction is equivalent to a random walk. In the simulation, it is necessary to set up a quit rule, which aims to replicate the asymmetric feature of cash flow distribution. One option of the quit rule could be that if the simulated series encounter some certain amount of, e.g. 3, negative cash flows one after one, the simulated series will be assumed to break up and all future simulation after then are zero. With the numbers in previous paragraph, model 2 could be easily simulated under the quit rules. The total number of simulated series is 10,000 and the maximum period is set to be 100. Figure 4.14 depicts the shape of simulated cash flow series. The two charts are produced by two rules that only differ in the number of negative cash flows allowed to appear one after another. The upper chart allows 3 (rule 1) and the lower allows 10 (rule 2) implying higher tolerance. Rule 2 seems unreasonable as the lower bound of the cash flow observations could reach over -10 per share and thus rule 1 will be used.

Unlike the DKW model, the Bayesian model and the grey-box model both need to predict the independent variables along with cash flow for making multi-periods-ahead prediction. The simulations by the Bayesian model and the grey-box model are plotted in Figure 4.15, along with the one generated by the DKW model. The two more advanced models have resulted in different shapes of the simulated cash flow series. The Bayesian model simulated observations have a tendency to converge to zero, which is not reasonable in the practical world. The grey-box model simulated data shows an almost constant range after several periods. The point predictions made for all six relevant variables by the grey-box model are plotted in Figure 4.16. The first sub-chart in Figure 4.16 is the multi-period-ahead predictions of cash flow. There is an upward trend for the prediction, which seems to converge to about $3 per share. The dashed line is the mean of the simulated series in Figure 4.15 lower chart. Clearly, the two lines do not converge, which indicates the influence of the simulation procedure, especially the quit rule.

4.7.2 Calculate Certainty Equivalent Cash Flows

In each period, the simulated series form a distribution of cash flows, replicating the possible observations in that particular period. In this stage, utility function equation (3.48) is applied to identify the certainty equivalent cash flow with regard of the distribution. The utility function is only controlled by the RRA coefficient ρ. Figure 4.17 depicts the chart of utility function when ρ takes values of 0, 0.25, 0.5 and 0.75. Zero coefficient means risk neutral and the chart is a straight line. For positive coefficients, the charts are concave curves, suggesting risk aversion.

In period 1, the simulated data for AAR Corp by DKW model follows almost normal distribution with mean equal to 0.4783. The utility of negative observations are all assigned to zero as the utility function cannot deal well with negative numbers. With ρ as 0.75, the certainty equivalent cash flow (CECF) in the first period is 0.1656, much lower than the mean cash flow. The series of CECF for the whole 100 periods are plotted in Figure 4.18. Although the simulated data has an upward trend (see Figure 4.15, upper chart), the CECF gradually declines after some periods and converges to some slightly positive value. This feature will promise that the theoretical stock price obtained by DCF methods converges. The CECF series by the Bayesian model and the grey-box model are plotted in Figure 4.19. It is noted that the grey-box model has a very slowly converging CECF series, which is due to the grey-box model's lower uncertainty in the simulation. Whether the speed of converging will affect the stock pricing procedure will be further discussed next.

4.7.3 Risk Free Rate Discounting

With the calculated CECF, the DCF method could be used to determine theoretical prices. The discount rate will be risk free rate which is observable and publicly available. Treasury bond rates are used as risk-free rates and they are available for 1, 2, 3, 5, 7, 10, 20 and 30 years. For the simulation series, there is a need for up to 100 years of interest rates, which will be interpolated and extrapolated using the method by Nelson and Siegel (1987). The disclosure date for AAR Corp's information of year 2005 is 31 May, 2006, and the corresponding yield curve is shown in Figure 4.20. The fitted yield curve is used to discount the CECF series and the discounted cash flows are summed to reach the firm's theoretical value. For example, using AAR Corp's data, the theoretical price is computed. By the DKW model, the theoretical price is 1.7; by the Bayesian model, it is 3.63 and by grey-box model, it is 32.63. The accumulated discounted cash flows by each period up to 100 years are plotted in Figure 4.21 to check the converging feature of each model. It takes about 60 years for the DKW model to reach a steady theoretical value by summing the discounted cash flows whereas the theoretical value by the Bayesian model converges in only about 30 years. However, the grey-box model theoretical value converges very slowly, which is consistent with the slow convergence of the CECF. It seems that the theoretical price based on the grey-box model converges over a longer period, over 100 years. Therefore, 32.63 might slightly be an underestimate of the theoretical value of the target firm's stock using grey-box model. As long as the CECF's growth rates in its steady state do not exceed the discount rates, the theoretical price of DCF will definitely converge. If a longer period of data is simulated, the resulting price will be more accurate.

4.7.4 Calibrate the Risk Aversion Coefficient

The actual market price on the data date[2] is 24.08, which implies that it is not accurate to assume risk aversion coefficient to be 0.75 for any model. The DKW model and the Bayesian model both produce much lower prices than the actual one; hence the market must be less risk averse if the market expectation of future cash flow is captured by the two models. For the DKW model, to let the theoretical value of stock equal the real market

[2] Data date is the item available in WRDS Compustat database.

price, the RRA coefficient needs to be 0.1891 and for the Bayesian model, it needs to be -3.116 which implies a risk seeking utility function. For the grey-box model, it should be the other way around: the market of which the expectation of cash flow is captured by the grey-box model needs to be more risk averse to price the stock at its market value. The calibrated RRA coefficient will be 0.8701 for the grey-box model.

Different individuals are likely to have various degrees of risk aversion, but the market as a whole need to be consistent in the attitude to risk. Therefore, it is reasonable to assume that the market only has a uniform RRA coefficient at one particular time point and all firms' stocks are priced according to this coefficient. The RRA coefficient is calibrated to meet condition (3.55). The rationale is that in the whole market the mispriced proportion on average is zero. From the January of 2006 to the December of 2012, there will be an RRA coefficient calibrated in each month on the firms making disclosure. Most firms, about 2000 of them, disclose data in December and March, June and September follows (roughly 200 in each). The calibration of RRA coefficients will be based on these firms' theoretical value and their market price. The RRA coefficients series based on DKW cash flow prediction is plotted in Figure 4.22. There are some months when the calibration fails as the resulted RRA coefficients are out of reasonable range. In these cases, the RRA coefficients are set to equal previous month's value. The failure months include January of 2006 which is thus excluded from the series. The RRA coefficients are presented along with the contemporary S&P index to see the relationship between the degree of market risk aversion and market price. Before 2010 there was an obvious negative relationship between the market index and the RRA coefficient. The lowest RRA coefficient appears during the peak of the S&P index, which corresponds to a period of least risk aversion for the market participants in general. Similarly, when the market index reached its trough, the RRA coefficient is accordingly very high. After 2010, although there is still a negative relationship between the two series, the RRA coefficients only move within a smaller range and thus the association between the series is no longer obvious. The RRA coefficient could be seen as a measure of market emotion. A high RRA implies that the market is valuing stocks in a very careful manner. During such a period, stocks may be relatively cheap, or in another word, their value may be less likely to experience dramatic decline. On the other hand, when the RRA coefficient is low, there is potential danger to enter the market. In Figure 4.22, it seems that 0.8 is approximately a benchmark value for the RRA coefficient.

Figure 4.23 plots the RRA coefficient series that are calibrated by grey-box model. Similar to Figure 4.23, the RRA coefficient reaches the lowest value in the June of 2007. When year 2008 comes, the RRA coefficients stay high and move only within a narrow range (between 0.8 and 1). Comparing it with Figure 4.22, the grey-box-resulted RRA coefficients are less volatile than that calibrated by DKW model. Besides, the former has shown a greater distance between the minimum and maximum values, which makes it easier to signal a market with lower risk aversion.

4.7.5 Exploit the Market Mispricing

Once the RRA coefficient is identified at a particular time, all firms in the market could then be priced accordingly. In the calibration procedure, equation (3.55) is applied as a condition to be met. As a result, there will be firms for which the theoretical value is greater than the market price. These firms will be seen as under-priced and their prices will be expected to rise in the future and vice versa. Table 4.12 summarises the numbers of firms for which the share prices are expected to move upwards or downwards. During the whole period, the DKW model makes an upward prediction for 5533 firms and a downward prediction for 10834 firms. Similarly, the valuation results on grey-box-model-predicted cash flows suggest that there are 5702 stocks undervalued and 10665 overvalued. There are much more overvalued stocks implied by the valuation method. This is because that there is a lower limit for the expected return of overpriced shares, i.e. -100% which is a loss of one's entire investment, There is however no upper limit for stocks considered under-priced. Therefore, to meet condition (5.1), the proportion of under-priced stocks is much lower. Table 4.12 indicates that there is only about 50-50 chance to make correct predictions, which is not encouraging. Pure random selection will have similar results as the total numbers of rising (8355) and declining (8012) stocks are roughly equal. However, it is noteworthy that the results are limited to the low frequency of data. The price movement direction is calculated on annual basis. The stock prices might have moved in the right direction at some point during the year but reverse at the end, which will dilute the result.

Market expectation of firms' future cash flows is unknown and unobservable. It is a concept that may not even exist. Market participants only know their own expectation of cash flows. The market only shows prices and volumes. Therefore, a cash flow prediction model that generates better matching of the market prices and theoretical prices should be considered to be closer to the market expectation. The variance of mispriced proportion, or say expected return calculated by equation (3.54), of all stocks in the market is used to measure the fitness of a cash flow model to unobservable market expectation. Lower variance indicates a better fit. In the most extreme case where a model could generate a variance of 0, the model will perfectly describe the market expectation because all stock prices are then equal to their theoretical value based on the cash flow expectations. The resulting variance by the DKW model is about 7×10^9 and that by grey-box model is about 6×10^7, which indicates that cash flow predictions made by grey-box model better describes market expectation than DKW model.

With the identification of under-priced stocks, it is examined whether excess profit could be made out of it. In every month there are firms making disclosure. The investment strategy is to buy the under-priced stocks among the firms making disclosure, hold them for a year and sell them in the same month of next year, when there will be newly selected under-priced shares according to the updated information. Table 4.13 lists the number of target firms (under-priced) selected by the DKW model and the one-year average return of them in each month from February 2006 to December 2011. Most firms make disclosure in December and thus there are much more targets in December than in other months. The numbers in bold indicate the returns that are higher than the S & P index annual return at the same month. Out of 71 months, there are 48 months in which the selected stocks outperform the market index on annual return. There are losses mainly concentrating in 2007 and the first three quarters of 2008. 12 portfolios are constructed corresponding to each month of year. Each portfolio only considers the firms making disclosure in the specific month and the portfolios are rebalanced annually. Table 4.14 shows the payoff of the 12 portfolio assuming the initial capital is 1. Simultaneously, the S & P index of the same period is normalised to have an initial value of 1 and is compared with the portfolio performance. Bold numbers indicate the time points when the portfolios have higher value than the index. The rightmost column reports the final payoffs of the portfolios and the S & P index at the end of 2012. The portfolio providing the highest return is the March portfolio, followed by the January and November portfolios, all of which more than double in value in six years. The portfolio with the least final payoff is the June portfolio, which has only generated a gross return of 6% and is the only portfolio that does not beat the contemporary market index. All portfolios have grown in value. The index is roughly on the same level in 2012 as in 2006 whereas the average final payoff of the 12 portfolios is 1.698, implying a gross return of nearly 70%. It is a promising performance given the low frequency of transactions. Even if transaction costs were considered, the strategy could still outperform the market index.

The same procedure is conducted again using the grey-box model to make cash flow predictions. The results is summarised in Table 4.15 and 4.16. Table 4.15 reports the number of target firms in each month during the period and the average one-year return of the target firms. There are 42 out of 71 months in which the selected stocks outperform the contemporary S & P index in one-year return. The proportion is smaller than that using DKW-model-predicted cash flow. Table 4.16 lists the payoff of the 12 portfolios, each with initial investment of 1. The two portfolios of stocks with disclosure date in February and November have underperformed the S & P index in 6 years, both of which have negative gross return. Portfolio 2 ends up with 0.7909 and portfolio 11 ends up with 0.7443. Over twenty percent of the initial investment is lost. Portfolios 7 and 9 have leading performance as their value are more than doubled. The average final payoff of the 12 portfolios is 1.334, which is still higher the performance of S & P index.

Although the grey-box model could better predict cash flow and better fit the market price data, the payoff by the basic strategy is worse than that when the same strategy is applied along with DKW mode instead. To examine whether a better cash flow prediction model could create economic profit, a second strategy is conducted. First use the DKW model to pick out the undervalued stocks and invest in the stocks which also meet another condition: the grey-box model generates higher one-year-ahead cash flow prediction than the DKW model. The results are summarised in Table 4.17 and 4.18. Table 4.17 lists the target firms by the two criteria and their one-year return. There are far fewer target firms due to the additional condition. Even in December, no more than 100 firms could pass the criteria. In some months there is no target firm at all. There are 57 months when investment opportunities exist. In 32 months the selected stocks have higher one-year returns than the S & P index. The proportion is lower than 60 percent. However, there are occasions where the selected stocks have very high returns. For instance, in February of 2009, only 4 stocks are considered as targets but their average one-year return exceeds 160%. There are 4 more months when the selected stocks' average returns are higher than 100%. Table 4.18 reports the performance of each portfolio when the new strategy is applied. There are only 8 portfolios providing higher final payoffs than the contemporary S & P index. However, the average final payoff is 1.96, which is higher than when the basic strategy is applied, no matter whether cash flows are predicted by the DKW model or the grey-box model. This is because the new strategy has resulted in extremely high returns for the leading portfolios, which compensate for the poorly performing portfolios. Portfolio 1 has a final payoff of 4.98, almost 5 times of the initial investment. Portfolio 3's final payoff is 4.5 and portfolio 4 has grown from 1 to 3.07. The best portfolio by the basic strategy provides no more than a 2.5 final payoff while the new strategy has made five portfolios more than double in value during the test period. In summary, the portfolios constructed on the new strategy have more disperse performances but on average higher return than the basic strategy. The new strategy depends on the better performance of grey-box model in making cash flow prediction and the results suggest that the combinative strategy is good at discovering undervalued stocks with high potential. Grey-box model thus has incremental power to improve investment decisions making.

4.8 Conclusion

This chapter conducts empirical analysis on U.S. data. First, in an attempt to understand the pattern of firms' cash flow, the cash flow observations of all firms are normalised to start from the same level. The results show that the realised paths of cash flow series disperse. The results also suggest that some firms' cash flows could grow to over 250 times their initial cash flow income over a 50 year period with some firms having negative cash flows. Therefore, it is difficult to reach a general conclusion about the cash flows' growth pattern. On average, cash flow has a growing trend but such a trend may be influenced by survivorship bias. Roughly 15% of sampled firms survived for 50 years and the average trend of cash flows is calculated on the observation of survival firms, therefore the raw conclusion ignoring such bias will overestimate U.S. firms cash flow' growth. Cash flows adjusted by multiplying by survival rate show that the average cash flow series still has an upward trend but with much lower growth rates than that suggested by the original data. It can be concluded that, on average, the U.S. firms' cash flow grows on an annual rate of 5 percent.

In addition, cash flow of sampled firms has an asymmetric distribution: there seems to be a lower bound but no upper limit for cash flow observations. This could be the results of artificial intervention, for instance, bankruptcy, enhancement of management and so on, which will be taken to avoid the occurrence of continuous large losses. This phenomenon is very important for the study in Section 4.7 as a quit rule is first added to the simulation procedure to replicate the asymmetry of cash flow distributions.

After the observation of the cash flow pattern, the main part of this chapter focuses on various models' performance in predicting cash flow. To conduct a comprehensive comparison, both in-sample and out-of-sample performances are examined, the latter of which includes not only the one-period-ahead prediction but also multi-period-ahead predictions. There are ten models that are examined in the study. The first three of the examined models have been applied in previous studies and hence could be considered as benchmark models. The rest of the models are implemented for the first time in a practical cash flow prediction study with panel data. The models used in the study could be summarised into three categories: panel data models that consider individual effects, Bayesian dynamic linear models and grey-box models. The application of the grey-box model is a major contribution of this thesis to the extant literature. The model performs well both in-sample and out-of-sample, especially in multi-period-ahead prediction competition where it has the best performance among all the candidate models.

The second contribution of this study is the examination of the novel discounted cash flow (DCF) method in pricing equities. Compared with the traditional way of applying DCF method, the one developed in this thesis is very different. First the cash flow to be discounted is not the raw predicted values but the certainty equivalent cash flow (CECF) transformed from the original predicted cash flow distribution by a certain form of utility function assumed to show constant relative risk aversion (CRRA). Simulation is conducted to implement the procedure. Secondly, the method applied in the thesis uses the risk-free interest rate that is easy to obtain to calculate discount factors. The novelty of this method lies in the incorporation of cash flow prediction including both the point prediction and the uncertainty of predictions into the framework of the DCF method in stock pricing. The portfolio performance by the new method has shown encouraging economic value despite the fact that it is a preliminary study. Although the results of predicted directions of price movements are not impressive, it is still too early to draw a negative conclusion before making more detailed studies.

Chapter 5 Empirical Study II: U.K. Listed Firms

Data for U.K. listed firms for this analysis is obtained from the Datastream database. The accounting information is collected from annual reports which span the period from 1989 to 2015. All sampled firms are traded on the London Stock Exchange. Similar to the U.S. sample, all financial firms in sectors of banking, investment or insurance are excluded. Usually financial firms do not follow similar operating procedures as firms in other sectors, e.g. trading, manufacturing etc. The cash flow models may thus not suitable for them. There are initially 1272 firms that have available observations for cash flow from operating activities.

The data for empirical study on U.S. firms suggests that a firm's age is a crucial variable in the analysis. First, age is used to plot cash flow trend and it is also used as an input variable into the grey-box model for cash flow prediction. The age of U.K. firms is calculated from the date of incorporation of that particular firm up to the year of data availability from the Financial Analysis Made Easy (FAME) database. After exclusion of firms that have unavailable or inconsistent ticker symbols recorded in the two databases, firm ages could be calculated for 1009 firms.

5.1 Pattern of Net Cash Flow to Total Assets

Figure 5.1 plots the mean, median and 2.5 to 97.5 percentiles of the distribution of net cash flow from operating activities divided by total assets, which are sampled by firm ages from 1 to 110 years. There are too few firms (less than 30) that are older than 110 years and thus the limit is set. The cash flows for sampled firms are not normalised as was the case for USA firms. The reason for the difference is because that firm ages are directly available for U.K. data whereas the ones used in the U.S. date analysis are proxy ages that actually indicate the number of years since a firm first earned a positive cash flow. In the case of U.K. data, there are no cash flow observations for firms that were founded earlier than 1987 when cash flow disclosure became a requirement. However, with exact firm age the observed pattern of cash flow to total assets can describe the cash flow behaviour from another angle. In the previous analysis of U.S.A. data, it is concluded that firms' cash flow tend to follow an ever-growing pattern as time evolves. The conclusion is however based on a biased sample of firm-year observations for firms that only report positive profit. The initial negative observations are excluded to make the normalisation procedure work. In this chapter, the negative cash flows are instead taken into account. The ratio of observations over 1 or below -1 are excluded, which do not occur frequently but severely affect the scale of the figure. Figure 5.1 shows a pattern that indicate that cash flows for firms at their beginning ages are averagely negative, or in another words, there is a net cash outflow. The median series also show the same pattern but there is a relatively big gap between the mean and median series in the beginning periods, implying heavily asymmetric distributions during firms' starting stage. The observed pattern, i.e. negative net cash flow to total assets in the first few years (age 1 through 9 for the mean series and age 1, 2 and 4 for the median), suggests that in the beginning periods there are large amounts of expenditure that need to be made in cash and/or the earned cash income is not as high and cannot cover the expenditure. Averagely the sample firms start to make positive cash income from age 10. The net cash flow to total assets ratio grows almost monotonically until age 20 when the ratio remains steady at about 8% thereafter. Over 20 years the two series start to get close and the ratio distribution thus becomes more symmetric. The 95% range shown by the dotted lines suggests that for most firms their annual net operating cash flows are no higher than 30% of the firms' total assets. If a simple world is assumed where firms' total assets increase exactly by the amount of net cash flow year after year, the cash/assets ratio will be equivalent to a firm's growth rate. Comparing the U.K. with the U.S.A. results, the two countries experience similar level of growth rate for survival firms, recalling that the mean normalised cash flow series of U.S. observations also have an annual growth rate of about 10% (Figure 4.2).

5.2 Cash Flow Prediction: Model Estimation

Cash flow prediction is based on model (3.11), which is different from (4.2) in that (3.11) uses depreciation and amortisation as two separate predictive variables therefore has one more parameter than (4.2). The difference is mainly due to variable availability for the different databases used in both studies.

Similar to the U.S.A. analysis, models placing different assumptions or specifications on the parameters are selected, which include the random walk model (Model 1), the theoretical DKW model that suggests the prediction of future cash flow is current cash flow plus the changes in working capital terms (Model 2), the BCN model estimated by pooled regression (Model 3), the panel model that assumes homogeneous and constant parameters except the intercept term estimated using demean (Model 4) and Arellano-Bond estimators (Model 5), the Bayesian model with prior from random parameter model (Model 6), the grey-box model using sales growth rates (Model 7) and firm age (Model 8) as additional input variables. Observations up to year 2006 are used for in-sample estimation and observations after 2006 are used for out-of-sample performance comparison.

5.2.1 Estimation Results of Panel Data Models

Before applying the cash flow models, all the variables are deflated by average total assets. The cash flow observations in the highest and lowest 1 percentile of the entire sample are excluded, resulting in a sample of 11837 firm-year observations. Table 5.1 lists the descriptive statistics for the deflated variables. The mean of deflated cash flow is 0.0291, only half of that of U.S. data, i.e. 0.0683 (see Table 4.1). However, the median is much higher, which reflects the high asymmetry of the distribution. The parameters estimated using panel models, including pooled regression, are listed in Table 5.2. Pooled regression and demean estimation are conducted on 4480 observations. The Arellano-Bond estimators are obtained by taking the first differences of the variables and thus are estimated on 3675 observations. Column 2 of Table 5.2 reports the results of pooled regression. The parameters on amortisation (β_6) and other components (β_7) in earnings are not significant either in statistical or economic sense. This is different from the observed result in U.S. data. Unlike amortisation, depreciation and depletion (β_5) is highly significant and positive. All the other variables are both statistically and economically significant and the signs of their parameters are consistent with the theoretical expectation. Columns 3 and 4 report the estimated results of demean estimation and the Arellano-Bond methods respectively. There is no substantial difference between the results by the two different methods, except that the Arellano-Bond estimator suggests that change in inventory (β_2) is not useful for predicting cash flow. Both methods support that variables of amortisation (β_6) and other components of earnings (β_7) are not predictive for cash flow. In general, panel models lead to lower parameters values than the pooled regression, the only exception being that on depreciation and depletion (β_5).

The results of the random parameter model are listed in Table 5.3. The estimated parameter distributions are represented by their mean values, variances and covariances with each other. The estimation is conducted on the in-sample data and the results are used as prior for Bayesian model. Ignoring the intercept term, the AR parameter, i.e. β_1, has the least dispersed distribution among all the parameters. It has a mean of 0.7453 and a standard deviation of 0.3715. On the contrary, the distribution of the parameter on amortisation has the largest variance, i.e. 1.0199 while the mean of the parameter is -0.2386.

5.2.2 Estimation of the Grey-box Models' Parameters

The grey-box model with sales growth rates in the black-box is estimated using the in-sample data. Figure 5.2 plots the functional association of the parameters on each variable and the growth rates of sales. In the plot, the growth rates of sales are restricted between -1 and 1. None of the 8 parameters has a monotonic pattern. Parameters on amortisation and other components of earnings, i.e. β_6 and β_7 respectively, are close to zero when the sales growth rate is positive. β_6 and β_7 reach their maximum values when the sales growth rate is around 0 and 20% respectively, and their maximum values are still very small in magnitude. Therefore, these two variables seem to have negligible impact on future cash flows, which is consistent with the panel model estimation results. The patterns of the parameters on other accrual terms suggest that when firms experience high growth in their annual sales, the cash flow tends to have higher dependence on the accrual terms in the previous year. The AR parameter stays above 0.8 with positive growth rates of sales. Comparing Figure 5.2 with Figure 4.8, the pattern of parameters on changes in working capital components, i.e. inventory, accounts payable and accounts receivable, are very different between the two countries. For instance, looking at the general trends, the U.K. result shows that the effect of change in inventory on future cash flow is positively related with sales growth rate whereas it is the opposite in the U.S. The effect of change in account receivable reaches its maximum when the sales growth rate is somewhere between -0.5 and 0 for the U.S. data. It reaches its minimum for U.K. data when the sales growth rate is about -0.5.

The parameters of the grey-box model having firm age as the black-box input variable are plotted in Figure 5.3. The 8 parameters in general have monotonic relationships with firm age and there seems to be steady levels for all of them to converge to as a firm grows to certain ages. For example, it took over 30 years for the AR parameter to converge to a steady value of around 0.75. The younger a firm is, the higher AR parameter the firm has in the cash flow prediction model. The AR parameter and the parameter on depreciation and depletion have slower convergence than the other parameters. The parameter on other components of earnings is almost always 0, where firm age seems to have little impact. The steady value of the parameter on amortisation is also close to zero. The intercept term is negative in the early ages and gradually becomes positive in about 10 years. The charts imply that firms that are older than roughly 30 years with all parameters reaching steady-states could be treated as homogeneous when they are sampled for cash flow prediction. Comparing Figure 5.3 to Figure 4.10 on U.S. data, their patterns of the parameters are obviously different, except the parameters on changes in account payable and changes in account receivable. However, it is noteworthy that the firm ages in the U.S. is a proxy variable while the U.K. firms' ages are real.

5.2.3 The In-sample Fitness of Various Models

The in-sample fitness of various models are compared by two criteria, i.e. mean squared error (MSE) and average rank of each model. A smaller numbers indicates a model's better performance in data fitting. In the Bayesian model, n_0 is set to 5 and S_0, as estimated by random parameter model, is set as 0.0028. The coefficients used in Bayesian model are assigned the same values as in the U.S.A., i.e. coefficient δ_B in (3.37) is set to 0.9; coefficient k in (3.42) is set to 2.5 and δ_C in (3.45) is set to 0.9; τ in step one of monitoring procedure is set to 0.2. The results of the 8 models are shown in Table 5.4. The calculation is based on 2643 observations. Model 1 fits the in-sample data better than Model 2 in both criteria, which is against the assertion made in Dechow et al. (1998) that cash flow plus changes in working capital terms better predict future cash flow than cash flow alone. The random walk model empirically generates lower prediction errors and its average rank also suggests that the random walk process describes the data well even for individual observations. Although Model 3 has lower MSE than Model 1 and 2 as estimated by OLS method, the average rank of Model 3 is higher. Model 4 and 5 have similar performance, the former performing marginally better than the latter. Both of them, taking account of individual effects of firms, have much lower MSE than other models, but their average ranks are among the highest of all. The Bayesian model, with the highest MSE, also has a high average rank that is only lower than the two panel models. The two grey-box models have quite balanced performance. Their MSE is comparable with the simple linear regression, i.e. Model 3, and they both have lower average ranks. Model 7 that applies sales growth rates as the black box input has the third best average rank, only worse than Model 1 and Model 8. Model 8 that uses firm age has obtained the lowest average rank among all models.

5.3 Cash Flow Prediction: Out-of-sample Performance

5.3.1 One-period-ahead Prediction Performance

Observations from 2007 to 2015 are used for out-of-sample prediction comparison. The one-period-ahead predictions made by the 8 models are compared in Table 5.5, which is calculated on 2661 observations. Although Model 1 provides a lower in-sample MSE than Model 2, it has a higher MSE than Model 2 in the out-of-sample period. It however still has lower average rank than most of the other models. The lowest MSE is obtained by Model 3, i.e. the pooled regression model, but it has the worst performance in individual predictions as implied by its highest average rank. Model 4, 5, and 6 have similar performance, none of which is impressive in either criterion. Model 7 and 8, i.e. the grey-box models, have good, balanced performance in the in-sample period. Their MSE is as low as that generated by Model 3 and the average ranks of the grey-box models are among the best of all 8 models. The lower half of Table 5.5 shows the ranks of each model according to their performance by each criterion. Model 1 and 3, although having leading performances in one criterion (Model 1 in average rank and Model 3 in MSE), have the poorest performance in the other criterion. Only the grey-box models could perform well for both criteria. Therefore, it can be concluded that their predictions are accurate and generally applicable for one-period-ahead use.

5.3.2 Multi-period-ahead Prediction Performance

Table 5.6 summarises the performance of the models in predicting cash flows of multiple periods ahead. Model 7 that applies sales growth rates as additional input variable is not applicable in this test because future sales growth rates are unknown *ex ante*. Therefore, the multi-period competition is among the other 7 models. The MSE and average ranks of each model are calculated for the 2867 observations in the study. The in-sample period ends in 2006 and therefore the predictive variables take their values in 2006 if available. The predictions made in 2007 are actually one-period-ahead predictions and thus not included in Table 5.6. Year 2015 is not listed either for there are only 3 observations. Bold numbers indicate the two models performing best for that criterion, which appear 9 times for Model 3 and Model 7. These two models outperform the other models in accuracy as they always provide the least MSE. Model 1 has bold numbers shown 7 times, all in the average rank criterion. Model 2 and 6 have shown none. Model 1 is also better than Model 2 by the criterion of MSE in every year of the out-of-sample period. The evidence is consistent with the in-sample result and is against the DKW assertion. Model 4, 5 and 6 do not provide impressive predictions, which is consistent with the results of one-period-ahead predictions. Comparing Model 3 and 7, the former performs better in predicting cash flows of further future. In the prediction for 2011 and 2013 (corresponding to 5 and 7 years ahead respectively), Model 3 has been one of the two best models by both criteria while Model 7 has such achievement for 2008 and 2009 (2 and 3 years ahead).

In summary, grey-box model could provide impressive and encouraging predictions for future cash flow, especially for the shorter future. Simple pooled regression model also has competitive performance and provides highly accurate predictions. The random walk model, as a benchmark model, is shown to be a general model.

5.4 Risk Aversion and Stock Valuation

In this section, the random walk model, DKW model, BCN model and grey-box model, which perform well according to the results of previous two sections, are examined for their performance in explaining market prices by the discounted simulated cash flow (DSCF) method developed in this thesis. A by-product of this method is the identification of the RRA coefficient in the market. The panel models and Bayesian model have poorer predictive power for cash flow in general, which hence are not considered in this section.

The study period spans from January 2007 to January 2015. Monthly prices of the sample firms are available and obtained from Datastream database. Disclosure date is not available in the database and the date of fiscal year end of each firm is applied as a proxy. The firms for which the fiscal years end in the same month (perhaps different days) constitute a sub-sample and the market risk aversion coefficient of that particular month is calibrated by matching the firms' theoretical values and their market prices. Similar to the U.S., most firms' fiscal years end in December, followed by March, June and September. There are a few observations in other months as well. Government bond rates are obtained from Datastream database and used as risk-free rate. There are available bond rates of 2, 3, 5, 7, 10, 15, 20 and 30 years, which are interpolated and extrapolated following Nelson and Siegel (1987) method to form the entire yield curve of required length.

5.4.1 RRA Coefficients and the Market Index

In the simulation procedure, the length of predicted and simulated cash flow series is set to 100 and the number of simulations is 1000. The calibrated RRA coefficients by the four examined cash flow prediction models are plotted in Figure 5.4. There are upward trends in the first two years period in the four charts, which correspond to the global financial crisis. In the following period, the RRA coefficient series by DKW, BCN and grey-box models are volatile. As a result, there is limited guidance they can provide in investment decision making. Nonetheless, the random-walk-model-resulted RRA coefficients show a slightly declining pattern since 2009. Figure 5.5 shows the association of random-walk-model-resulted RRA coefficients and the FTSE All-Share index of the same period. There is a clearly negative relationship between the two series. In previous chapter, there is no such clear negative association after 2009 between U.S. market index and the RRA coefficients, the latter of which are relatively constant after reaching the peak. In both U.S. and U.K. markets, the stock indexes have been rising since 2009. However the empirical study on the market risk aversion suggests that the U.K. market's recovery seems to be driven more by a loosening attitude towards risk of the market than the U.S.

5.4.2 Stock Price Calculation and Moving Direction Prediction

The theoretical values of stocks are calculated by the DSCF method with the calibrated RRA coefficients. The variances of mispriced proportions of all firms are calculated in consequence, which are 5.18, 60.41, 43.57 and 87.24 for random walk model, DKW model, BCN model and grey-box model respectively. Therefore, the random walk model better fit the market prices than the other three models and could be considered as the best that capture the market expectation of future cash flows. The other three models do not have as good fitness for the unobservable market expectation from this point of view. The share prices are expected to move in the direction towards their theoretical values. Table 5.7 lists the number of correct and wrong predictions of share price directions according to the theoretical values generated by the four cash flow prediction models. Since monthly prices are available, correct predictions are thus defined as follows:

- Undervalued stocks: theoretical value > market price; the share price in any one of the following 12 months is greater than current market price.
- Overvalued stocks: theoretical value < market price; the share price in any one of the following 12 months is lower than current market price.

In Table 5.7, the proportions of making correct prediction for the price moving direction are much higher than the results in previous chapter of U.S. study. It is because the U.S. results are limited to moving directions of share prices in exactly one year’s time, which is relatively long and thus is usually accompanied with more uncertainty. The monthly observations suggest that the share prices do have the tendency to converge to their theoretical values no matter which model is used to predict cash flows. The DSCF works better to discover undervalued stocks as the ratio of making correct predictions to wrong predictions for undervalued stocks is at least over 6 for BCN model and can even reach 7.62 for the grey-box model. The correct/wrong ratios for the category of overpriced stocks are no higher than 4.15 that is achieved by random walk model. This is a very good result in direction prediction and the DSCF model appears to be useful in pricing stocks. With such high rate of correct predictions, many trading strategies could be developed upon the model, which could be left as extensive study.

5.5 Conclusion

This chapter studies the U.K. firms listed in the London stock exchange. Various models are selected to predict net operating cash flow for the sample firms. In addition, the novel DSCF model is also applied to this dataset to examine the association between firms’ cash flow and their market share prices.

Cash flows of individual firms have different scales depending on the firm sizes, therefore it becomes difficult to plot the general trend of cash flows. Instead of normalising different firms’ cash flow series to the same level, this chapter applies a different method. The cash flows are divided by the contemporary total assets and the trend of the resulted cash flow to asset ratio is plotted. In firms' starting stage, the net operating cash flow remains negative for 9 years on average. During this period, the distribution of cash to asset ratio among the sample firms is highly asymmetric. The cash flow to asset ratio has a clear growing pattern with age and converges to a level around 8%. After then, there is no declining pattern clearly shown in the sample. It could be deduced that the U.K. firms cash flows tend to grow at an annual rate near 8%.

The models applied in cash flow prediction include the panel models, Bayesian model and grey-box model apart from selected benchmark models. Panel models allow the intercept of individual firms to be heterogeneous and thus generate lower in-sample MSE than other models. However, the out-of-sample performances of panel models are poor. There is hence problem of over-fitting for the panel models. The out-of-sample MSE of pooled regression is smaller despite that it does not assume individual intercepts. The Bayesian model's performance is not impressive either in-sample or out-of-sample. There may need further tuning for it to improve in such applications. On the other hand, the grey-box models have encouraging performance. When taking two measures of performance into account, the two grey-box models have the most balanced result in the one-period-ahead out-of-sample competition. In the multi-period context, the grey-box model with sales growth rate is not utilised and the one with firm age still shows promising performance. However, with the U.K. data, there is no single model that can consistently outperform others. The simple pooled regression model also has shown a hard-to-beat performance. The random walk model is shown to be very general in describing the cash flow dynamics of U.K. listed firms. It outperforms the theoretical DKW model, which is against the assertion that with accrual terms the cash flow predictability could be improved.

In the application of DSCF model, the market RRA coefficient is calibrated by matching the theoretical values of stocks and their market prices. Four cash flow prediction models are used in the DSCF model framework. The random walk model has generated smoother series of the market RRA coefficients. The market index has been rising since 2009, accompanied with a declining pattern in the RRA coefficients. Therefore, economic recovery or growth cannot take the whole credit for the market boom, which is at least partially driven by the change in the market's attitude towards risk. Besides, the random walk model also has better fitness of the observed market price than others. Therefore it could be concluded that the market expectation of future cash flows is better captured by a simple random walk process. The resulted theoretical value could provide guidance for predicting future moving direction of share price.

Chapter 6 Empirical Study III: U.K. Unlisted Firms

This chapter is aimed to show the behaviour of cash flows of the unlisted firms in U.K. There are far more unlisted firms than the listed ones. To have a fuller view of U.K. firms, it is informative to conduct studies on the dataset of unlisted firms. In addition, better prediction of cash flows for the unlisted firms could also lead to relevant applications, such as Initial Public Offering (IPO) pricing. Differences that exist in the pattern of listed and unlisted firms' cash flows will be shown in this chapter.

Financial data of U.K. unlisted firms is obtained from the FAME database. The firms to be selected have to be active and their locations are limited to England, Scotland, Wales and Northern Ireland. Firms in financial sectors are excluded. The database only provides the most recent 10 years' data for each firm, therefore the sample period spans from 2005 to 2014. After missing values are excluded, the primary cash flow sample consists of 37574 firm-year observations. These observations come from 12756 unlisted firms, approximately 10 times the number of listed firms.

6.1 Pattern of Net Operating Cash Flow to Total Assets Ratio

Figure 6.1 plots the distribution of net operating cash flow divided by total assets with firm ages up to 93 years. The samples that consist of observations over 93 years have sizes smaller than 30 and thus are not included. In the upper chart, the mean, median and 2.5 to 97.5 percentiles of the sample distribution corresponding to certain firm ages are plotted. The mean and median of the ratio distributions are close to each other, which implies roughly symmetric distributions. The mean series is shown alone in the lower chart so that the pattern could be seen more clearly without the impact of larger scales of other curves, especially the 2.5 to 97.5 percentiles. For firms under the age of 20, the mean cash flow to asset ratio seems to be on a slightly upward trend, growing from 8% to 10% and then the ratio declines gradually but not in a smooth fashion. Between 70 and 80 years, the net cash flows of the sample firms are jumping between 2% and 8% of their total assets. The pattern of unlisted firms' net cash flow to total assets ratio is different from that of listed firms studied in previous chapter (Figure 5.1). In firms' early years the listed firms will experience a period with negative operating cash flow while the mean cash to asset ratio of unlisted firms are always positive. A potential explanation could be that the data is concentrated in the period between 2005 and 2014 and as a result the sample firms that provide data illustrating the cash flow behaviours in the early ages are mostly founded during or not much before the same period. Therefore the cash flow behaviour for firms early ages as shown in Figure 6.1 may not be representative given the limited amount of data. Apart from the diverging cash flow patterns in the early ages between listed and unlisted firms, more differences can be observed for the two samples. There is no particular declining pattern for listed firms and the steady levels of the cash flow to asset ratio of listed firms are a little higher than that of unlisted firms. Therefore, not considering survivorship bias, it could be concluded that on average the listed firms outperform the privately owned companies in the cash flow generating capacity measured in total assets, particularly when they have grown into mature or steady stages.

6.2 In-sample Estimation and Fitness of Cash Flow Models

As Figure 6.1 shows the pattern of cash flows for unlisted firms, which is not similar to their listed counterparts. This section will take further steps to see the features of cash flow dynamics as captured by predictive models. The models to be examined are the same as in previous chapter for the listed firms. The in-sample data spans the period from 2005 to 2012 and the data in year 2013 and 2014 is left for out-of-sample use. Such a partition is to guarantee that there are sufficient observations for model estimation, but on the other hand, the out-of-sample results will be limited by fewer observations.

Table 6.1 summarises the descriptive statistics of the variables in the cash flow models. All variables are deflated by the average total assets of each firm and the calculation is based on 24081 firm-year observations. The mean of net cash flow is 0.0933 with a standard deviation of 0.162. The sizes of the other variables are not as large as net cash flow. The variable *Other* denotes other components in earnings and is calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – (net cash flow + changes in account receivable + changes in inventory – changes in account payable – depreciation – amortisation). It has a mean of 0.0503, which is second only to net cash flow.

6.2.1 Estimation of the Panel Data Models' Parameters

Table 6.2 shows the in-sample estimation of model (3.4) by pooled regression, demean method and Arellano-Bond estimator from the second to the fourth column respectively. The pooled regression results suggest that depreciation, amortisation and other earnings components, as predictive variables, are not statistically significant. There is a lack of theoretical explanation why these three variables should have predictive power for future cash flows. However, it is noteworthy that in the previous two chapters where different datasets are studied, depreciation (depreciation and amortisation merged in U.S. data) is statistically significant as estimated by pooled regression. For the unlisted firms, only the lagged cash flow and changes in working capital terms are statistically significant. The parameter sign on lagged changes in payables is negative, which is consistent with theoretical expectation. The parameter on lagged cash flow is 0.53, which is smaller than that of U.S. and U.K. listed data but still exceeds the magnitudes of other parameters. The third column lists the estimation results of the fixed effect model by removing group means of each variable. The demean estimation method has led to different conclusions from the pooled regression. The lagged cash flow and changes in inventory are no longer significant statistically whereas the other earnings component becomes significant. The Arellano-Bond estimator applies the principle of the GMM method and the results are listed in the rightmost column. The number of observations for this method is much smaller than the other two methods because the first difference is taken for the variables. The magnitudes of the parameters thus estimated are smaller than those of pooled regression except other earnings component and again the lagged cash flow is not statistically significant. The two variables that are statistically significant by the Arellano-Bond estimation method are changes in account receivable and other earnings. The in-sample estimation results of U.K. unlisted firms show different pattern from using U.S. and U.K. listed firms' data. Especially the AR parameter is not estimated to be statistically significant when panel models are applied, which implies that the unlisted firms' cash flow might be less persistent and harder to predict than listed firms.

Table 6.3 lists the estimation results of the random parameter model on the in-sample data. The mean value, standard deviation of each parameter and the variance-covariance matrix of the parameters are reported. In general, the mean values of the parameters have larger sizes than the point estimates using pooled regression. However, each parameter is accompanied with high standard deviation as well, which suggests that the distributions of the random parameters are dispersed. For instance, the AR parameter by random parameter model has a mean value of 0.6724 and a standard deviation of 0.5518. The parameter on depreciation has a mean value of -0.2289 but its standard deviation reaches 1.1847, which is also the highest standard deviation among all the parameters. The distributions of the parameters will be used as prior of Bayesian forecasting model.

6.2.2 The Grey-box Models' Estimations

Two grey-box models are estimated, using sales growth rates and firm age respectively as input variables to the black-box part. Figure 6.2 plots the functional form of the parameters on the predictive variables of sales growth rates. The sales growth rates take values from -1 to 1. The patterns of all parameters are nonlinear and none of them is monotonic. The AR parameter reaches its peak when sales grow at a rate between 20% and 40%. Similarly, the parameters on changes in working capital terms, i.e. inventory, account payable and account receivable, also reach their respective peak or trough (for negative parameters) values when the sales growth rate is around 0.2. Figure 6.3 instead shows the impact of firm ages on the parameters. Ages up to 150 years are plotted. There are steady-state values for the parameters to converge to when firms reach certain ages. The intercept term takes higher values in early ages and in later ages it declines and seems to converge to zero. The AR parameter is on a monotonically increasing trend with growing ages. In the early ages the AR parameter is low, which implies less persistent cash flow series. Once firms grow for about 50 years, the AR parameter converges to around 0.5. The other parameters also converge in similar ages. The steady-state value for the parameter on changes in inventory is low, only slightly above 0.1 whereas the parameter values for changes in account receivable and payable reach around 1 and -1 respectively. The other predictive variables' parameters converge to values that are no close to zero. Parameter on depreciation has a steady-state value marginally below 0.4 while it is roughly -0.5 for amortisation. It is 0.2 for the parameter on other earnings components. Comparing the steady-stage value of the parameters in Figure 6.3 to the pattern in Figure 6.2, the steady stage as firms grow over, e.g. 50 years, correspond to a growth rate in sales that is approximately 0, where the parameter values in both figures are close or at least the parameters signs are consistent with each other.

6.2.3 The In-sample Fitness of Various Models

The estimation sample consists of about 3700 observations. The in-sample fitness of the 8 models are compared in Table 6.4. The two measures used to compare model performance are mean squared errors (MSE) and average rank of the models over each observation. Lower values for the two measures indicate better performance. The models resulting in the least MSE are the two panel models, i.e. Model 4 and 5, because of their consideration of individual effects. The grey-box model using sales growth rates follows, with an MSE of 0.0149 and then the pooled regression with an MSE of 0.0158. Grey-box model using firm ages has generated an MSE of 0.017, slightly exceeding that of pooled regression. The MSE of other models are above 0.02, the highest of which comes from the Bayesian model.

The grey-box model with firm ages has the lowest average rank of all models, while grey-box model with sales growth rates, i.e. Model 7, despite its lower MSE, is higher ranked on average than the former. Model 7 however still shows good performance by this criterion, only worse than Model 8 and Model 1. Panel models, i.e. Model 4 and 5, although showing the lowest MSE, are not competitive in the measure of average rank. Model 2, the theoretical DKW model, predict the cash flow using current cash flow plus changes in working capital terms, has a poorer data fitting performance than the simple random walk model, according to both measures. It underperforms the pooled regression as well.

6.3 Out-of-sample Prediction Performance of Cash Flow Models

Due to the fact that the data availability is limited to the most recent 10 years, the out-of-sample period only includes the years of 2013 and 2014. There are 2735 firm-year observations in 2013 and 2336 observations in 2014. All the out-of-sample observations are used to test the one-period-ahead predictive ability of the 8 models. Besides, observations in 2014 are additionally used to examine the models (excluding Model 7) in two-period-ahead predictions.

The one-period and two-period results are both summarised in Table 6.5. In 2013, the predictions are made based on the observations in 2012. The lowest MSE is generated by the grey-box model with sales growth rates, followed by the pooled regression. Bayesian model performs well as its MSE is comparable with that of pooled regression. Grey-box model with firm age has an MSE higher than the mentioned three models but lower than the rest. The panel models, although taking account of individual effects, have higher out-of-sample MSE. The MSE generated by Random walk model and the DKW model are even higher. The average rank results suggest that Model 8 is the most general model to make individual cash flow prediction, followed by Model 7. The result shows superior power of the grey-box models in capturing cash flow dynamics. Apart from the two grey-box models, the random walk model has the third lowest average rank in 2013. Taking both measures into account, Model 7 in general has the best performance in 2013.

The third and fourth rows in Table 6.5 panel A are the comparison result of one-period-ahead predictions for the observations in 2014. The prediction is based on the predictive variables in 2013. The performances of the two grey-box models are similar to that in 2013. Model 7 again generates the least MSE in 2014 and has the second lowest average rank. Model 8 has the lowest average rank but the MSE is not as competitive. The Bayesian model has very low MSE, second only to Model 7. On the other hand, the two panel models have shown the largest MSE, even larger than Model 1 and 2. It could be seen as a sign of over-fitting for the panel models, i.e. although their in-sample MSE is the lowest of all models, their out-of-sample performance is inconsistently poor. The conclusion hence can be drawn that in general Model 7 outperforms the other models in making one-period-ahead prediction in the out-of-sample period.

Panel B in Table 6.5 lists the results of the models in predicting the cash flows in 2014 using the observations in 2012. Therefore, the performance of the models in making two-period-ahead prediction is examined and Model 7 is excluded. Below the two measures are listed the rank of each model in each performance measure. The least MSE is generated by Bayesian model, followed by pooled regression and grey-box model with firm ages. The lowest average rank is achieved by the demean panel model, followed by the grey-box model and the Bayesian model. Therefore, it is hard to identify which model dominantly outperforms the others. According to the two measures, the Bayesian model seems to have a generally better performance in this sample and the grey-box model also performs well in both measures. Besides, the random walk model outperforms the theoretical DKW model, which is consistent with the one-period result.

6.4 Conclusion

The unlisted U.K. firms are studied in this chapter. The cash flow pattern of unlisted firms is different from that of listed firms. Firstly, cash flows of unlisted firms are on average positive regardless of age. There is no phenomenon of net cash outflow in early ages as in the U.K. listed sample. Secondly, the unlisted firms' cash flow to asset ratio tend to decline after 20 years whereas the listed firms does not. The listed firms' ratio on average stays higher than the unlisted firms.

In addition, the cash flow series of unlisted firms are likely to be less persistent than the listed firms, as implied by the lower AR parameter value in the cash flow prediction models. Models applied for the listed data are also applied in this chapter. The over-fitting problem of the fixed effect model and the GMM model prevails. According to currently available sample data, grey-box model has shown its strength in making individual predictions with high accuracy. In the one-period prediction comparison, the best performance for both criteria is achieved by the two grey-box models separately. In the two-period result, the grey-box also provides good performance. Besides, there is also an encouraging result for the application of the Bayesian model. However, no single model could dominate the other competitors. Limited by the data availability, there are only one year's observations in the study of multi-period prediction. The result cannot be treated as very representative. To reach a sounder conclusion, there is need for more data collected in the future.

Chapter 7 Empirical Study IV: Chinese Firms

China, currently as the second largest economy, has been attracting the attention of investors and researchers. The high growth of China's economy and the possibility of maintaining it has long been a hot and controversial topic in the media as well. China's economy has been developing only for a relatively short period of about 40 years. The Chinese economic system is different from U.K. and U.S.in many aspects including the laws, the role of government and so on. The stock market also has its specialities. The majority of the listed firms are state-owned. Thus there is demand for studies into Chinese firms to see what pattern their incomes follow as in the rapidly developing country. This chapter is aimed to study the cash flow behaviour of China listed firms. There are two stock exchanges in Shanghai and Shenzhen, both of which were founded in 1990. The subjects of this chapter are the firms listed as A stock in the two stock exchanges of China. A-stock firms are the majority of China listed firms and they are the main investment targets for domestic investors in China.

7.1 Pattern of Net Operating Cash Flow to Total Assets Ratio

The financial data of the sample firms is obtained from the RESSET database. The sample firms are limited to A stock firms and non-bank firms. The number of sample firms is 3460. Annual cash flow data is collected from 1997 to 2015. Thus there are 36047 net operating cash flow observations in total, from which the trend of cash flows' evolution could be examined. These cash flow observations are divided by firms' contemporary total assets, forming cash to asset ratio and the values of the ratio of the observations are sorted into sub-samples according to the firms' ages in the particular observation year. The distributions of the sub-samples are summarised by mean, median and 2.5 to 97.5 percentiles, which are plotted in Figure 7.1. The samples of observations over age 30 consist of less than 30 observations, which are thus excluded from the chart.

The upper chart shows that the mean and median cash flow are always positive. The distributions are approximately symmetric. It is difficult to see the evolving pattern of the mean cash to asset ratio along with growing firm ages. Therefore, the lower chart separately plots the mean series and the pattern becomes clearer. Although the mean cash flow is always positive, the ratio declines as firm grows. In the early ages, the mean cash to asset ratio stays above 0.04, while for firms over 20 years old, it declines to about 0.03. The cash to asset ratios of Chinese firms are much lower than the U.K. firms of the same ages. The U.K. firms, when growing to 30 years, still could generate cash flows of about 8% of the total assets. On the other hand, China firms' cash flow generating ability could not reach 6% of the total assets even in their early ages when they are expected to experience a high growth stage. Moreover, the declining speed of the cash to asset ratio for U.K. firms is generally lower than China firms. China firms' mean cash flow to asset ratio declines from 0.05 to 0.03 (40% decline) in about 20 years while Figure 6.1 shows that it takes about 50 years for the mean cash to asset ratio of U.K. unlisted firms to decline from 0.1 to 0.06. The cash flow to asset ratio could be seen as a measure of a firm's efficiency and from this point of view, the efficiency of listed China A-stock firms still need to improve as their performance is not consistent with the positive view of China's growth. Their role as investment targets for Chinese investors might be in doubt too.

7.2 Cash Flow Prediction: In-sample Model Estimation

The RESSET database provides the direct method statement of cash flow, hence with the additional data it could be examined whether the disclosed cash flow components could improve the predictability of future cash flows. The disaggregated cash flow model will be in the form of:

$$\begin{aligned} CF_{t+1} = {} & \beta_0 + \beta_1 CF\_Crec_t + \beta_2 CF\_Cpaid_t + \beta_3 CF\_NCrec_t + \beta_4 CF\_NCpaid_t \\ & + \beta_5 \Delta INV_t + \beta_6 \Delta AP_t + \beta_7 \Delta AR_t + \beta_8 DEP_t + \beta_9 AMORT_t + \beta_{10} OTHER_t + \varepsilon_{t+1} \end{aligned} \tag{7.1}$$

where CF_Crec denotes cash received from the sales of goods and rendering of services, CF_Cpaid denotes cash paid for goods and services, CF_NCrec denotes other cash receipt calculated as sub-total cash inflows from operating activities minus CF_Crec and CF_NCpaid denotes other cash paid out calculated as sub-total cash inflows from operating activities minus CF_Cpaid. Thus net operating cash flow is disaggregated into four components. The models examined in this section are the random walk model (Model 1), the theoretical DKW model (Model 2), the BCN model, i.e. equation (3.4) estimated by pooled regression (Model 3), demean method (Model 4) and GMM method (Model 5), model (7.1) estimated by pooled regression pooled regression (Model 6), demean method

(Model 7) and the GMM method (Model 8), the Bayesian model (Model 9), the grey-box model with sales growth rate (Model 10) and firm ages (Model 11) separately as additional input variable to the black box. Models 6, 7 and 8 have more predictive variables than the other models. Their performance can be directly compared with Model 3, 4 and 5 respectively to see whether the disaggregation of net cash flow will benefit the forecast.

Table 7.1 summarises the descriptive statistics of the variables in use. All variables have been deflated by average total assets. The statistics are calculated over a sample of 30596 firm-year observations. The deflated net operating cash flow has a sample mean of 0.0492, which is lower than that of U.S. data and U.K. unlisted firms but higher than the U.K. listed data. The standard deviation of the China data is the lowest among all the datasets studied in this thesis. The third to sixth columns summarise the cash flow components' descriptive statistics. From the mean of the four variables, it can be roughly seen the source of net operating cash flow. The cash received is mostly from selling goods or delivering services, about 80% of average total assets. Cash received from somewhere else only has a mean of 0.0566. The structure of cash outflow is a little different as the cash purchasing goods or services is only about 0.6 along with cash paid out for other purposes of about 0.2, one third of the size of the former. It is worth noting that the changes in working capital terms have relatively high mean values compared with the U.S. and U.K. data. The three variable in U.S. and U.K. samples are mostly smaller than 0.01. There are exceptions that are still smaller than 0.015. However, in the China sample, all the three variables have mean values above 0.02 and the mean of changes in account receivable even reaches 0.0364, which is comparable to that of net cash flow. As a result, the proportion of accruals for China listed firms is relatively large, which indicate that the operations of China listed firms might involve large amounts of trade credit.

7.2.1 Estimation of the Original Panel Data Models' Parameters

The in-sample period is set as from 1997 to 2006 and thus the out-of-sample period is from 2007 to 2015. The in-sample estimation results of Model 3, 4 and 5 are listed in Table 7.2, which are based on 8046 observations. The second column lists the estimated parameter values using pooled regression. All the variables except other earnings components are statistically significant. However, the parameters on the first four variables have low values. The AR parameter is only 0.239 and the parameters on the changes in working capital terms are around 0.1 or -0.1. Nonetheless, the parameters on depreciation and amortisation take much higher values, both exceeding 1 in absolute value. The effect of depreciation to predict future cash flows is positive whereas amortisation has negative effect. The contradiction is hard to explain because the two variables are recorded for similar purpose and are only different on the types of assets applied. The third column reports the results using demean method to take account of individual intercepts. Only parameters on changes in account payable and depreciation are statistically significant by this method. The rightmost column reports the parameter values estimated by Arellano-Bond method and none of the parameters is statistically significant by this method. The pattern of the estimation results using panel models is similar to that on U.K. unlisted data, where most of the parameters becomes insignificant too. This pattern might imply that the China firms' cash flows are more difficult to predict than those for other countries.

Table 7.3 lists the distribution of the parameters according to random parameter model. The mean, standard deviation and the matrix of correlation coefficients between each parameter are reported. The standard deviations of all parameters are higher than their respective means, partially because of the fact that the latter are low in general. The parameter on amortisation has the largest standard deviation, followed by depreciation. There are also high correlations between each other for almost all parameters. Relatively, only the parameter on amortisation has lower correlation with other parameters. The result of random parameter model will be used as prior for the Bayesian model.

7.2.2 Estimation of the Extended Linear Models' Parameters

Parameters in (7.1) estimated by three different methods are listed in Table 7.4. The second column lists the results by pooled regression. With disaggregated cash flow components as additional predictors, the parameter on changes in accounts receivable becomes 0.0491 and statistically insignificant. It is 0.0838 and significant in the model with aggregated net cash flow as predictor. The parameter on other earnings components as statistically insignificant in the BCN model remains so in the more detailed model. The parameters on the four cash flow components are not all significant. The cash received from other sources only has a parameter value of 0.0784 and it is insignificant. The two components, cash involving goods or services, have higher absolute parameter value than that on cash paid for other purposes, although the latter is still significant. It thus suggests that the persistence of different cash flow components are not equal and disaggregating them if data is available might improve the predictive power of cash flows. Similar to the BCN model, when the two panel model is applied, most parameters become insignificant. Especially in the fourth column where the GMM estimation is applied, none of the parameters is significant any more.

7.2.3 Estimation of the Grey-box Models' Parameters

The grey-box models with different input variables are also estimated on the in-sample data. The cash flow model parameters are considered as the output of the black box model and are assumed to be functions of the extra input variable. The function is not clearly known but the approximated pattern could be plotted numerically. Figure 7.2 plots the association of the parameter values in the BCN model with sales growth rates, the latter of which are limited between -1 and 1. It is noteworthy that the AR parameter, as shown in the second chart takes negative values when experiencing dramatic drop in sales, which implies that such abnormal situation tends to be reversed in the next period. The AR parameter in general increases with sales growth rate but does not exceed 0.5. It is also worth noting that the parameter on amortisation is positive no matter whether sales increase or decline, which is inconsistent with the pooled regression result. Figure 7.3 plots the evolvement of the parameter values with growing firm ages. All parameters except that on amortisation seem to converge within a certain period. The AR parameter converges to roughly 0.2 very quickly, which only takes about 20 years. The steady-state value of the depreciation parameter can reach 3, which is very high value. The parameter on amortisation takes negative values and does not converge to a steady level within 100 years. The unreasonably high values for the two parameters might be due to the fact that the firms included in the sample are mostly no older than 30 years. Therefore, the resulted parameter values of over 30 years are not as reliable as otherwise. This drawback might lead to difficulties and inaccuracy for long-term predictions.

7.2.4 The In-sample Fitness of Various Models

The in-sample fitness of all the models are compared in Table 7.5. The two measures, i.e. MSE and average rank, are used to compare the models' performance. The models that generate the lowest MSE are Models 4 and 7, followed by Models 5 and 8. It is consistent with expectation because they estimated intercept terms for individual firms. Apart from them, the lowest MSE is generated by Model 6, the cash disaggregation model, and Model 10, the grey-box model using sales growth rates. Comparing Models 6 and 3, it can be seen that the disaggregation of cash flow results in better fitness of the in-sample data, as both measures of Model 6 are superior to Model 3. Model 10 has not only low MSE but also low average rank. Its average rank is only higher than Model 1 and Model 11. The comparison suggests that grey-box models in general better fit the in-sample data.

7.3 Cash Flow Prediction: Out-of-sample Performance

7.3.1 One-period-ahead Prediction Performance

The data from 2007 to 2015 is used of out-of-sample comparison. The one-period-ahead results are summarised in Table 7.6. The MSE and average ranks are calculated on 10104 observations. Model 6 has the lowest MSE. Besides, it outperforms Model 3 in both criteria. Similarly, Models 7 and 8 respectively outperform Models 4 and 5. The comparison shows the incremental power of using disaggregated components of net cash flow as predictors. Model 11 has the lowest average rank, followed by the random walk model and then Model 10. Model 6 has a moderate average rank although it performs best according to the MSE criterion. Model 1's low average rank suggests that it could well describe the general pattern of the sample firms' cash flow dynamics. However, it has the largest MSE of all the models. Therefore, no single model could outperform others in both criteria. The Grey-box models have encouraging performance as they have low MSE and low average rank as well. Although Model 10 does not outperform Model 6 in the MSE measure, it is worth noting that Model 6 has more variables than Model 10 and hence more detailed information is exploited by Model 6. Model 10 still has an MSE which is very close to that of Model 6 despite leaving fewer predictors. Model 1 is the simplest random walk model, which is usually used as benchmark model. In this case, Model 11 is the only model that beats Model 1 in both criteria. If Model 2, the theoretical DKW model, is used as benchmark, there will be only two models that could beat it in the two measures, i.e. Model 10 and Model 11.

7.3.2 Multi-period-ahead Prediction Performance

The multi-period prediction results are summarised in Table 7.7. The grey-box model using sales growth rate is excluded. The data used is from 2008 to 2015 and year 2007 is excluded. The MSE and average ranks of the models are calculated over 7146 observations. The grey-box model performs abnormally poor in this application, which is not consistent with its encouraging one-period result. It is the worst model of all and the MSE is unexplainably high. One potential reason for the grey-box model's inconsistent performance in one-period and multi-period application could be that the grey-box model fails in capturing the dynamics of other predictive variables, which is required for multi-period-ahead cash flow prediction. Therefore, a solution for the problem may to replace the prediction of other variables with a simpler model, e.g. Model 3, and still use the grey-box model to predict cash flow. Thus, a combinative model will still keep the superior performance of the grey-box model. The result in applying the combinative model is listed in Table 7.8. As expected, there is a remarkable improvement of the grey-box model when it is combined with pooled regression. The combinative model performs well in terms of both measures but does not rank the first in either. Model 6 has the lowest MSE and Model 1 has the lowest average rank. As the same as in the one-period prediction test, none of the models could outperform the others in both criteria, but Model 10 provides the most balanced performance. Although it has higher MSE than Model 6, fewer variables are adopted by Model 10 and thus Model 10 has an advantage over Model 6 that it can apply to the situation where direct method disclosure of cash flow is unavailable. On the other hand, the differences in the MSE generated by the models except the Bayesian model are small and negligible. For instance, the MSE of Model 6 is 0.0291, merely 1% smaller than that of Model 10 (0.0293) and 5% smaller than that of Model 1 (0.0306). MSE could be used to measure the predictability of cash flows. Higher MSE usually indicates more difficult predictions. Comparing the MSE in multi-period performance of various dataset in this thesis, it can be seen that the MSE of China and U.K. unlisted data are generally higher than that of U.S. and U.K. listed data, which implies that firms in the former samples have less predictability in future cash flows.

7.4 Risk Aversion in the Stock Market

This section will demonstrate the result of applying the DSCF method in China stock market. The random walk model is applied as the predictive model for cash flow. It has been shown in previous section that random walk model could well capture the general pattern of China listed firms' cash flows and the accuracy is also comparable with other models. Besides, its application is the simplest of all models. For instance, Model 6 involves four cash flow components, which will incur more complexity and computing burden in the simulation procedure. The grey-box model also has its drawbacks in long-term prediction as shown in previous sections. Especially the ages of the sample firms are in a narrow range. Simple cash flow model, as demonstrated in the U.S. and U.K. data studies, could also have encouraging result in evaluating the markets risk attitude and stocks' theoretical values.

The calibration of the RRA coefficient is done in monthly frequency. The sample firms that make disclosure in the same month will be used to calculate the market RRA coefficient in that particular month. Not like U.S. and U.K. firms, most China firms make disclosure in the first four months of year, leaving almost none observations for the other eight months. Therefore, the RRA coefficients are only calibrated in January, February, March and April of each year in the out-of-sample period. The RRA coefficients of other months will be linearly interpolated. The calculated RRA coefficients from February, 2007 (there is no observation in January) to March, 2015 is plotted in Figure 7.4, along with the contemporary HUSHEN 300 index that includes the largest 300 firms in Shanghai and Shenzhen stock exchanges. In 2007, the market index soared from lower than 2000 to nearly 6000. The RRA coefficient in the beginning of the period is relatively high, staying above 0.7. In 2008, as affected by the global financial crisis, the market index dived as quickly to below 2000. The RRA coefficients calibrated in the beginning of 2008 are lower, reflecting a decline in the market's risk aversion. As the index drops, the market risk aversion bounced back to high position again in 2009. There has been a clear declining trend in the RRA coefficient since 2012, implying an increasing intention to invest. A new bull market started in the late 2014 accompanied, or even possibly driven, by historically low risk aversion coefficients.

7.5 Theoretical Value and Return Predictability

The theoretical values of the stocks are calculated by the DCSF method using the calibrated RRA coefficients. The sample is again limited to the first four months of year. The stocks of which the current prices exceed their theoretical values are considered as overvalued and are expected to have lower than current prices in the future, vice versa. The prediction is treated as successful if the expectation actually realises in any one of the following 12 months. To allow for sufficiently long following period for each observation, the firms making disclosure in 2015 are not considered. The predicted numbers are listed in Table 7.9. For comparison, the result of naive prediction is listed first. There are in total 12466 observations and the naive guess is that all the shares' prices will go up or down in the future. The possibility for the guess to be true is very high. There is only about 20% possibility that share prices go to one particular direction and never return to their current levels. That is also the probability of making wrong prediction. In the majority of the cases, the share prices will return to their current levels at some time during the following year no matter in which direction they initially move. The numbers of correct predictions are over 3.9 times that of wrong predictions for both directions. The DSCF model identifies 3206 under-priced shares and 9260 over-priced shares. For both categories, the ratios of correct predictions to wrong predictions are higher than when no strategy is applied. Two-sample *t*-test is conducted to test whether the ratio using the DSCF method is statistically higher than that involving no strategy. The bootstrapping method is adopted for the purpose. 1000 sub-samples are drawn from the whole sample, each containing 1000 randomly selected observations. The ratios are calculated for each sub-sample and the samples of ratios are used for the *t*-test. For both over-priced and under-priced categories, the *t*-test rejects the null hypothesis that the correct to wrong ratios using DSCF and no strategy are equal at very high probability. Therefore, the DSCF model can provide incremental information to investment decision making. Besides, there is a higher probability for the under-priced shares than over-priced shares to actually realise the expectation, which is consistent with the evidences of other datasets.

To examine the stock return predictability, the following regression is run on the available observations:

$$r_{t+k} = \alpha_k + \beta_k E_t(r) + \varepsilon_{t+k} \tag{7.2}$$

where r_{t+k} denotes accumulated stock return in *k* months; *k* takes integer values from 1 to 12; $E_t(r)$ denotes expected return by the DSCF model, calculated by (3.54); α_k and β_k are parameters. The significance of β_k will indicate predictability of stock returns. The estimated parameters of 12 regressions are listed in Table 7.10. β_k for *k* greater than 1 are all statistically significant and the parameter value increases with *k*. Therefore, the theoretical value obtained by the DSCF model could provide guidance for judging the future pattern of share prices.

7.6 Conclusion

As the largest economy among developing countries, China listed firms are studied in this chapter. The cash flow to total assets ratio shows poorer cash generating ability than U.K. firms. As discussed before, the ratio could be seen as a proxy for firms' growth rate. From this point of view, the China listed firms' performance does not match the impression given of the country's high growth level. Deduced from that, their efficiency might be outperformed by unlisted firms, which is only a hypothesis left for future studies. Another potential problem lies in the finding that the working capital terms on average have high proportions in the total earnings. There might be more space for earnings management.

The application of cash flow prediction models have generally tested two questions: whether the panel models, Bayesian models and grey-box models could outperform the simpler benchmark models and whether the incremental information in direct method cash flow disclosure could enhance the cash flow predictability. Panel models and Bayesian model' performances are not impressive and they do not show a leading position in either criterion. The grey-box models have the most balanced performance, but still lack dominancy, especially with fewer predictive variables than the newly added competitors. In addition, the grey-box model seems to fail in predicting other variables despite its strong power in cash flow prediction. Supported by the empirical result, the answer to the second question is positive. The disaggregation of cash flow does improve the model's performance. The model with more variables estimated by pooled regression has the lowest MSE in the out-of-sample test. The finding supports the necessity of direct method disclosure.

The calibrated RRA coefficients of the stock market show that the recent boom since later 2014 is associated with very low market risk aversion. In the framework of the DSCF model, stock prices are driven by two factors only, the cash flow prospect of a firm and the risk aversion of the market. If the China economy and thus firms' incomes did not keep growing, the stock market's rise will fully depend on further reduction of the risk aversion, which will become more and more difficult as the current RRA is already low. The random-walk-model predicted theoretical values are compared with their actual prices and predictions of future price. From a two-sample *t*-test on bootstrapped sub-samples, the DSCF model outperforms a naive guess in predicting share price movement. Additional regressions are conducted to test the predictive power of stocks' expected return by the model. The results favour the application of DSCF model as the implied expected return by the method has a statistically significant effect in predicting future accumulated return. The effect is not statistically significant though for one-month-ahead return prediction perhaps because prices need a longer period to move towards the theoretical values.

Chapter 8 Comparisons and Discussions

8.1 Comparisons

This study has been undertaken using data collected from U.K, U.S.A. and China. A battery of models are applied to cash flow prediction and their precision in cash flow prediction are also evaluated. Four datasets have been used in this study to investigate the performance of the models proposed in this study and the results are reliable and sufficiently robust. In this chapter, the performances of the various models are compared and the resultant conclusions using different datasets as well as any differences are highlight.

Chapter 4 reports the results of cash flow prediction models and the DSCF model using U.S. data. Chapter 5 and 6 focus on U.K. listed firms and unlisted firms, respectively. Chinese data are analysed in Chapter 7. It is interesting to see how the average pattern of cash flow series of sample firms grow with age in general, as well as what differences there can be between the applied datasets.

The distributions of normalised cash flow (cash flow to total asset ratio), as measured by mean, median and 2.5 to 97.5 percentiles, are plotted along with firm age for U.S. (U.K. and China) data. Although the plots of U.S. and U.K. listed firms are not directly comparable, they both lead to a similar conclusion. On average, cash flow is on a rising trend. Not considering survivorship bias, the two samples have similar growth rate of cash flow if cash flow to total asset ratio could be used as approximate for cash flow growth rate. A noteworthy observation is that U.K. listed firms show that the cash flows in firms' early ages are negative on average and the cash flow to total asset ratio tends to stay constant when growth rate reaches about 9%. This is not the same pattern for U.S. sample firms for which the initial negative observations have been excluded from the sample. However, U.K. unlisted firms show different patterns. The average cash flow to total asset ratio is always positive throughout the ages for sampled firms and has a gradually declining pattern after 20 years.

The contrast of listed and unlisted samples may provide indications for predicting cash flows. For instance, making long-term forecast of unlisted firms' cash flows may be more difficult than for the listed firms. Similarly, China listed firms also shows positive average cash flows in their early ages. It could be due to the different operational strategies or styles between U.K. and Chinese firms. Besides, it is also likely that the listed firms in China have been required to meet certain conditions and thus the firms with negative cash flows have not been allowed to list publicly in the first place. The average cash flow pattern can be seen as an indicator for the firms' operational quality and competitive ability. Compared to U.S. and U.K. firms, the average cash flow to total asset ratio of the Chinese firms not only is lower in general, but also, declines more quickly and sharply. Therefore, the forecast of China listed firms' cash flows might be even less optimistic. The difference in patterns of behaviour in the datasets implies that practical situations need to be taken into account for such applications and the specialities of certain dataset that come from systematic factors such as culture, economy and regulations etc., which is likely to result in different conclusions and also varying effects of models.

The descriptive study of cash flow patterns is followed by the implementation and comparison of predictive models. The proposed models in Chapter 3 are compared with the benchmark models including the pooled regression. The panel models have very different estimation results from the pooled regression method. In general, the parameters estimated by the panel models are smaller in magnitude than that estimated by pooled regression. This phenomenon could be explained by the fact that ignoring individual effects may introduce upward bias in the parameters.

As a result, a variable in the cash flow prediction model that is statistically significant when estimated by pooled regression may become insignificant when the panel estimation methods are applied. Besides, the variables that are shown to be useful for predicting cash flow as in BCN model may not apply to different countries. For instance, in the U.K. listed sample, the parameters on 'amortisation' and 'other components in earnings' are statistically insignificant as estimated by pooled regression.

The grey-box model plot also confirms the limited information these two variables could contribute in making cash flow predictions. In the U.K. unlisted sample, in addition to these two variables, the parameter on 'depreciation' is also statistically insignificant as estimated by pooled regression. Apart from the significance of variables, there are also differences in the relative weights on the statistically significant variables among different countries. For instance, the magnitude of the parameters on ‘depreciation’ and ‘amortisation’ are extremely large for China data, both of which have absolute values of over 1 while the AR parameter is merely 0.239 (Table 7.2). For the other three datasets, it is always the AR parameter that has the greatest magnitude. There are also findings comparing the U.K. listed and unlisted firms. The AR parameter by pooled regression for listed firms is much higher (0.83, Table 5.2) than for the unlisted firms (0.53, Table 6.2). When panel models are applied, the AR parameter becomes insignificant in statistical terms for unlisted firms. As such, it can be concluded that the listed firms tend to report more persistent cash flow series than the unlisted firms.

Bayesian modelling applied in this thesis assumes the model parameters follow random walk process and thus does not ‘estimate’ the parameters as the panel models do. The prior of the Bayesian model is taken as the estimation result by random parameter model (RPM). The estimated parameters using the RPM show how the parameters behave for different firms. Especially, the variances of the parameters can be seen as a measure of heterogeneity. For instance, the AR parameters estimated by RPM for the U.K. listed firms have a mean value of 0.7453 and a standard deviation of 0.3715 (Table 5.3). The mean is roughly twice the standard deviation, which implies a relatively concentrate distribution. Therefore, firms in the U.K. listed sample have higher persistent cash flows in general. On the contrary, the AR parameters of Chinese firms have a mean value of 0.3293 whereas a standard deviation as high as 0.6323 (Table 7.3). It can be deduced that a larger proportion of Chinese firms have low persistent cash flows. For the rest two samples, i.e. the U.S. listed firms and U.K. unlisted firms, the mean and standard deviation of AR parameters are closer to each other, which implies less extreme distributions than the above two samples. The other parameters can be interpreted in the same way.

8.2 Further Discussions

Despite the different results in parameter estimation, the performance of various models is interesting to discuss, especially in the out-of-sample tests. Counter-intuitively, the panel models have poorer performance than the simple pooled regression despite the consideration of individual intercept terms, with no exception in the four datasets. Therefore, it can be concluded that the panel models, even if capturing partial heterogeneity of individual firms, brings more bias into the models and makes them impractical for prediction. The potential heterogeneity should be more complicated than that can be explained by individual intercept terms. Bayesian modelling does not place such restrictive assumptions on the parameters of the model. It allows every parameter to differ between individuals and even between an individual's consequential observations. The RPM prior is compared against another more 'naive' prior in the Bayesian modelling performance and the result is in favour of the RPM prior.

However, the application of Bayesian model in cash flow prediction cannot be considered as successful. In general, the accuracy of Bayesian model is worse than the other models, including the benchmark models. The Bayesian model does not show encouraging result on the criterion of average rank either. There is only one exception in the study of U.K. unlisted firms, where the Bayesian model shows superior power to other competitors. This thesis is among the first studies to apply the Bayesian model to the panel data setting. The feature of panel data would limit the power of the Bayesian model because of several subjective factors. For example, the determination of coefficients and prior information would cause deteriorating performance of the model in fitting empirical data. The good result for U.K. unlisted firms' data indicate potential of the Bayesian model, which is worth exploring further. The results suggest that there may be particular peculiarities with unlisted firms' cash flows which may differ from the listed firms. Therefore, there is a need for future studies on unlisted firms to investigate this phenomenon.

8.3 Grey-Box Models

The Grey-box model is shown to be a superior cash flow prediction model by the results of all the four datasets. The functional form of grey-box is relatively simple. The processes of parameters in the clear box, or the first-principle based model, are captured by local black box models, each of which is determined by 5 coefficients. The variables entering the black box are selected to be the sales growth rate and the firm's age, the former of which is selected by theoretical derivations and the latter is used as a proxy for the former. The black box is used to model the relationship of the input variable and the output, i.e. the parameters. There is no analytical form for the relationship, but it can be shown in plots. The resulting plots of the black box are smooth and regular. There are clear nonlinear relationships between the input variables and the output parameters. The patterns of the two input variables with the output parameters are different. The function of sales growth rate on the parameters has a less regular shape than firm age, where the latter tends to converge to steady levels. As there is no difficulty in predicting firm age, it offers advantages in making multi-period-ahead cash flow predictions. Fruther, the functional plots of the grey-box model are also different across the four datasets. For instance, for U.S. listed firms, U.K. unlisted firms and China listed firms have shown an increasing pattern for the AR parameters in the cash flow model when firm age is increasing. That is to say, as firms grow older, their cash flows will become more persistent. However, that of the U.K. listed firms is the opposite. The AR parameter declines with the firms' age.

In addition to the shapes, the time for parameter convergence is also crucial for analysts. According to the grey-box model with a firm's age, observations older than the convergence time of the parameters could be treated as homogeneous in cash flow prediction. There is therefore an implication that the sooner the parameters converge, the less gain can be achieved by the grey-box model over pooled regression. The implication is supported by the practical performance of the models in the four datasets. The U.K. data for both listed and unlisted firms shows quicker convergence of the parameters with firm age. It takes about 20 years for the U.K. unlisted firms to have converged parameters while it is even sooner for the U.K. listed firms. However, for U.S. and China data, the convergence time exceeds 50 years in general. In the U.S. sample, the grey-box models are the best models according to the two evaluation criteria. For the in-sample fitness, the lowest average rank is obtained by the grey-box model with firm age. For the one-period out-of-sample predictions, the two grey-box models have the lowest MSE and the one with firm age simultaneously has the lowest average rank. For the multi-period out-of-sample predictions, the grey-box model with firm age again has the lowest MSE and the lowest average rank. Therefore, the grey-box model is proven to be a remarkably more successful model to predict cash flow for U.S. firms. Similarly, in the China sample, the grey-box model with a firm's age beats the pooled regression model in all cases, the only exception being in the MSE for one-period out-of-sample prediction with negligible difference, which is 0.0231 for pooled regression and 0.0233 for the grey-box model. On the other hand, the U.K. samples do not favour the grey-box models as much as those for the other countries. In the sample of listed firms, the least MSE for one-period out-of-sample prediction is obtained by the pooled regression. In the multi-period setting, pooled regression performs as well as the grey-box model. Similarly, in the unlisted firms sample, the grey-box model with firm age still cannot beat pooled regression in the MSE measure but the grey-box model with sales growth rate has the lowest MSE in one-period out-of-sample competition. Therefore, the comparative results have supported the implication that slower convergence of grey-box model parameters with firm age should lead to more gain in using grey-box model over pooled OLS.

In general, the grey-box models can generate very low MSE and simultaneously has low average rank. It is the most balanced-performing model. The two grey-box models perform among the best of all models. When they cannot dominate, neither could any other model. Comparing the performance of the two grey-box models, the one with sales growth rates is better in accuracy as it more often has lower MSE in one-period out-of-sample tests. On the other hand, the grey-box model with firm age usually has lower average rank, showing its general fitness to the empirical data.

8.4 Empirical Data and Random-Walk

In addition to the performance of various models, a few minor conclusions are drawn from analysing the empirical datasets. First, in the China sample, due to the availability of direct method disclosure of cash flow, the incremental power of cash flow disaggregation is also examined. The empirical results suggest that the direct disclosure of cash flow is really helpful in predicting future cash flows. The conclusion is consistent with previous studies such as Cheng and Hollie (2008) and Orpurt and Zang (2009), but this thesis is among the first to examine this argument based on out-of-sample comparison. In addition, the DKW assertion that using earnings could result in better prediction of future cash flow than using cash flow is not supported by all empirical datasets. In the U.K. listed sample, the random walk model outperforms the DKW model in the multi-period competition by both criteria.

Apart from the panel models, the Bayesian model and the grey-box model, it is noteworthy that the benchmark models, i.e. random walk model, theoretical DKW model and pooled regression BCN model, despite their simplicity, provide good performance difficult to surpass in many cases. For instance, for the China data, the random walk model has the lowest average rank in multi-period out-of-sample predictions. Therefore, the simple models are neither useless nor weak, especially when considering that the amount of calculation required is minimal.

Following the comparison of cash flow prediction models, the DSCF model is also applied for the listed firms of the three countries. The outputs of the model are theoretical stock values. A by-product of the DSCF model is the market RRA coefficient at a particular time. All the three datasets include the 2008 global financial crisis. During the financial crisis, the market indexes are closely related with the market RRA coefficients. The peaks of the market indexes correspond to a local minimum of the RRA coefficients and the fall of the market indexes is accompanied by an increasing RRA coefficient. Based on this implication, even though the stock markets of the three countries are all on obviously upward trends, it is perhaps more risky in U.K. and China stock markets than in U.S. market because the RRA coefficients of U.S. market are not on an obviously declining trend.

Through the DSCF model, cash flow prediction models could be examined for approximating the unobservable market expectation. The cash flow prediction model that results in theoretical stock values fitting the actual share prices better should be seen as a closer process to the market expectation of cash flow. It is not necessary that a better cash flow prediction model will better fit market expectation, but any exception will be a challenge towards efficient market hypothesis. The results of U.S. and U.K. markets suggest that a random walk model has a better fitness for the market expectation of future cash flows. This could be a good implication to investors because there will be great scope for strategy development, exploiting superior cash flow prediction models that may not be taken into account by the market. The theoretical value of a share can be used to determine a stock's expected return and thus predict the price's direction of movement. In the U.S. study, simple portfolios are constructed on conditions that are designed to select undervalued stocks. The portfolios have high returns compared with the contemporary market index. The portfolios are updated at a very low frequency and thus the return will be minimally impacted even if transaction costs are involved. When the superior power of the grey-box model in predicting cash flow is employed to design the investment strategy, the constructed portfolios could have even achieved a better performance. For the China data, an additional test has been conducted on the directions of price movement. The DSCF model has been applied to predict the directions of price movement, which is compared with the result of naive prediction, i.e. all shares will go up (or down). A two-sample *t*-test suggests that the DSCF model could have superior power in making predictions for market prices.

In conclusion, the above comparisons of the empirical results by analysing four different datasets show that it is important to look into special features within each dataset, which might result in different modelling performances and conclusions. Cash flow behaviours of different countries have their own particular patterns, even though some similarities are shared. Even for firms within the same country, the cash flows may be captured by different models according to the firms' operating conditions. It should also be noted that the samples used in this thesis have limitations. For instance, the U.K. unlisted firms only provide 10 years of observations, which restricts the extent to which they can be compared with their listed counterparts. Similarly, it should be noted that the disclosure behaviours of Chinese firms are different from those in U.K. and U.S. markets, mainly in the period of making disclosure. Chinese firms tend to disclose financial information during the first four months of year while U.K and U.S. firms' disclosures span the whole year. As a result, the analysis of Chinese market for the rest 8 months starting from May becomes difficult and mostly uncertain as there are hardly any firms making disclosure during such periods. This is an obstacle of undertaking continuous study and analysis when the stock market is to be studied. It may require more assumptions to validate the study and make further analysis.

Chapter 9 Conclusions, Limitations and Future Work

This thesis aims to develop models to improve firm-level annual cash flow prediction. Cash flow is an important measure of a firm's income. Although the legal disclosure of earnings has a longer history for all registered firms under company law, recent requirement for cash flow disclosure for firms has proven to provide incremental information for investors along with earnings (see Bowen et al., 1987 and Sloan, 1996 for instance). Therefore, making predictions of cash flow is as crucial as earnings.

9.1 Problems Addressed in This Thesis

This thesis starts from reviewing relevant literature in the relationship and informational content of cash flows and earnings, and then critically analyses and discusses the advantages and disadvantages of the models that have been applied to make cash flow prediction. The review of the models has lead to further original derivation and initial application of the novel cash flow models in this thesis. In academia, the models used to predict cash flow are different from those predicting earnings. In the literature, it has been recognised that cash flow has a relationship with other accounting variables. Therefore, the studies using univariate models have been extended to multiple-variables models. Apart from lagging cash flow, the literature has discussed several additional accounting accrual variables that are useful for predicting cash flows. For instance, DKW model suggests that earnings, apart from cash flow, can provide incremental information for future cash flow prediction. Later models such as BCN model disaggregate relevant predictive variables into components according to their accounting relationships and thus use the resulted more detailed model in prediction.

However, there are at least two important aspects in the task of predicting cash flows that the extant literature has not shed light on. First, the heterogeneity across firms may have an impact on the model used to make a prediction. The existing studies mostly apply pooled regression when studying cash flow prediction. Barth et al. (2001) noted that pooled regression in prior studies only examines the sign of the parameters for each variable in the cash flow model and in practice, it may be better to individually model cash flow for each firm. Even so, there is an inevitable problem in individual estimation when the number of parameters exceeds that of observations. Second, the problem ignored by most of the extant literature is the dynamics of cash flow models, i.e. instead of remaining constant, the parameters might be driven by exogenous factors and hence be time-varying. Individual estimation, even being feasible for firms with sufficient observations, does not solve the second problem. Recognising these problems, this thesis selects a particular path to conduct the whole study. To be specific, this study explores the nonlinear and dynamic features of the model components, especially the parameters.

9.2 Methods and Research Design

This thesis has employed panel data models and a Bayesian forecasting model, and has developed a novel grey-box model to improve cash flow prediction modelling. Panel data models calculate individual intercept terms and this thesis selected three different estimation methods, i.e. demean, first difference and the Arellano-Bond estimator. It is expected that individual effects could partially capture the heterogeneity of firms and hence could enhance performance of cash flow models over pooled regression in making individual predictions. A Bayesian model is applied in this thesis because it enables the model parameters to be dynamic, i.e. time-varying. Besides, the Bayesian model also has the power to consider the heterogeneity problem for it estimates, or updates, parameters based on individual observations.

The grey-box model is the major contribution of this thesis. It treats the parameters in the cash flow model as nonlinear processes and uses a Padé approximant to capture the processes. The model is termed a 'grey-box' for it is a combination of a clear-box cash flow model and a black-box parameter process. There are two variables employed separately in the grey-box model to capture the parameters' dynamics, i.e. sales growth rate and the firm's age. The grey-box model is the first attempt to model the parameter dynamics in the cash flow prediction model and show that the dynamics are related to the two selected variables.

The cash flow prediction models are compared according to their performance in the in-sample and out-of-sample periods. Two measures of model performance evaluation are investigated in this thesis, MSE and average rank. MSE is usually selected as a criterion when evaluating and comparing predictive models in various applications. However, this thesis deals with panel data, which is comprised of time-series data of many individuals.

Therefore, a good model is not only required to generate low prediction errors for the whole dataset, but also expected to provide accurate predictions for as many individuals as possible. That is to say, a good model should be both accurate and general. If MSE is used as the only criterion, the generality of models will be overlooked. On the other hand, using two measures in parallel occasionally results in difficulty in reaching a clear conclusion as a model that has the lowest MSE may not necessarily have the lowest average rank as well, which is confirmed by the empirical results in this thesis. A model can be regarded as good when it could achieve leading positions in both criteria.

The second major contribution of this thesis is the extension for the basic DCF model. The DCF model is used to value an asset given the forecast of the cash flows the asset will generate. The use of the DCF model is simple and the model is introduced as a fundamental pricing tool in almost all finance textbooks. There are two new features added to the model so that the DCF model could be jointly used with the cash flow prediction models. The first feature involved is Monte Carlo simulation designed to approximate the asymmetry of cash flow distribution. The second is the expected utility theory applied to bring the many simulated cash flow series to a single value to be discounted. The novel model (DSCF) is applied along with cash flow prediction models, taking the prediction uncertainty as an additional input. The DSCF model could partially eliminate the subjective impact in the traditional DCF model implementation because the cash flow prediction models are estimated on actual data by statistical methods. Based on finance theory that market share prices reflect cash flow expectations that are unobservable the DSCF model could be used to show which cash flow prediction model is closer to the market's cash flow expectations. This thesis then investigates whether superior models for cash flow prediction could have investment value.

9.3 Empirical Evidence

This thesis has made thorough analysis on four different datasets and the selected and developed methods are applied to them. This thesis is the first study to plot the pattern of cash flow distributions along with firm age (or time measure approximating that). Different from previous studies which tend to plot the patterns according to specific years, this thesis aims to show whether there is any underlying natural rules for the evolution of firms' cash flow. The patterns vary across different datasets, as summarised in previous chapter. It is expected that the observation of the cash flow pattern of a particular sample would help in modelling the cash flow process in predictive applications.

Among all the datasets, the grey-box model has proved its superior power over the other applied models. Therefore, it should be recognised that a nonlinear parameter model can capture the real data more accurately. It is a crucial implication that a firm's growth state is a significant factor that should be taken into account when making predictions for cash flow or making other decisions. Knowledge is also obtained from the empirical observation that there exists a natural trend for firms' growth to be closely related to firm age. Therefore, firm age can be used as a proxy variable for firm growth, which can lead to promising results without introducing heavier working burden.

The panel data models and the Bayesian model, unlike the grey-box model, cannot show superior performance in making practical prediction. Despite their features such as taking account of heterogeneity and/or time-varying parameters, the two models do not outperform the simple regression model consistently. It can be seen from the result that a good model is not necessarily complicated. Simpler model may perform better in some certain cases. Secondly, previous studies tend to focus on the in-sample data fitness of relevant models and thus the conclusions might be partial or biased. This thesis examines the models using out-of-sample data and look at predictions of both one period and multiple periods ahead. The model that fit the in-sample data best may not be a good model in practical applications, e.g. making predictions for 2 years ahead. Similarly, a model that generates lower MSE for all sample firms may not be the most general model to describe the cash flow behaviour. It thus can be shown that the research design of this thesis helps in drawing sounder conclusions when comparing the predictive models.

Relating cash flow prediction to stock pricing, the DSCF model in this thesis is designed for the purpose of exploiting the economic value of cash flow models. Besides, the application of the DSCF model could also provide indication about which cash flow models could capture the market expectation better. This topic is relatively new because the market expectation is not observable. Thus, there is hardly any way to measure it effectively. The best model for predicting future cash flow may not necessarily capture the market expectation better. In this thesis, the grey-box model has been shown to be superior to other models. However, it is shown to fit market expectation of future cash flows better only for U.S. markets. For both U.K. and China markets, simpler predictive models describe the market expectation better, even though their predictive performances are not as good as the grey-box model. When it comes to the stock market, the joint application of cash flow prediction models and the DSCF model could produce many possibilities for further study. This thesis conducts some experiments in portfolio construction in order to demonstrate how superior power of cash flow prediction can lead to profit opportunities in the equity market. By analysing the datasets of U.S., U.K. and China markets, this work shows that there exists possible gain once suitable strategies are applied along with the models developed in this thesis. Economic profit is likely to be generated through some well-designed procedure.

9.4 Limitations

A major limitation in cash flow studies is the availability of comprehensive data. It is only 30 years since the statement of cash flow first became compulsory. On the other hand, from the trend of cash flow model development, it can be deduced that the number of variables included is highly likely to increase in the future.

As a result, the availability of a large number of observations to facilitate conducting model estimations for individual firm level data is limited. Therefore, the focus of cash flow prediction studies will be put more on the procedure of looking for common features in different individuals, or in sampling firms that could be treated as homogeneous.

Fundamental factors, such as sectors, operating cycle and so on, may also be studied in the future to determine whether cash flow behaviours of firms are affected by these factors and if so, in what way are they affected.

In the study of equity pricing using DSCF method, portfolios are constructed in Chapter 4 to examine the profit opportunity of the novel method. Although the results are encouraging as the portfolios by certain strategies in this thesis perform well, it should be noted that the risk factors of the portfolios are not taken into account.

9.5 Future Work

Following the analysis of the limitation, heterogeneity across firms should be studied further. The factors that may influence cash flow predictions could be incorporated into the grey-box model as additional input variables.

The grey-box model could also be further developed for incorporating multiple variables into the black box to take account of heterogeneity in empirical observations. One-dimensional heterogeneity, e.g. firm age, will still leave many individuals indifferent. As the structure of the grey-box model in this thesis is relatively simple, sophisticated structures could be developed with the assistance of increasing computing power to suit more situations.

The study of cash flow pattern in this thesis is limited to sampled firms and overlooked some empirical issues. First the study does not consider the effect of inflation. To be specific, inflation can affect growth and hence the growth pattern calculated directly using the accounting data without considering inflation might bias the reported results. Inflation has impact on both cash flows and share pricesand it will be worthwhile to incorporate the effects of inflation in any future study on the modelling.

In addition, this thesis put emphasis on model construction. For such a practical topic, i.e. cash flow prediction, there is very much information, apart from the accounting variables, that is closely related to cash flows of firms. People in practice will collect the information and make analysis accordingly. Future work, to make the study more practical, could take account of the extra information from social media or qualitative information.

In this thesis, the calibration of DSCF model for RRA coefficients is conducted monthly and restricted on the sample firms making disclosure. To make the procedure more precise, the sample restriction could be relaxed and include all the stocks in the market. The calibration procedure could thus be conducted on a daily frequency. It will require prediction of cash flows for all firms and involve much more computation. The output will be a daily index reflecting current market risk aversion, which could guide decision making in investment and trading.

Databases

Chapter 4: WRDS Compustat database (https://wrds-web.wharton.upenn.edu/wrds/) for U.S. firms' accounting variables data and stock market data; Datastream database (subscribed by the University of Glasgow library) for bond yields data.

Chapter 5: Datastream database for U.K. listed firms' accounting data, share price data and bond rates data; FAME database (https://fame2.bvdep.com/ukfederation.aspx) for the data of 'incorporation date'.

Chapter 6: FAME database (https://fame2.bvdep.com/ukfederation.aspx) for U.K. unlisted firms' fundamental data and information.

Chapter 7: RESSET database (http://www3.resset.cn:8080/product/) for Chinese listed firms' accounting data and stock market data; the bond rates are from http://www.chinabond.com.cn/d2s/index.html.

References

Albrecht, W. S., Lookabill, L. L., McKeown, J. C. (1977). The time-series properties of annual earnings. *Journal of Accounting Research*, *15*(2), 226-244.

Badertscher, B. A., Phillips, J. D., Pincus, M., Rego, S. O. (2009). Earnings management strategies and the trade-off between tax benefits and detection risk: To conform or not to conform? *The Accounting Review*, *84*(1), 63–97.

Baginski, S. P., Branson, B. C., Lorek, K. S., Willinger, G. L. (2003). A time-series approach to measuring the decline in quarterly earnings persistence. *Advances in Accounting*, *20*, 23-42.

Bakshi, G., Chen, Z. (2005). Stock valuation in dynamic economies. *Journal of Financial Markets*, *8*(2), 111-151.

Ball, R., Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, *6*(2), 159-178.

Ball, R., Watts, R. (1972). Some time series properties of accounting income. *The Journal of Finance*, *27*(3), 663-681.

Barth, M. E., Cram, D. P., Nelson, K. K. (2001). Accruals and the prediction of future cash flows. *The Accounting Review*, *76*(1), 27-58.

Bathke Jr, A. W., Lorek, K. S., Willinger, G. L. (1989). Firm-size and the predictive ability of quarterly earnings data. *Accounting Review*, *64*(1), 49-68.

Beaver, W. H. (1970). The time series behavior of earnings. *Journal of Accounting Research*, *8,* 62-99.

Beaver, W. H., Dukes, R. E. (1972). Interperiod tax allocation, earnings expectations, and the behavior of security prices. *Accounting Review*, 47(2), 320-332.

Bernard, V., Noel, J. (1991). Do inventory disclosures predict sales and earnings? *Journal of Accounting, Auditing & Finance*, *6*(2), 145-181.

Bernard, V. L., Stober, T. L. (1989). The nature and amount of information in cash flows and accruals. *Accounting Review*, *64*(4), 624-652.

Bowen, R. M., Burgstahler, D., Daley, L. A. (1986). Evidence on the relationships between earnings and various measures of cash flow. *Accounting Review*, *61*(4), 713-725.

Bowen, R. M., Burgstahler, D., Daley, L. A. (1987). The incremental information content of accrual versus cash flows. *Accounting Review, 62*(4), 723-747.

Box, G. E., Jenkins, G. M., Reinsel, G. C. (2013). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brochet, F., Nam, S., Ronen, J. (2008). The role of accruals in predicting future cash flows and stock returns. Working Paper, available at SSRN 1126022 .

Brown, L. D., Rozeff, M. S. (1979). Univariate time-series models of quarterly accounting earnings per share: A proposed model. *Journal of Accounting Research*, *17*(1), 179-189.

Burgstahler, D., Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, *24*(1), 99-126.

Cameron, A. C., Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. New York, NY: Cambridge University Press.

Chan, L. K., Karceski, J., Lakonishok, J. (2003). The level and persistence of growth rates. *The Journal of Finance*, *58*(2), 643-684.

Cheng, C. A., Hollie, D. (1996). The Persistence of Cash Flow Components into Future Cash Flows. Working Paper, available at SSRN 541293.

Cheng, C. A., 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*(1), 29-53.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Aystems*, *2*(4), 303-314.

Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics*, *18*(1), 3-42.

Dechow, P. M., Ge, W., 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies*, *11*(2), 253–296.

Dechow, P. M., Kothari, S. P., L Watts, R. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics*, *25*(2), 133-168.

Dechow, P. M., Richardson, S. A., Tuna, I. (2003). Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies*, *8*(2), 355-384.

Dechow, P. M., Sloan, R. G., Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review*,*70*(2), 193–225.

DeFond, M. L., Hung, M. (2003). An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics*, *35*(1), 73-100.

DeFond, M. L., Jiambalvo, J. (1994). Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics*, *17*(1), 145-176.

DeFond, M. L., Park, C. W. (1997). Smoothing income in anticipation of future earnings. *Journal of Accounting and Economics*, *23*(2), 115-139.

DeFond, M. L., Park, C. W. (2001). The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review*, *76*(3), 375-404.

Dichev, I. D., Tang, V. W. (2008). Matching and the changing properties of accounting earnings over the last 40 years. *The Accounting Review*, *83*(6), 1425-1460.

Dichev, I. D., Tang, V. W. (2009). Earnings volatility and earnings predictability. *Journal of Accounting and Economics*, *47*(1), 160-181.

Dopuch, N., Watts, R. (1972). Using time-series models to assess the significance of accounting changes. *Journal of Accounting Research*, *10*(1), 180-194.

Drtina, R. E., Largay III, J. A. (1985). Pitfalls in calculating cash flow from operations. *Accounting Review*, *60*(2), 314-326.

Farshadfar, S., Monem, R. (2011). Discretionary accruals and the predictive ability of earnings in the forecast of future cash flows: Evidence from Australia. *Corporate Ownership and Control*, *9*(1), 597-608.

Farshadfar, S., Monem, R. (2013). The usefulness of operating cash flow and accrual components in improving the predictive ability of earnings: a re-examination and extension. *Accounting & Finance*, *53*(4), 1061-1082.

Farshadfar, S., Monem, R. (2013b). Further evidence on the usefulness of direct method cash flow components for forecasting future cash flows. *The International Journal of Accounting*, *48*(1), 111-133.

Finger, C. A. (1994). The ability of earnings to predict future earnings and cash flow. *Journal of Accounting Research*, *32*(2), 210-223.

Foster, G. (1977). Quarterly accounting data: Time-series properties and predictive-ability results. *Accounting Review*, *52*(1), 1-21.

Francis, R. N., Eason, P. (2012). Accruals and the naïve out-of-sample prediction of operating cash flow. *Advances in Accounting*, *28*(2), 226-234.

Francis, R. N. (2011). Out-of-sample cash flow prediction and cash distributions to shareholders. *Advances in Accounting*, *27*(1), 1-9.

Francis, J., Olsson, P., Oswald, D. R. (2000). Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. *Journal of Accounting Research*, *38*(1), 45-70.

Fudenberg, D., Tirole, J. (1995). A theory of income and dividend smoothing based on incumbency rents. *Journal of Political Economy*, *103*(1), 75-93.

Fuller, R. J., Huberts, L. C., Levinson, M. J. (1992). It's not higgledy-piggledy growth! *The Journal of Portfolio Management*, *18*(2), 38-45.

Gordon, M. J., Shapiro, E. (1956). Capital equipment analysis: the required rate of profit. *Management Science*, *3*(1), 102-110.

Gordon, M. J. (1959). Dividends, earnings, and stock prices. *The Review of Economics and Statistics*, *41*(2), 99-105.

Greenberg, R. R., Johnson, G. L., Ramesh, K. (1986). Earnings versus cash flow as a predictor of future cash flow measures. *Journal of Accounting, Auditing & Finance*, *1*(4), 266-277.

Griffin, P. A. (1977). The time-series behavior of quarterly earnings: Preliminary evidence. *Journal of Accounting Research*, *15*(1), 71-83.

Hales, J., Orpurt, S. F. (2013). A Review of Academic Research on the Reporting of Cash Flows from Operations. *Accounting Horizons*, *27*(3), 539-578.

Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, *7*(1), 85-107.

Holthausen, R. W., Larcker, D. F., Sloan, R. G. (1995). Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics*, *19*(1), 29-74.

Hopwood, W. S., McKeown, J. C. (1992). Empirical evidence on the time-series properties of operating cash flows. *Managerial Finance*, *18*(5), 62-78.

Hribar, P., Collins, D. W. (2002). Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research*, *40*(1), 105-134.

Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, *29*(2), 193–228.

Kalman, Rudolph Emil. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35-45.

Kaplan, S. N., & Ruback, R. S. (1995). The valuation of cash flow forecasts: An empirical analysis. *The Journal of Finance*, *50*(4), 1059-1093.

Ketz, J. E., J.A. Largay. (1987). Reporting income and cash flows from operations. *Accounting Horizons*, *1*(2), 9-17.

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*(5), 753-780.

Kothari, S. P., Leone, A. J., Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, *39*(1), 163–197.

Lee, C. W. J., Chen, C. (1990). Structural changes and the forecasting of quarterly accounting earnings in the utility industry. *Journal of Accounting and Economics*, *13*(2), 93-122.

Lev, B., Li, S., Sougiannis, T. (2010). The usefulness of accounting estimates for predicting cash flows and earnings. *Review of Accounting Studies*, *15*(4), 779-807.

Lipe, R. C. (1986). The information contained in the components of earnings. *Journal of Accounting Research*, *24*, 37-64.

Little, I. M. (1962). Higgledy piggledy growth. *Bulletin of the Oxford University Institute of Economics & Statistics*, *24*(4), 387-412.

Livnat, J., Zarowin, P. (1990). The incremental information content of cash-flow components. *Journal of Accounting and Economics*, *13*(1), 25-46.

Lookabill, L. L. (1976). Some additional evidence on the time series properties of accounting earnings. *Accounting Review*, *51*(4), 724-738.

Lorek, K. S., Schaefer, T. F., Willinger, G. L. (1993). Time-series properties and predictive ability of funds flow variables. *Accounting Review*, *68*(1), 151-163.

Lorek, K. S., Willinger, G. L. (1996). A multivariate time-series prediction model for cash-flow data. *Accounting Review*, *71*(1), 81-102.

Lorek, K. S., Willinger, G. L. (2008). Time-series properties and predictive ability of quarterly cash flows. *Advances in Accounting*, *24*(1), 65-71.

Lorek, K. S., Willinger, G. L. (2010). Time series versus cross-sectionally derived predictions of future cash flows. *Advances in Accounting*, *26*(1), 29-36.

Manegold, J. G. (1981). Time-series properties of earnings: a comparison of extrapolative and component models. *Journal of Accounting Research*, *19*(2), 360-373.

McInnis, J., Collins, D. W. (2011). The effect of cash flow forecasts on accrual quality and benchmark beating. *Journal of Accounting and Economics*, *51*(3), 219-239.

Nelson, C. R., Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business*, *60*(4), 473-489.

Ogneva, M. (2012). Accrual quality, realized returns, and expected returns: The importance of controlling for cash flow shocks. *The Accounting Review*, *87*(4), 1415–1444.

Orpurt, S. F., Zang, Y. (2009). Do direct cash flow disclosures help predict future operating cash flows and earnings? *The Accounting Review*, *84*(3), 893-935.

Pae, J., Yoon, S.-S. (2012). Determinants of analysts cash flow forecast accuracy. *Journal of Accounting, Auditing & Finance*, *27*(1), 123–144.

Pástor, Ľ., Pietro, V. (2003). Stock valuation and learning about profitability. *The Journal of Finance*, *58*(5), 1749-1790.

Pinheiro, J., Bates, D. (2006). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

Rayburn, J. (1986). The association of operating cash flow and accruals with security returns. *Journal of Accounting Research*, *24,* 112-133.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, *42*(3), 335–370.

Salamon, G. L., Smith, E. D. (1977). Additional evidence on the time series properties of reported earnings per share: Comment. *The Journal of Finance*, *32*(5), 1795-1801.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425-442.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1-48.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, *71*(3), 289-315.

Tan, K. C., Li, Y. (2002). Grey-box model identification via evolutionary computing. *Control Engineering Practice*, *10*(7), 673-684.

Watts, R. L., Leftwich, R. W. (1977). The time series of annual accounting earnings. *Journal of Accounting Research*, *15*(2), 253-271.

West, M., Harrison, J. (1997). *Bayesian forecasting and dynamic models*. New York, NY: Springer.

Williams, J. B. (1938). *The theory of investment value*. Cambridge, MA: Harvard University Press.

Wilson, G. P. (1987). The Incremental Information Content of the Accrual and Funds Components of Earnings after Controlling for Earnings. *Accounting Review*, *62*(2), 293-322.

Appendices

A. A Method of Discounting Predicted Cash Flows

Discounted cash flow (DCF) is a traditional method of asset evaluation, presuming the value of an asset (*V*) is the summation of cash flows (*CF*) it generates in its life, discounted at particular rates (*r*):

$$V_t = \sum_{i=t+1}^{T} \frac{CF_i}{(1+r_i)} \qquad \text{(A1)}$$

In stock valuation, there are challenges in applying this DCF model. On one hand, cash flows need to be predicted to reflect the asset's income streams. On the other hand, the discount rate also requires specification. For equities there is no maturity *T* as firms are considered to survive forever. Infinite numbers of cash flows predictions therefore are required to enter the model. People who apply it empirically will make subjective assumptions such as 'in five years the cash flow will stay at a constant level' and so on. For such reasons, this model has remained on the theoretical level.

Cash flow risk, as a by-product of cash flow modelling, should have an impact on the discount rate of the DCF model. A model of the relationship between the discount rate and cash flow risk will be derived below.

1. a. Assume there is a utility function in the market for cash flow $u(c)$;
   b. Assume the predicted cash flow follows normal distribution in the form of $c \sim N(\bar{c}, \sigma^2)$; c. Assume cash flow risk is reflected by its variance σ^2.

2. According to expected utility theory, calculate the expected utility of the risky cash flow
   $$E(u(c)) = \int p(c)u(c)dc \qquad \text{(A2)}$$
   where *p(c)* denotes the probability of having cash flow *c*.
   There will be a certainty equivalent cash flow c^* that provides the same expected utility:

$$u(c^*) = E(u(c)) \tag{A3}$$

c^* could be seen as the price the representative investor is willing to pay to get the risky cash flow income c.

3. According to Pratt (1964)[3], the certainty equivalent cash flow is approximately:

$$c^* = \bar{c} - \frac{\sigma^2}{2}\left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \tag{A4}$$

The term in the bracket is the coefficient of absolute risk aversion (ARA)[4].

4. The utility function should meet the condition:

$$u(\mathrm{n}\,c^*) = E(u(\mathrm{n}\,c)) \tag{A5}$$

where n can be any positive number. This condition intuitively means that for buying n shares of some firm, the total price will be n times the price of one share. Therefore:

$$nc^* = n\bar{c} - \frac{(n\sigma)^2}{2}\left(-\frac{u''(\mathrm{n}\,\bar{c})}{u'(\mathrm{n}\,\bar{c})}\right), \tag{A6}$$

Recall in 3,

$$\begin{aligned} c^* &= \bar{c} - \frac{\sigma^2}{2}\left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \\ nc^* &= n\bar{c} - n\frac{\sigma^2}{2}\left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \end{aligned} \tag{A7}$$

Therefore,

$$\begin{aligned} n\bar{c} - \frac{(n\sigma)^2}{2}\left(-\frac{u''(\mathrm{n}\,\bar{c})}{u'(\mathrm{n}\,\bar{c})}\right) &= n\bar{c} - n\frac{\sigma^2}{2}\left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \\ n \times \left(-\frac{u''(\mathrm{n}\,\bar{c})}{u'(\mathrm{n}\,\bar{c})}\right) &= -\frac{u''(\bar{c})}{u'(\bar{c})} \\ n\bar{c} \times \left(-\frac{u''(\mathrm{n}\,\bar{c})}{u'(\mathrm{n}\,\bar{c})}\right) &= \bar{c} \times \left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \end{aligned} \tag{A8}$$

[3] Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*(1/2) 122-136.

[4] See footnote 3 and also Arrow, K. J. (1971). Essays in the Theory of Risk Bearing, Chicago: Markham Publ. Co., 90–109.

This expression implies the utility function to have the feature of constant relative risk aversion (CRRA).

5. The isoelastic utility function:

$$u(c) = \frac{c^{1-\rho}}{1-\rho} \tag{A9}$$

exhibits constant relative risk aversion (CRRA) with

$$\bar{c} \times \left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) = \rho \tag{A10}$$

With this utility function, we could calculate certainty equivalent cash flow using

$$\begin{aligned} c^* &= \bar{c} - \frac{\sigma^2}{2}\left(-\frac{u''(\bar{c})}{u'(\bar{c})}\right) \\ &= \bar{c} - \frac{\rho\sigma^2}{2\bar{c}} \end{aligned} \tag{A11}$$

and simply discount the certainty equivalent cash flow by risk free interest rates.

B. Tables

Ta**ble 2.1: Comparison of the models in the studies of cash flow prediction.**

Paper	Greenberg1986	Lorek1993	Finger1994	Lorek1996	DKW1998	BCN2001	Cheng2008	Farshadfar 2011
model	single regression	ARX/ARIMA[a]	ARX	ARX	ARX/na ïve	ARX/pooled regression	ARX/pooled regression	ARX/pooled regression
TS/CS[b]	TS	CS/TS	TS	TS	TS	CS	CS	CS
frequency	annual	quarterly	annual	quarterly	annual	annual	annual	annual
data	Compustat U.S	Compustat U.S	Compustat U.S	Compustat U.S	Compustat U.S	Compustat U.S	Compustat U.S	Australian
out-of-sample		y	y	y	y		y	
multiyear-ahead	y		y		y			
Variables as predictors	income	net income	earnings	operating income	earnings	d(account payable)[c] d(account receivable) d(inventory) depreciation amortisation other component in earnings	d(account payable); d(account receivable); d(inventory); Depreciation; amortisation other component in earnings	discretionary accruals; non-discretionary accruals
	or	and	and/or		and/or			
	cash flow	cash flow	cash flow	cash flow	cash flow	cash flow	cash flow from sales;	cash flow

							cash flow from cost; cash flow from operating expense; cash flow interest; cash flow tax; other cash flow	
				account payable; account receivable; inventory				

Note: a. ARX for autoregressive model with exogenous variables; ARIMA for autoregressive integrated moving average model.
b. TS for time series analysis; CS for cross sectional analysis.
c. d() indicates the changes in relevant variable

Table 4.1: Descriptive statistics for the variables in cash flow prediction model (U.S. listed firms).

	CF	*DA*	*ΔAR*	*ΔAP*	*ΔINV*	*OTHER*
Mean	0.0683	0.0459	0.0142	0.0077	0.0094	0.0214
Standard Error	0.0004	0.0001	0.0002	0.0002	0.0002	0.0003
Median	0.0631	0.0359	0.0068	0.0033	0.0008	0.0196
Mode	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Standard Deviation	0.1318	0.0461	0.0722	0.0483	0.0592	0.0892
Sample Variance	0.0174	0.0021	0.0052	0.0023	0.0035	0.0080
Kurtosis	26.6543	122.0470	34.8779	160.6502	148.3367	165.1803
Skewness	-0.1282	6.8188	0.7104	2.0117	2.2320	-2.1291
Range	6.3604	1.6376	2.8081	4.1359	5.3354	7.4330
Minimum	-1.8893	-0.0034	-1.3301	-1.9632	-2.2495	-4.9773
Maximum	4.4711	1.6342	1.4779	2.1727	3.0859	2.4557

Note: The variables are defined as follows:

CF : Net operating cash flow;

DA: Depreciation and amortisation;

ΔAR : Changes in account receivable;

ΔAP : Changes in account payable;

ΔINV : Changes in inventory;

OTHER: Other accruals, calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – ($CF + \Delta AR + \Delta INV - \Delta AP - DA$).

All variables are deflated by average total assets of each firm.

Table 4.2: The cash flow prediction model parameters estimated with panel data methods (U.S. firms).

	Pooled Regression		Demean	Difference	A-B
	(1987-1996)	(1987-2005)	(1987-2005)	(1987-2005)	(1987-2005)
β_1	0.613***	0.692***	0.392***	-0.303***	0.197***
	(39.94)	(82.46)	(33.02)	(-20.75)	(6.55)
β_2	0.247***	0.323***	0.273***	0.0785***	0.0742*
	(10.81)	(17.61)	(15.68)	(4.70)	(2.51)
β_3	-0.534***	-0.594***	-0.491***	-0.240***	-0.292***
	(-13.05)	(-23.24)	(-18.90)	(-12.05)	(-8.10)
β_4	0.399***	0.447***	0.416***	0.199***	0.195***
	(19.49)	(27.53)	(22.17)	(11.49)	(6.71)
β_5	0.253***	0.140***	0.293***	0.0581	0.0374
	(8.40)	(7.06)	(7.99)	(1.72)	(0.73)
β_6	0.290***	0.351***	0.338***	0.142***	0.132***
	(16.35)	(24.10)	(19.53)	(7.99)	(4.47)
β_0	0.00336**	0.00439***			
	(3.17)	(5.27)			
N	27630	63120	63120	52884	52884

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model estimated:

$$CF_{i,t+1} = \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} + \beta_5 DA_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon_{i,t+1}$$

See Table 4.1 for variable definitions.

Table 4.3: The covariance-correlation matrix of parameters in model (4.2) estimated individually for available sample firms (U.S. firms).

	β_0	β_1	β_2	β_3	β_4	β_4	β_6
β_0	0.0017	-0.0034	0.0012	0.0021	-0.0058	-0.0558	-0.0026
β_1	**-0.1596**	0.2595	0.0345	-0.2089	0.1889	-0.5379	0.1153
β_2	**0.0079**	**0.0184**	13.5099	-0.0877	-0.5180	-0.2613	0.0606
β_3	**0.0358**	**-0.2847**	**-0.0166**	2.0747	-0.5809	0.4516	-0.1606
β_4	**-0.1320**	**0.3521**	**-0.1338**	**-0.3830**	1.1088	-0.4566	0.2339
β_5	**-0.5583**	**-0.4392**	**-0.0296**	**0.1304**	**-0.1804**	5.7806	-0.5902
β_6	**-0.0924**	**0.3337**	**0.0243**	**-0.1645**	**0.3276**	**-0.3621**	0.4597

Notes: 1. Parameters are defined in model (see Table 4.1 for variable definitions): $CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$ 2. Bold numbers (lower triangle) are correlation coefficients while the plain numbers (upper triangle) are covariance and variance on the diagonal.

Table 4.4: The comparison of the means of the parameter distributed among different individuals and the variances of residuals obtained by two methods (U.S. firms).

	Individual Estimation	Random Parameter Model
β_0	0.008845	0.006972
β_1	0.513972	0.596072
β_2	0.207829	0.284259
β_3	-0.68456	-0.71851
β_4	0.629093	0.479734
β_5	0.584838	0.226338
β_6	0.419901	0.336922
$\hat{\sigma}^2(\varepsilon)$	0.002682	0.0035

Notes: Parameters are defined in model (see Table 4.1 for variable definitions):

$$CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

Table 4.5: The covariance-correlation matrix of cash flow model parameters according to random parameter model (U.S. firms).

	β_0	β_1	β_2	β_3	β_4	β_5	β_6
β_0	0.0002	-0.0023	-0.0035	0.0030	-0.0018	-0.0008	-0.0042
β_1	**-0.5029**	0.1366	0.0759	-0.0938	0.0479	-0.1403	0.0469
β_2	**-0.4567**	**0.3252**	0.3992	-0.3167	-0.0418	-0.0679	0.0899
β_3	**0.2264**	**-0.2365**	**-0.4669**	1.1524	-0.2550	0.1542	-0.1253
β_4	**-0.2304**	**0.2037**	**-0.1038**	**-0.3730**	0.4055	-0.0701	0.1132
β_5	**-0.0979**	**-0.5448**	**-0.1542**	**0.2062**	**-0.1580**	0.4857	-0.1201
β_6	**-0.6990**	**0.2602**	**0.2916**	**-0.2393**	**0.3645**	**-0.3532**	0.2380

Notes: 1. Parameters are defined in model (see Table 4.1 for variable definitions):
$CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$ 2. Bold numbers (lower triangle) are correlation coefficients while the plain numbers (upper triangle) are covariance and variance on the diagonal.

Table 4.6: The Bayesian update of parameters vector for AAR Corp with monitoring.

Year	β_0	β_1	β_2	β_3	β_4	β_5	β_6	$\hat{\sigma}^2(\varepsilon)$
Prior	0.0070	0.5961	0.2843	-0.7185	0.4797	0.2263	0.3369	0.0035
1989	0.0070	0.5960	0.2777	-0.6998	0.4545	0.2299	0.3306	0.00296
1990	0.0068	0.6037	0.3084	-0.7248	0.4504	0.2187	0.3374	0.002549
1991	0.0068	0.5763	0.3272	-0.7269	0.4408	0.2565	0.3414	0.002324
1992	0.0068	0.5766	0.3252	-0.7305	0.4428	0.2565	0.3410	0.002066
1993	0.0067	0.5751	0.2946	-0.6627	0.4846	0.2740	0.3445	0.001888
1994	0.0066	0.5756	0.2966	-0.6678	0.4837	0.2739	0.3442	0.001717
1995	0.0066	0.5761	0.3034	-0.6654	0.4733	0.2729	0.3466	0.001574
1996	0.0070	0.5644	0.3007	-0.7477	0.4943	0.2889	0.3281	0.001481
1997	0.0070	0.5643	0.2980	-0.7497	0.4947	0.2881	0.3282	0.001375
1998	0.0072	0.5605	0.2930	-0.7729	0.4880	0.2861	0.3265	0.001284
1999	0.0114	0.5005	0.0921	-0.7290	0.5154	0.3337	0.1699	0.001254
2000	0.0100	0.5071	0.1977	-0.8946	0.5149	0.3215	0.2317	0.001188
2001	0.0100	0.5071	0.1977	-0.8946	0.5149	0.3215	0.2317	0.001188
2002	0.0093	0.4544	0.1495	-0.8809	0.5352	0.3607	0.3117	0.001127
2003	0.0105	0.4603	0.2390	-1.0577	0.5986	0.2343	0.3190	0.001086
2004	0.0097	0.4511	0.2143	-1.1175	0.7138	0.3036	0.3356	0.001037
2005	0.0097	0.4511	0.2143	-1.1175	0.7138	0.3036	0.3356	0.001284

Notes: Parameters are defined in model (see Table 4.1 for variable definitions):

$$CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

Table 4.7: The Bayesian update of parameters vector for AAR Corp without monitoring.

Year	β_0	β_1	β_2	β_3	β_4	β_5	β_6	$\hat{\sigma}^2(\varepsilon)$
Prior	0.0070	0.5961	0.2843	-0.7185	0.4797	0.2263	0.3369	0.0035
1989	0.0070	0.5960	0.2777	-0.6998	0.4545	0.2299	0.3306	0.00296
1990	0.0068	0.6037	0.3084	-0.7248	0.4504	0.2187	0.3374	0.002549
1991	0.0068	0.5763	0.3272	-0.7269	0.4408	0.2565	0.3414	0.002324
1992	0.0068	0.5766	0.3252	-0.7305	0.4428	0.2565	0.3410	0.002066
1993	0.0067	0.5751	0.2946	-0.6627	0.4846	0.2740	0.3445	0.001888
1994	0.0066	0.5756	0.2966	-0.6678	0.4837	0.2739	0.3442	0.001717
1995	0.0066	0.5761	0.3034	-0.6654	0.4733	0.2729	0.3466	0.001574
1996	0.0070	0.5644	0.3007	-0.7477	0.4943	0.2889	0.3281	0.001481
1997	0.0070	0.5643	0.2980	-0.7497	0.4947	0.2881	0.3282	0.001375
1998	0.0072	0.5605	0.2930	-0.7729	0.4880	0.2861	0.3265	0.001284
1999	0.0114	0.5005	0.0921	-0.7290	0.5154	0.3337	0.1699	0.001254
2000	0.0100	0.5071	0.1977	-0.8946	0.5149	0.3215	0.2317	0.001188
2001	0.0155	0.1874	-0.0450	0.4299	0.0085	0.4429	0.2391	0.001531
2002	0.0154	0.1814	-0.0506	0.4294	0.0119	0.4475	0.2490	0.001451
2003	0.0177	0.1778	0.0784	0.2177	0.0853	0.2562	0.2573	0.001419
2004	0.0164	0.1552	0.0317	0.1435	0.2736	0.3772	0.2847	0.001366
2005	0.0250	0.0468	0.0726	-0.2686	0.2930	-0.2201	0.3722	0.001632

Notes: Parameters are defined in model (see Table 4.1 for variable definitions):

$$CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

Table 4.8: The Bayesian update of parameters vector for AAR Corp with naive prior and monitoring.

Year	β_0	β_1	β_2	β_3	β_4	β_5	β_6	$\hat{\sigma}^2(\varepsilon)$
Prior	0	1	1	-1	1	0	0	0.01
1989	-0.0278	1.0000	0.9994	-1.0001	0.9990	-0.0002	-0.0003	0.005007
1990	-0.0151	1.0161	1.0031	-1.0030	0.9747	0.0005	0.0046	0.003427
1991	-0.0190	1.0024	1.0166	-1.0019	0.9818	-0.0003	0.0041	0.002587
1992	-0.0134	0.9893	1.0245	-0.9925	0.9712	0.0008	0.0121	0.00212
1993	-0.0100	0.9877	1.0125	-1.0059	0.9477	0.0001	0.0153	0.001789
1994	-0.0061	0.9740	1.0005	-0.9802	0.9729	-0.0001	0.0241	0.001568
1995	-0.0067	0.9737	1.0042	-0.9795	0.9670	-0.0010	0.0284	0.001373
1996	-0.0074	0.9688	1.0044	-0.9822	0.9710	-0.0022	0.0229	0.001223
1997	-0.0102	0.9794	0.9179	-1.0157	0.9517	-0.0127	0.0447	0.00114
1998	-0.0108	0.9353	0.8255	-1.1029	0.8582	-0.0291	0.0656	0.00111
1999	-0.0103	0.8911	0.7491	-1.1165	0.8952	-0.0571	-0.0451	0.001074
2000	-0.0079	0.8451	0.8064	-1.2189	0.7906	-0.0262	0.0875	0.001052
2001	-0.0204	0.5052	1.0318	-0.8756	0.8317	-0.1379	0.2312	0.001372
2002	-0.0135	0.1817	0.7525	-0.8676	0.8311	-0.0711	0.6204	0.001484
2003	-0.0125	0.2873	0.6255	-0.7814	0.8114	0.0140	0.6370	0.001445
2004	-0.0082	0.2507	0.5440	-0.8019	0.9859	0.1059	0.6234	0.001404
2005	-0.0074	-0.0055	0.5698	-0.9487	0.9553	-0.0894	0.6537	0.001523

Notes: Parameters are defined in model (see Table 4.1 for variable definitions):

$$CF_{i,t+1} = \beta_{i,0} + \beta_{i,1}CF_{i,t} + \beta_{i,2}\Delta INV_{i,t} + \beta_{i,3}\Delta AP_{i,t} + \beta_{i,4}\Delta AR_{i,t} + \beta_{i,5}DA_{i,t} + \beta_{i,6}OTHER_{i,t} + \varepsilon_{i,t+1}$$

Table 4.9: The in-sample fitness of various cash flow models (U.S. firms).

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0101	0.0125	0.0076	0.0053	0.0078	0.0057	0.0147	0.0088	0.0075	0.0075
Av. Rank	4.9777	5.5063	5.6012	5.8425	5.9641	5.8913	5.8367	5.3475	5.2823	4.7504

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 7: Bayesian model with naive prior;
Model 8: Bayesian model with prior from random parameter model;
Model 9: grey-box model using sales growth rates;
Model 10: grey-box model using firm age.

Table 4.10: The out-of-sample one-period-ahead cash flow prediction performance of various models (U.S. firms).

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0150	0.0150	0.0113	0.0140	0.0376	0.0198	0.0150	0.0126	**0.0111**	**0.0109**
Av. Rank	**4.9476**	5.4358	5.8546	6.1324	6.2193	5.8686	5.5645	5.0467	5.3747	**4.5557**

Note: Bold numbers indicate two models that perform best in that criterion.
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 7: Bayesian model with naive prior ;
Model 8: Bayesian model with prior from random parameter model;
Model 9: grey-box model using sales growth rates;
Model 10: grey-box model using firm age.

Table 4.11: The out-of-sample multi-period-ahead cash flow prediction performance of various models (U.S. firms).

	Models	1	2	3	4	5	6	8	10
2006~ 2012	MSE	0.0205	0.0208	**0.0190**	0.0217	0.0262	0.0239	0.0422	**0.0177**
	Av. Rank	**4.3647**	4.3991	4.5984	4.6508	4.3927	4.4868	4.8149	**4.2927**
2007~ 2012	MSE	0.0227	0.0227	**0.0214**	0.0244	0.0278	0.0268	0.0497	**0.0199**
	Av. Rank	4.4292	**4.3701**	4.5719	4.6276	4.3770	4.4751	4.8152	**4.3338**

Note: Bold numbers indicate two models that perform best in that criterion.
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 8: Bayesian model with prior from random parameter model;
Model 10: grey-box model using firm age as additional input.

Table 4.12: The comparison of expected and actual price movement directions by using DKW and grey-box models respectively along with the cash flow discounting model (U.S. data).

DKW	Actual up	Actual down	Sum
Expected up	2848	2685	5533
Expected down	5507	5327	10834
Sum	8355	8012	
Grey-box	Actual up	Actual down	Sum
Expected up	2966	2736	5702
Expected down	5389	5276	10665
Sum	8355	8012	

Note: figures in the table are the number of observations in the particular categories.

Table 4.13: The number of stocks selected by the first investment strategy (using DKW model to predict cash flows) and the average one-year return of the long-position portfolio (U.S. data).

	2006		2007		2008		2009		2010		2011	
	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return
Jan			46	-7.22%	10	**-13.35%**	32	**84.45%**	40	**30.32%**	48	**24.92%**
Feb	10	**19.67%**	14	-16.91%	6	**-33.82%**	10	**104.16%**	9	3.59%	11	**17.31%**
Mar	72	**13.72%**	76	**-0.65%**	51	**-25.68%**	39	**94.47%**	52	**23.71%**	47	**19.83%**
Apr	20	**24.60%**	23	-9.76%	16	**-18.37%**	12	**70.20%**	19	16.20%	15	**5.71%**
May	20	**26.55%**	14	**-4.71%**	4	**0.01%**	13	**28.52%**	11	18.57%	13	-4.45%
Jun	72	**18.97%**	84	-17.33%	8	-25.61%	60	**23.37%**	56	**29.51%**	51	-9.21%
Jul	5	2.74%	17	**-3.71%**	11	-24.60%	15	**32.89%**	13	**25.42%**	14	-3.28%
Aug	21	**28.85%**	24	**-14.87%**	15	-17.71%	13	7.34%	13	**26.42%**	16	13.24%
Sep	85	**20.34%**	66	**-23.56%**	64	**11.56%**	60	**58.81%**	42	-4.87%	51	**26.88%**
Oct	22	1.01%	22	**-33.25%**	21	19.47%	16	**21.47%**	16	-7.09%	14	**25.39%**
Nov	8	**31.04%**	14	**-37.04%**	12	**81.79%**	8	**23.39%**	8	-9.25%	6	**22.33%**
Dec	514	**8.64%**	922	**-39.33%**	382	**71.58%**	630	**20.08%**	629	-6.85%	548	12.39%

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 4.14: Comparison of annual payoffs between the strategies of investing in the portfolios of under-priced stocks (DKW model) and investing in the contemporary S & P index.

	Year	2006	2007	2008	2009	2010	2011	2012
Jan	Portfolio 1		1	0.9278	**0.8039**	**1.4829**	**1.9325**	**2.4141**
	S&P Index		1	0.9458	0.5225	0.7851	0.9434	0.9708
Feb	Portfolio 2	1	**1.1967**	0.9944	**0.6581**	**1.3436**	**1.3919**	**1.6328**
	S&P Index	1	1.0973	1.0215	0.6162	0.9031	1.0239	1.0877
Mar	Portfolio 3	1	**1.1372**	**1.1298**	**0.8396**	**1.6328**	**2.0199**	**2.4205**
	S&P Index	1	1.1311	1.0572	0.6660	0.9054	1.0404	1.0666
Apr	Portfolio 4	1	**1.2460**	**1.1243**	**0.9178**	**1.5620**	**1.8151**	**1.9188**
	S&P Index	1	1.2051	1.1026	0.7237	0.8577	1.0591	1.0317
May	Portfolio 5	1	**1.2655**	**1.2058**	**1.2060**	**1.5499**	**1.8378**	**1.7560**
	S&P Index	1	1.1836	1.0077	0.7238	0.8115	1.0397	1.0724
Jun	Portfolio 6	1	**1.1897**	0.9835	0.7316	**0.9026**	**1.1689**	1.0613
	S&P Index	1	1.1399	0.9927	0.7735	0.8629	1.0122	1.0804
Jul	Portfolio 7	1	1.0274	**0.9892**	0.7459	**0.9913**	**1.2433**	**1.2025**
	S&P Index	1	1.1305	0.9839	0.7828	0.8048	0.9349	1.0788
Aug	Portfolio 8	1	**1.2885**	**1.0969**	**0.9026**	**0.9688**	**1.2247**	**1.3868**
	S&P Index	1	1.1429	0.8731	0.7913	0.8543	0.8470	1.0785
Sep	Portfolio 9	1	**1.2034**	**0.9199**	**1.0262**	**1.6298**	**1.5505**	**1.9673**
	S&P Index	1	1.1244	0.7030	0.7520	0.8587	0.9095	1.0248
Oct	Portfolio 10	1	1.0101	**0.6742**	**0.8055**	**0.9785**	**0.9091**	**1.1399**
	S&P Index	1	1.0575	0.6399	0.7822	0.8429	0.8903	1.0111
Nov	Portfolio 11	1	**1.3104**	**0.8250**	**1.4998**	**1.8506**	**1.6795**	**2.0545**
	S&P Index	1	1.0353	0.6369	0.7862	0.8867	0.8867	1.0056
Dec	Portfolio 12	1	**1.0864**	**0.6591**	**1.1310**	**1.3581**	**1.2651**	**1.4218**
	S&P Index	1	0.9585	0.5742	0.7467	0.8942	0.9125	1.0416

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 4.15: The number of stocks selected by the first investment strategy (using grey-box model to predict cash flows) and the average one-year return of the long-position portfolio (U.S. data).

	2006		2007		2008		2009		2010		2011	
	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return
Jan			53	-26.29%	48	**-44.47%**	30	**84.06%**	39	**32.54%**	26	**10.79%**
Feb	13	**19.83%**	12	-30.73%	5	**-24.67%**	6	35.40%	7	-5.89%	6	-0.73%
Mar	55	**11.81%**	60	-10.40%	48	-37.70%	29	**79.95%**	39	**15.37%**	47	-2.97%
Apr	17	**20.74%**	17	-9.28%	15	**-29.47%**	5	**56.90%**	12	2.37%	11	**3.65%**
May	22	**22.95%**	19	-21.23%	19	**-19.66%**	14	**27.75%**	13	10.42%	16	-0.30%
Jun	78	**19.55%**	73	-20.68%	32	-25.73%	54	**27.65%**	44	**38.52%**	50	-4.57%
Jul	16	**19.67%**	17	30.54%	15	**-13.56%**	15	**32.56%**	11	**26.72%**	12	1.71%
Aug	17	**33.63%**	20	-19.53%	19	-19.15%	14	**14.69%**	12	**30.10%**	13	11.98%
Sep	71	**22.97%**	73	**-19.29%**	52	**7.07%**	42	**76.97%**	56	-6.67%	62	**26.46%**
Oct	20	3.40%	19	**-35.51%**	14	20.62%	14	**28.80%**	14	-6.14%	8	**22.20%**
Nov	10	***2.12%***	9	-51.06%	8	18.57%	8	**13.24%**	7	-3.43%	6	**14.86%**
Dec	702	**9.35%**	915	**-39.91%**	844	**62.79%**	473	**21.22%**	502	-7.26%	486	**14.19%**

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 4.16: Comparison of annual payoffs between the strategies of investing in the portfolios of under-priced stocks (grey-box model) and investing in the contemporary S & P index.

	Year	2006	2007	2008	2009	2010	2011	2012
Jan	Portfolio 1		1	0.7371	0.4093	0.7534	**0.9985**	**1.1063**
	S&P Index		1	0.9458	0.5225	0.7851	0.9434	0.9708
Feb	Portfolio 2	1	**1.1983**	0.8300	**0.6253**	0.8466	0.7967	0.7909
	S&P Index	1	1.0973	1.0215	0.6162	0.9031	1.0239	1.0877
Mar	Portfolio 3	1	1.1181	1.0018	0.6242	**1.1232**	**1.2958**	**1.2573**
	S&P Index	1	1.1311	1.0572	0.6660	0.9054	1.0404	1.0666
Apr	Portfolio 4	1	**1.2074**	1.0953	**0.7726**	**1.2122**	**1.2409**	**1.2862**
	S&P Index	1	1.2051	1.1026	0.7237	0.8577	1.0591	1.0317
May	Portfolio 5	1	**1.2295**	0.9685	**0.7781**	**0.9940**	**1.0976**	**1.0944**
	S&P Index	1	1.1836	1.0077	0.7238	0.8115	1.0397	1.0724
Jun	Portfolio 6	1	**1.1955**	0.9483	0.7043	**0.8990**	**1.2454**	**1.1884**
	S&P Index	1	1.1399	0.9927	0.7735	0.8629	1.0122	1.0804
Jul	Portfolio 7	1	**1.1967**	**1.5621**	**1.3503**	**1.7899**	**2.2682**	**2.3070**
	S&P Index	1	1.1305	0.9839	0.7828	0.8048	0.9349	1.0788
Aug	Portfolio 8	1	**1.3363**	**1.0753**	**0.8693**	**0.9970**	**1.2972**	**1.4526**
	S&P Index	1	1.1429	0.8731	0.7913	0.8543	0.8470	1.0785
Sep	Portfolio 9	1	**1.2297**	**0.9924**	**1.0626**	**1.8805**	**1.7550**	**2.2194**
	S&P Index	1	1.1244	0.7030	0.7520	0.8587	0.9095	1.0248
Oct	Portfolio 10	1	1.0340	**0.6669**	**0.8044**	**1.0361**	**0.9725**	**1.1884**
	S&P Index	1	1.0575	0.6399	0.7822	0.8429	0.8903	1.0111
Nov	Portfolio 11	1	1.0212	0.4998	0.5926	0.6710	0.6480	0.7443
	S&P Index	1	1.0353	0.6369	0.7862	0.8867	0.8867	1.0056
Dec	Portfolio 12	1	**1.0935**	**0.6571**	**1.0697**	**1.2967**	**1.2025**	**1.3732**
	S&P Index	1	0.9585	0.5742	0.7467	0.8942	0.9125	1.0416

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 4.17: The number of stocks selected by the second investment strategy and the average one-year return of the long-position portfolio (U.S. data).

	2006		2007		2008		2009		2010		2011	
	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return	No. of targets	Average one-year return
Jan			6	**26.01%**	0	0	4	**132.05%**	4	15.48%	3	**47.55%**
Feb	2	3.01%	3	-15.48%	0	0	4	**160.80%**	3	-0.65%	2	-8.43%
Mar	18	**19.85%**	9	**7.48%**	4	**-28.84%**	9	**125.81%**	4	1.17%	2	**114.97%**
Apr	5	**21.42%**	2	**-3.81%**	3	**-13.44%**	2	**87.80%**	4	16.34%	2	**39.00%**
May	4	**58.98%**	3	-35.40%	0	0	4	**33.71%**	2	15.75%	2	-29.04%
Jun	3	**24.80%**	9	-23.84%	1	-27.45%	15	**33.09%**	3	**24.31%**	0	0
Jul	0	0	1	**48.02%**	0	0	5	**36.96%**	0	0	0	0
Aug	0	0	1	-33.59%	0	0	2	-19.64%	2	**22.82%**	0	0
Sep	8	2.11%	7	**-23.67%**	9	**24.78%**	13	12.29%	2	-31.98%	2	7.22%
Oct	1	-43.91%	3	**-20.23%**	5	16.63%	2	-5.89%	2	-0.61%	1	**60.68%**
Nov	1	**11.24%**	2	-53.39%	0	0	1	12.69%	0	0	0	0
Dec	40	**3.17%**	91	**-38.42%**	38	**100.21%**	75	**20.78%**	79	-9.56%	51	**30.42%**

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 4.18: Comparison of annual payoffs between the strategies of investing in the portfolios constructed by the second strategy and investing in the contemporary S & P index.

	Year	2006	2007	2008	2009	2010	2011	2012
Jan	Portfolio 1		1	**1.2601**	**1.2601**	**2.9241**	**3.3769**	**4.9825**
	S&P Index		1	0.9458	0.5225	0.7851	0.9434	0.9708
Feb	Portfolio 2	1	1.0301	0.8707	**0.8707**	**2.2707**	**2.2560**	**2.0659**
	S&P Index	1	1.0973	1.0215	0.6162	0.9031	1.0239	1.0877
Mar	Portfolio 3	1	**1.1985**	**1.2881**	**0.9167**	**2.0699**	**2.0941**	**4.5017**
	S&P Index	1	1.1311	1.0572	0.6660	0.9054	1.0404	1.0666
Apr	Portfolio 4	1	**1.2142**	**1.1680**	**1.0109**	**1.8986**	**2.2088**	**3.0703**
	S&P Index	1	1.2051	1.1026	0.7237	0.8577	1.0591	1.0317
May	Portfolio 5	1	**1.5898**	**1.0270**	**1.0270**	**1.3732**	**1.5895**	**1.1279**
	S&P Index	1	1.1836	1.0077	0.7238	0.8115	1.0397	1.0724
Jun	Portfolio 6	1	**1.2480**	0.9505	0.6896	**0.9177**	**1.1409**	**1.1409**
	S&P Index	1	1.1399	0.9927	0.7735	0.8629	1.0122	1.0804
Jul	Portfolio 7	1	1.0000	**1.4802**	**1.4802**	**2.0273**	**2.0273**	**2.0273**
	S&P Index	1	1.1305	0.9839	0.7828	0.8048	0.9349	1.0788
Aug	Portfolio 8	1	1.0000	0.6641	0.6641	0.5336	0.6554	0.6554
	S&P Index	1	1.1429	0.8731	0.7913	0.8543	0.8470	1.0785
Sep	Portfolio 9	1	1.0211	**0.7794**	**0.9726**	**1.0921**	0.7429	0.7965
	S&P Index	1	1.1244	0.7030	0.7520	0.8587	0.9095	1.0248
Oct	Portfolio 10	1	0.5609	0.4475	0.5219	0.4911	0.4882	0.7844
	S&P Index	1	1.0575	0.6399	0.7822	0.8429	0.8903	1.0111
Nov	Portfolio 11	1	**1.1124**	0.5185	0.5185	0.5843	0.5843	0.5843
	S&P Index	1	1.0353	0.6369	0.7862	0.8867	0.8867	1.0056
Dec	Portfolio 12	1	**1.0317**	**0.6353**	**1.2720**	**1.5362**	**1.3894**	**1.8121**
	S&P Index	1	0.9585	0.5742	0.7467	0.8942	0.9125	1.0416

Note: the bold numbers indicate the times when the constructed portfolios outperform the contemporary market index.

Table 5.1: Descriptive statistics of the variables in the cash flow model (U.K. listed firms).

	CF	*ΔINV*	*ΔAP*	*ΔAR*	*DEP*	*AMORT*	*OTHER*
Mean	0.0291	0.0046	0.0090	0.0112	0.0277	0.0134	0.0430
Standard Error	0.0015	0.0004	0.0008	0.0007	0.0003	0.0003	0.0016
Median	0.0482	0.0000	0.0046	0.0050	0.0188	0.0019	0.0483
Mode	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Standard Deviation	0.1614	0.0384	0.0816	0.0766	0.0315	0.0377	0.1697
Sample Variance	0.0261	0.0015	0.0067	0.0059	0.0010	0.0014	0.0288
Kurtosis	5.0458	69.1757	273.3624	71.2522	28.9258	182.0565	118.2252
Skewness	-1.5650	-1.2422	-6.1853	-0.2835	3.5440	9.9871	-6.7490
Range	1.3417	1.4034	4.3218	3.3780	0.6727	1.4640	6.2672
Minimum	-0.8745	-0.8564	-3.2472	-1.9674	0.0000	-0.3753	-3.7004
Maximum	0.4672	0.5470	1.0746	1.4106	0.6727	1.0887	2.5669

Note: The variables are defined as follows:

CF : Net operating cash flow;

ΔINV : Changes in inventory;

ΔAP : Changes in account payable;

ΔAR : Changes in account receivable;

DEP: Depreciation;

AMORT: Amortisation;

OTHER: Other accruals, calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – ($CF + \Delta AR + \Delta INV – \Delta AP – DEP – AMORT$).

All variables are deflated by average total assets of each firm.

Table 5.2: The cash flow prediction model parameters estimated with panel data methods (U.K. listed firms, in-sample).

	Pooled Regression (1990-2006)	Demean (1990-2006)	A-B (1990-2006)
β_1	0.830^{***}	0.346^{***}	0.343^{***}
	(33.40)	(8.44)	(6.29)
β_2	0.261^{*}	0.151^{*}	0.104
	(2.13)	(2.03)	(1.25)
β_3	-0.601^{***}	-0.403^{***}	-0.349^{***}
	(-6.73)	(-4.62)	(-3.57)
β_4	0.478^{***}	0.355^{***}	0.300^{***}
	(6.13)	(5.00)	(4.13)
β_5	0.278^{**}	0.575^{***}	0.468^{**}
	(3.09)	(4.84)	(2.94)
β_6	0.0255	-0.0334	-0.138
	(0.43)	(-0.73)	(-1.88)
β_7	-0.0166	-0.0220	-0.0303
	(-0.88)	(-0.88)	(-0.93)
β_0	0.000697		
	(0.32)		
N	4480	4480	3675

Note: *t* statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The parameters are defined as in the model (See Table 5.1 for variable definitions): $CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 5.3: The mean, standard deviation and the variance-covariance matrix of the parameters distributed among individual firms estimated by random parameter model (U.K. listed firms, in-sample).

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Mean	0.0050	0.7453	0.5013	-0.7594	0.6137	0.2139	-0.2386	0.1000
St. Dev.	0.0100	0.3715	0.6740	0.9840	0.7096	0.7053	1.0099	0.3731
Variance-Covariance matrix								
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_0	0.0001	-0.0036	-0.0012	0.0052	-0.0021	0.0006	-0.0061	-0.0002
β_1	-0.0036	0.1380	0.1414	-0.1323	0.0816	-0.0257	0.1476	-0.0433
β_2	-0.0012	0.1414	0.4543	-0.3099	0.1670	0.2024	0.0418	-0.1956
β_3	0.0052	-0.1323	-0.3099	0.9682	-0.4026	-0.0596	-0.1650	0.0447
β_4	-0.0021	0.0816	0.1670	-0.4026	0.5035	0.0788	0.1004	-0.0920
β_5	0.0006	-0.0257	0.2024	-0.0596	0.0788	0.4975	0.2315	-0.2075
β_6	-0.0061	0.1476	0.0418	-0.1650	0.1004	0.2315	1.0199	-0.1007
β_7	-0.0002	-0.0433	-0.1956	0.0447	-0.0920	-0.2075	-0.1007	0.1392

Note: The parameters are defined as in the model (See Table 5.1 for variable definitions): $CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 5.4: The in-sample fitness of various cash flow prediction models (U.K. listed firms).

Models	1	2	3	4	5	6	7	8
MSE	0.0069	0.0076	0.0054	0.0037	0.0038	0.0084	0.0052	0.0052
Av. Rank	4.1445	4.5199	4.7268	4.8838	4.9644	4.7280	4.3383	3.6943

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: Bayesian model with prior from random parameter model;
Model 7: grey-box model using sales growth rates;
Model 8: grey-box model using firm age.

Table 5.5: The out-of-sample (one period) performance of various cash flow prediction models (U.K. listed firms).

Models	1	2	3	4	5	6	7	8
MSE	0.0136	0.0117	0.0101	0.0125	0.0128	0.0130	0.0101	0.0102
Av. Rank	3.9406	4.4810	4.8730	4.8873	4.8256	4.8643	4.3574	3.7708
Models' ranks in the two measures of performance								
Rank(MSE)	8	4	1	5	6	7	2	3
Rank(Av. Rank)	2	4	7	8	5	6	3	1

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: Bayesian model with prior from random parameter model;
Model 7: grey-box model using sales growth rates;
Model 8: grey-box model using firm age.

Table 5.6: The out-of-sample (multi-period) performance of various models (U.K. listed firms).

	Models	1	2	3	4	5	6	7
2008	MSE	0.0154	0.0162	**0.0132**	0.0163	0.0164	0.0142	**0.0134**
	Av. Rank	**3.5361**	3.8578	4.2013	4.2407	4.1028	4.1969	**3.8643**
2009	MSE	0.0202	0.0218	**0.0165**	0.0208	0.0206	0.0211	**0.0164**
	Av. Rank	**3.8206**	4.0175	4.0175	4.1772	4.0153	3.9891	**3.9628**
2010	MSE	0.0211	0.0242	**0.0180**	0.0227	0.0224	0.0788	**0.0179**
	Av. Rank	**3.8568**	3.9802	3.9846	**3.9515**	3.9537	4.1806	4.0925
2011	MSE	0.0220	0.0257	**0.0190**	0.0256	0.0253	0.1949	**0.0187**
	Av. Rank	**3.8838**	3.9795	**3.9499**	3.9704	3.9522	4.1891	4.0752
2012	MSE	0.0250	0.0273	**0.0214**	0.0269	0.0267	0.5887	**0.0212**
	Av. Rank	**3.8333**	4.1765	4.0515	3.8995	**3.7941**	4.2181	4.0270
2013	MSE	0.0205	0.0218	**0.0195**	0.0203	0.0204	1.9715	**0.0199**
	Av. Rank	**3.8385**	4.0708	**3.8669**	3.8725	3.9972	4.2125	4.1416
2014	MSE	0.0221	0.0252	**0.0205**	0.0245	0.0243	6.8115	**0.0209**
	Av. Rank	**3.7726**	4.1204	4.0368	3.9097	**3.8562**	4.2140	4.0903

Note: Bold numbers indicate two models that perform best in that criterion.
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: Bayesian model with prior from random parameter model;
Model 7: grey-box model using firm age.

Table 5.7: The comparison of expected and actual price movement directions using random walk, DKW, BCN and grey-box models respectively to predict cash flows (U.K. listed firms).

Random Walk	Actually occur	Not occur	Correct/Wrong
Predicted to rise	1256	170	7.39
Predicted to decline	2106	507	4.15
Sum	3362	677	
DKW	Actually occur	Not occur	Correct/Wrong
Predicted to rise	1320	182	7.25
Predicted to decline	2033	504	4.03
Sum	3353	686	
BCN	Actually occur	Not occur	Correct/Wrong
Predicted to rise	1097	180	6.09
Predicted to decline	2206	585	3.77
Sum	3303	765	
Grey-box	Actually occur	Not occur	Correct/Wrong
Predicted to rise	1075	141	7.62
Predicted to decline	2229	594	3.75
Sum	3304	735	

Note: Correct/Wrong is the ratio calculated as the number of 'Actually occur' divided by the number of 'Not occur'.

Table 6.1: Descriptive statistics of the variables in the cash flow model (U.K. unlisted data).

	CF	*ΔINV*	*ΔAP*	*ΔAR*	*DEP*	*AMORT*	*OTHER*
Mean	0.0933	0.0079	0.0089	0.0115	0.0325	0.0155	0.0503
Standard Error	0.0010	0.0004	0.0005	0.0005	0.0002	0.0002	0.0016
Median	0.0874	0.0011	0.0035	0.0035	0.0235	0.0060	0.0442
Mode	0.0828	0.0000	0.0000	0.0000	0.0004	0.0002	0.0000
Standard Deviation	0.1620	0.0597	0.0746	0.0842	0.0339	0.0349	0.2536
Sample Variance	0.0262	0.0036	0.0056	0.0071	0.0011	0.0012	0.0643
Kurtosis	283.9409	30.1920	18.5729	21.3246	39.0656	242.7998	9553.6167
Skewness	-5.8488	0.3410	0.2336	0.5204	4.1274	8.8046	79.1298
Range	11.3565	2.0053	1.8870	2.6756	0.9862	2.1823	36.1177
Minimum	-7.9937	-0.9551	-0.8669	-1.3868	-0.2246	-0.7951	-4.9184
Maximum	3.3627	1.0502	1.0201	1.2888	0.7616	1.3872	31.1994

Note: The variables are defined as follows:

CF : Net operating cash flow;

ΔINV : Changes in inventory;

ΔAP : Changes in account payable;

ΔAR : Changes in account receivable;

DEP: Depreciation;

AMORT: Amortisation;

OTHER: Other accruals, calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – ($CF + \Delta AR + \Delta INV – \Delta AP – DEP–AMORT$).

All variables are deflated by average total assets of each firm.

Table 6.2: The cash flow prediction model parameters estimated with panel data methods (U.K. unlisted firms, in-sample).

	Pooled Regression (2005-2012)	Demean (2005-2012)	A-B (2005-2012)
β_1	0.530***	-0.0219	-0.00546
	(11.18)	(-0.21)	(-0.04)
β_2	0.245***	-0.0446	-0.0879
	(3.51)	(-0.47)	(-0.73)
β_3	-0.370***	-0.221**	-0.143
	(-4.97)	(-2.59)	(-1.54)
β_4	0.398***	0.378***	0.243*
	(5.79)	(3.92)	(2.11)
β_5	0.197	0.110	0.107
	(1.91)	(1.23)	(1.70)
β_6	-0.177	0.0867	-0.0751
	(-1.88)	(0.30)	(-0.12)
β_7	0.0839	0.184***	0.126*
	(1.13)	(4.17)	(2.03)
β_0	0.0343***		
	(6.83)		
N	3701	3701	435

Note: *t* statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The parameters are defined as in the model (See Table 6.1 for variable definitions):
$CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 6.3: The mean, standard deviation and the variance-covariance matrix of the parameters distributed among individual firms estimated by random parameter model (U.K. unlisted firms, in-sample).

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Mean	0.0152	0.6724	0.4965	-0.7039	0.7028	-0.2289	-0.3339	0.5422
St. Dev.	0.0619	0.5518	0.7606	0.8869	0.5263	1.1847	1.0724	0.7178
Variance-Covariance matrix								
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_0	0.0038	-0.0204	-0.0184	0.0039	-0.0001	-0.0053	0.0054	-0.0126
β_1	-0.0204	0.3045	0.0745	-0.1552	0.0873	-0.3587	-0.3882	0.1538
β_2	-0.0184	0.0745	0.5786	-0.5715	0.2733	-0.2455	-0.2245	0.2956
β_3	0.0039	-0.1552	-0.5715	0.7866	-0.3977	0.6995	0.6257	-0.4308
β_4	-0.0001	0.0873	0.2733	-0.3977	0.2770	-0.5060	-0.3773	0.2907
β_5	-0.0053	-0.3587	-0.2455	0.6995	-0.5060	1.4035	1.2010	-0.6710
β_6	0.0054	-0.3882	-0.2245	0.6257	-0.3773	1.2010	1.1501	-0.6191
β_7	-0.0126	0.1538	0.2956	-0.4308	0.2907	-0.6710	-0.6191	0.5152

The parameters are defined as in the model (See Table 6.1 for variable definitions): $CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 6.4: The in-sample fitness of various cash flow prediction models (U.K. unlisted data).

Models	1	2	3	4	5	6	7	8
MSE	0.0223	0.0229	0.0158	0.0010	0.0010	0.0240	0.0149	0.0170
Av. Rank	4.1355	4.8656	4.2849	4.7442	5.1752	5.0206	4.2366	3.5375

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: Bayesian model with prior from random parameter model;
Model 7: grey-box model using sales growth rates;
Model 8: grey-box model using firm age.

Table 6.5: The out-of-sample performance of various cash flow prediction models (U.K. unlisted data).

Models		1	2	3	4	5	6	7	8
Panel A: One-period-ahead prediction comparison									
2013	MSE	0.0227	0.0229	**0.0168**	0.0213	0.0213	0.0169	**0.0158**	0.0179
	Av. Rank	4.4984	4.8069	4.7931	4.6457	4.5448	4.5064	**4.3967**	**3.8080**
2014	MSE	0.0300	0.0298	0.0244	0.0323	0.0313	**0.0242**	**0.0238**	0.0260
	Av. Rank	4.2320	4.5681	4.9204	4.9148	4.6789	4.6734	**4.2235**	**3.7890**
Panel B: Two-period-ahead prediction comparison									
2014	MSE	0.0315	0.0326	0.0280	0.0297	0.0295	0.0279	-	0.0293
	Av. Rank	4.0086	4.1328	4.1915	3.7301	4.1508	3.9722	-	3.8141
	Rank(MSE)	6	7	2	5	4	1	-	3
	Rank(Av. Rank)	4	5	7	1	6	3	-	2

Note: Bold numbers indicate two models that perform best in that criterion.
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: Bayesian model with prior from random parameter model;
Model 7: grey-box model using sales growth rates;
Model 8: grey-box model using firm age.

Table 7.1: Descriptive statistics for variables in the cash flow prediction models (China data).

	CF	*CF_Crec*	*CF_Cpaid*	*CF_NCrec*	*CF_NCpaid*	*ΔINV*	*ΔAP*	*ΔAR*	*DEP*	*AMORT*	*OTHER*
Mean	0.0492	0.7926	0.5956	0.0566	0.2046	0.0262	0.0237	0.0364	0.0243	0.0027	-0.0147
Standard Error	0.0007	0.0059	0.0054	0.0015	0.0018	0.0006	0.0007	0.0007	0.0002	0.0000	0.0004
Median	0.0385	0.5394	0.3579	0.0229	0.1405	0.0095	0.0102	0.0170	0.0166	0.0011	-0.0080
Mode	0.0237	0.4493	0.1356	0.0000	0.1234	0.0000	0.0249	-0.0010	0.0024	0.0000	-0.0092
Standard Deviation	0.1258	1.0315	0.9521	0.2544	0.3176	0.0978	0.1265	0.1198	0.0263	0.0064	0.0690
Sample Variance	0.0158	1.0639	0.9066	0.0647	0.1009	0.0096	0.0160	0.0143	0.0007	0.0000	0.0048
Kurtosis	30.7337	140.4783	192.5368	6972.2132	3055.7259	162.2159	119.4941	77.8484	26.4867	344.5215	277.3589
Skewness	-0.3651	8.1920	9.8995	71.5607	39.6663	8.0217	4.1721	4.2391	3.5050	12.7674	-8.6087
Range	4.3855	32.7547	32.3130	28.8558	28.6675	4.5714	6.5914	5.0912	0.5322	0.3394	4.1013
Minimum	-2.4818	0.0000	-0.1280	-0.4938	-0.0351	-1.0838	-2.2548	-1.5780	-0.0250	-0.0267	-2.6304
Maximum	1.9037	32.7547	32.1851	28.3621	28.6325	3.4876	4.3365	3.5132	0.5072	0.3127	1.4709

Note: The variables are defined as follows:

CF : Net operating cash flow;

CF_Crec : Cash received from the sales of goods and rendering of services;

CF_Cpaid : Cash paid for goods and services;

CF_NCrec : Other cash receipt calculated as sub-total cash inflows from operating activities minus CF_Crec;

CF_NCpaid : Other cash paid out calculated as sub-total cash inflows from operating activities minus CF_Cpaid ;

ΔINV : Changes in inventory;

ΔAP : Changes in account payable;

ΔAR : Changes in account receivable;

DEP: Depreciation;

AMORT: Amortisation;

OTHER: Other accruals, calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – (*CF + ΔAR + ΔINV – ΔAP – DEP – AMORT*).

All variables are deflated by average total assets of each firm.

Table 7.2: The cash flow prediction model parameters estimated with panel data methods (China listed firms, in-sample).

	Pooled Regression (1998-2006)	Demean (1998-2006)	A-B (1998-2006)
β_1	0.239***	-0.00912	0.00923
	(5.53)	(-0.27)	(0.17)
β_2	0.124**	0.0924	-0.0453
	(2.83)	(1.83)	(-0.76)
β_3	-0.138**	-0.0794*	0.0135
	(-3.23)	(-2.08)	(0.24)
β_4	0.0838*	0.0470	-0.00398
	(2.10)	(1.25)	(-0.08)
β_5	1.115***	0.910***	0.352
	(11.66)	(5.11)	(0.69)
β_6	-1.244*	-0.468	-0.208
	(-2.07)	(-0.66)	(-0.21)
β_7	0.0105	-0.0127	-0.0171
	(0.62)	(-0.70)	(-0.67)
β_0	0.00999***		
	(6.81)		
N	8246	8246	6613

Note: t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
The parameters are defined as in the model (see Table 7.1 for variable definitions):
$CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 7.3: The mean, standard deviation and the correlation matrix of the parameters distributed among individual firms estimated by random parameter model (China listed firms, in-sample).

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Mean	0.0082	0.3293	0.2018	-0.2985	0.2642	0.8573	-1.0588	-0.0323
St. Dev.	0.0116	0.6323	0.7107	0.7046	0.6760	1.5198	5.3704	0.1895
Correlation matrix								
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
β_0	1.0000	-0.6226	-0.5082	0.2693	-0.3180	-0.1282	0.1162	0.7947
β_1	-0.6226	1.0000	0.6502	-0.7404	0.8091	-0.6033	0.1487	-0.4149
β_2	-0.5082	0.6502	1.0000	-0.7292	0.6501	-0.4639	0.1355	-0.1976
β_3	0.2693	-0.7404	-0.7292	1.0000	-0.8018	0.5321	0.0651	0.1651
β_4	-0.3180	0.8091	0.6501	-0.8018	1.0000	-0.7052	0.1217	-0.2222
β_5	-0.1282	-0.6033	-0.4639	0.5321	-0.7052	1.0000	-0.6053	-0.2039
β_6	0.1162	0.1487	0.1355	0.0651	0.1217	-0.6053	1.0000	0.2790
β_7	0.7947	-0.4149	-0.1976	0.1651	-0.2222	-0.2039	0.2790	1.0000

Note: The parameters are defined as in the model (see Table 7.1 for variable definitions):

$CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t + \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1}$

Table 7.4: The disaggregated cash flow model parameters estimated with panel data methods (China listed firms, in-sample).

	Pooled Regression (1998-2006)	Demean (1998-2006)	A-B (1998-2006)
β_1	0.250***	0.00211	0.0491
	(6.00)	(0.06)	(0.84)
β_2	-0.250***	0.000186	-0.0576
	(-5.88)	(0.00)	(-0.95)
β_3	0.0784	-0.105*	-0.0935
	(1.54)	(-2.28)	(-1.35)
β_4	-0.165***	0.0595	-0.0228
	(-3.73)	(1.71)	(-0.43)
β_5	0.110**	0.0822	-0.0172
	(2.66)	(1.68)	(-0.29)
β_6	-0.122**	-0.0686	-0.0144
	(-3.06)	(-1.91)	(-0.25)
β_7	0.0491	0.0199	-0.00105
	(1.32)	(0.55)	(-0.02)
β_8	0.963***	0.763***	0.649
	(9.27)	(3.47)	(0.96)
β_9	-1.039*	-0.486	-0.0375
	(-2.08)	(-0.74)	(-0.04)
β_{10}	0.0140	-0.0132	-0.0207
	(0.77)	(-0.64)	(-0.74)
β_0	0.00954***		
	(5.29)		
N	8211	8211	6576

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The parameters are defined as in the model (see Table 7.1 for variable definitions):

$$CF_{t+1} = \beta_0 + \beta_1 CF\_Crec_t + \beta_2 CF\_Cpaid_t + \beta_3 CF\_NCrec_t + \beta_4 CF\_NCpaid_t$$
$$+\beta_5 \Delta INV_t + \beta_6 \Delta AP_t + \beta_7 \Delta AR_t + \beta_8 DEP_t + \beta_9 AMORT_t + \beta_{10} OTHER_t + \varepsilon_{t+1}$$

Table 7.5: The in-sample fitness of various cash flow prediction models (China data).

Models	1	2	3	4	5	6	7	8	9	10	11
MSE	0.0086	0.0080	0.0053	0.0041	0.0042	0.0052	0.0041	0.0042	0.0065	0.0052	0.0055
Av. Rank	5.2605	6.0713	6.2420	6.3027	6.0512	6.1217	6.3869	6.4778	6.3512	5.9825	4.7520

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: CH model (pooled regression),
Model 7: CH model (demean);
Model 8: CH model (Arellano-Bond estimators);
Model 9: Bayesian model with prior from random parameter model;
Model 10: grey-box model using sales growth rates;
Model 11: grey-box model using firm age.

Table 7.6: The out-of-sample (one period) performance of various cash flow prediction models (China data).

Models	1	2	3	4	5	6	7	8	9	10	11
MSE	0.0283	0.0260	0.0231	0.0270	0.0274	0.0224	0.0261	0.0261	0.0244	0.0224	0.0233
Av. Rank	4.9053	5.9974	6.3955	6.5928	6.4578	6.3325	6.3822	6.4176	6.3396	5.8698	4.3095
Rank(MSE)	11	6	3	9	10	1	7	8	5	2	4
Rank(Av. Rank)	2	4	8	11	10	5	7	9	6	3	1

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: CH model (pooled regression),
Model 7: CH model (demean);
Model 8: CH model (Arellano-Bond estimators);
Model 9: Bayesian model with prior from random parameter model;
Model 10: grey-box model using sales growth rates;
Model 11: grey-box model using firm age.

Table 7.7: The out-of-sample (multi-period) performance of various cash flow prediction models (China data).

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0306	0.0300	0.0294	0.0297	0.0301	0.0291	0.0296	0.0300	0.0677	0.2552
Av. Rank	5.1558	5.2863	5.4660	5.6318	5.3781	5.4754	5.6413	5.3865	5.3357	6.2431
Rank(MSE)	8	6	2	4	7	1	3	5	9	10
Rank(Av. Rank)	1	2	6	8	4	7	9	5	3	10

Note: Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: CH model (pooled regression),
Model 7: CH model (demean);
Model 8: CH model (Arellano-Bond estimators);
Model 9: Bayesian model with prior from random parameter model;
Model 10: grey-box model using firm age.

Table 7.8: The out-of-sample (multi-period) performance of various cash flow prediction models (China data) with adjustment made to the grey-box model.

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0306	0.0300	0.0294	0.0297	0.0301	0.0291	0.0296	0.0300	0.0677	0.0293
Av. Rank	4.7684	5.4712	5.7305	6.0113	5.5127	5.5742	5.9149	5.6597	5.5799	4.7772
Rank(MSE)	9	7	3	5	8	1	4	6	10	2
Rank(Av. Rank)	1	3	8	10	4	5	9	7	6	2

Note: 7146 observations; year 2007 excluded;
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (Arellano-Bond estimators);
Model 6: CH model (pooled regression),
Model 7: CH model (demean);
Model 8: CH model (Arellano-Bond estimators);
Model 9: Bayesian model with prior from random parameter model;
Model 10: Combinative model of grey-box model using firm age and Model 3.

Table 7.9: The comparison of two strategies in predicting directions of stock price movement and the actual price movement directions for China market.

No strategy	Sum	Actually occur	Not occur	Correct/Wrong
Predict all shares to rise	12466	9952	2514	3.96
Predict all shares to decline	12466	9927	2539	3.91
DSCF model	Sum	Actually occur	Not occur	Correct/Wrong
Predicted to rise	3206	2633	573	4.60
Predicted to decline	9260	7435	1825	4.07

Note: Correct/Wrong is the ratio calculated as the number of 'Actually occur' divided by the number of 'Not occur'.

Table 7.10: The accumulated (up to 12 months) stock return prediction regression results (China market).

k	1	2	3	4	5	6	7	8	9	10	11	12
β_k	0.000665	0.00169**	0.00372***	0.00516***	0.00490***	0.00533***	0.0114***	0.0122***	0.0158***	0.0195***	0.0202***	0.0207***
	(1.54)	(2.94)	(5.22)	(6.20)	(4.83)	(4.63)	(8.44)	(8.81)	(10.22)	(11.15)	(10.53)	(11.44)
α_k	0.0164***	-0.00560***	-0.00231	0.0148***	0.0393***	0.0574***	0.0841***	0.0879***	0.108***	0.157***	0.218***	0.160***
	(12.92)	(-3.32)	(-1.10)	(6.02)	(13.15)	(16.88)	(21.09)	(21.43)	(23.59)	(30.34)	(38.30)	(28.72)
N	12624	12503	12431	12454	12449	12454	12379	12355	12341	12332	12317	11199

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The regression equation is:

$$r_{t+k} = \alpha_k + \beta_k E_t(r) + \varepsilon_{t+k} \ ,$$

where *k* takes values from 1 to 12, indicating the number of months for the return accumulation period.

C. Figures

Figure 3.1: Examples of Monte Carlo method in simulating cash flow distributions.

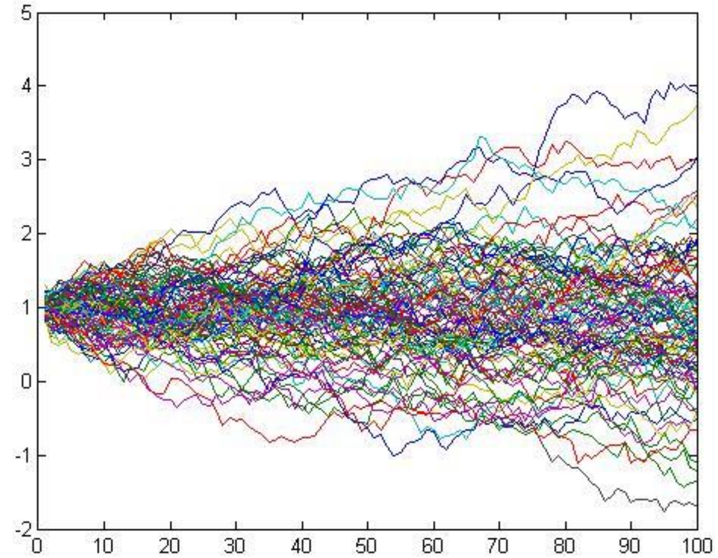


Fig. 3.1a: no quit rule is applied .

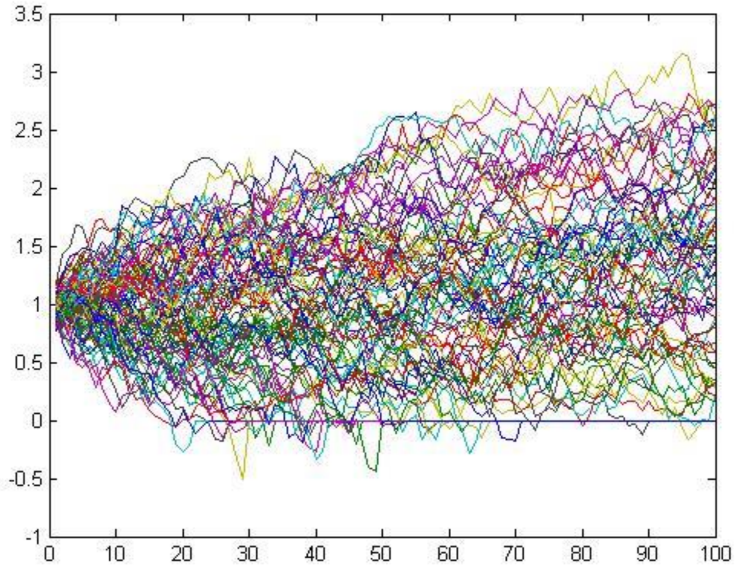


Fig. 3.1b: a quit rule is added to the simulation procedure.

Figure 4.1: The distribution of the number of firms with different sample lengths.

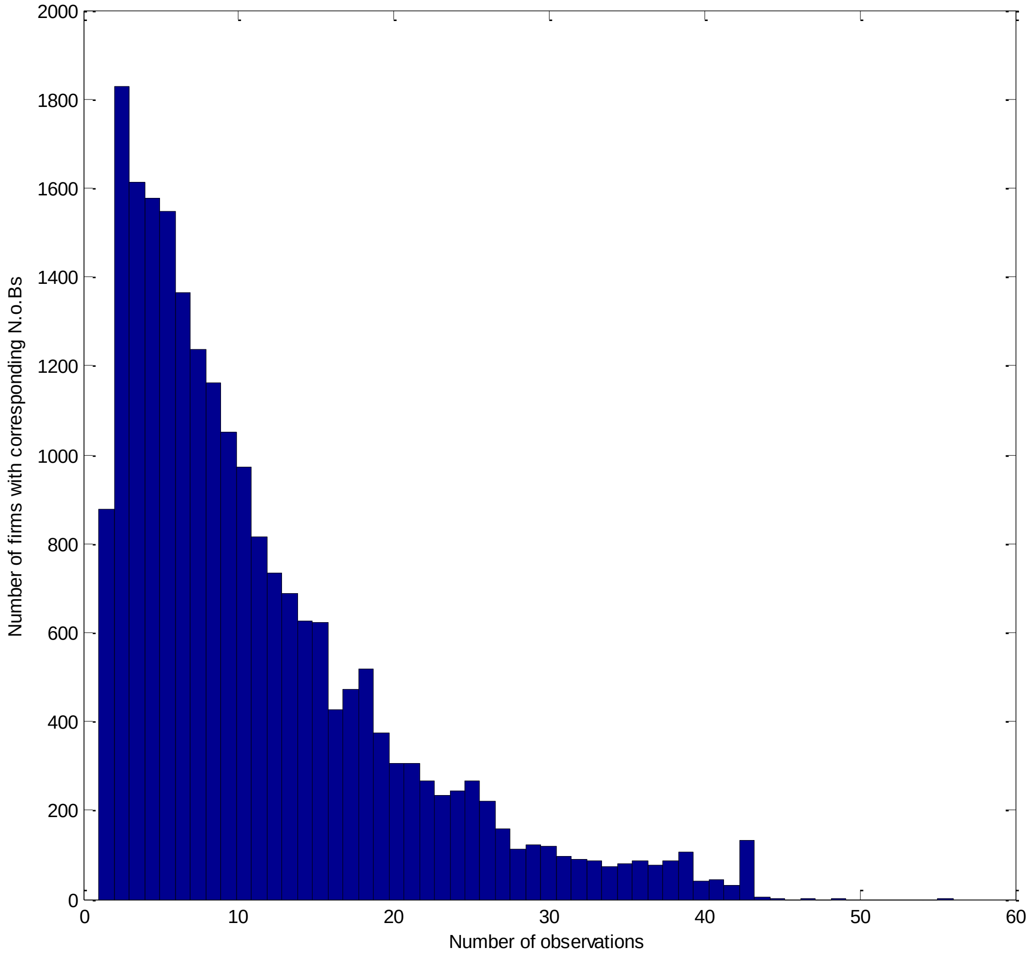


Figure 4.2: Mean, median and the 2.5 to 97.5 percentiles of the normalised cash flow series of all firms.

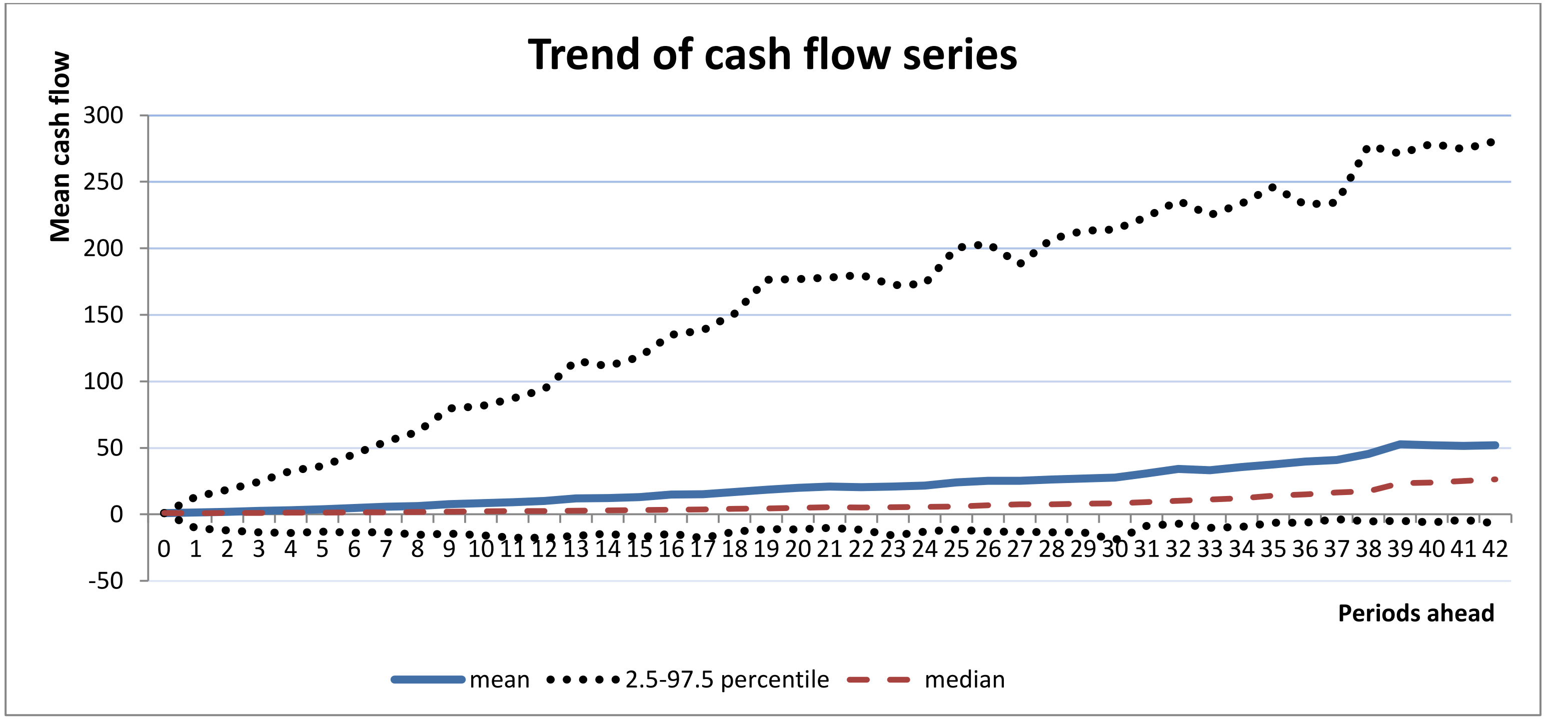


Figure 4.3: Mean, median and the 2.5 to 97.5 percentiles of the simulated cash flow series.

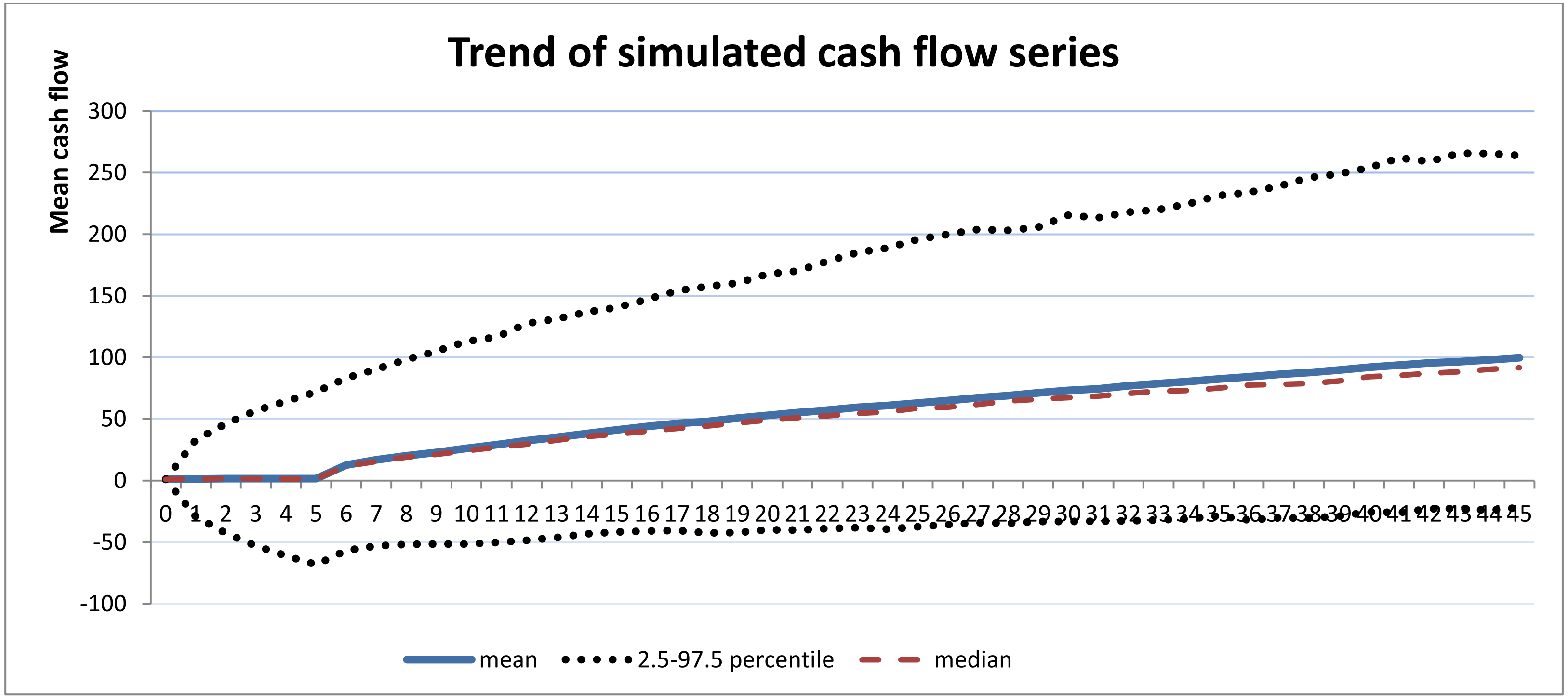


Figure 4.4: Survival rates of U.S. sample firms.

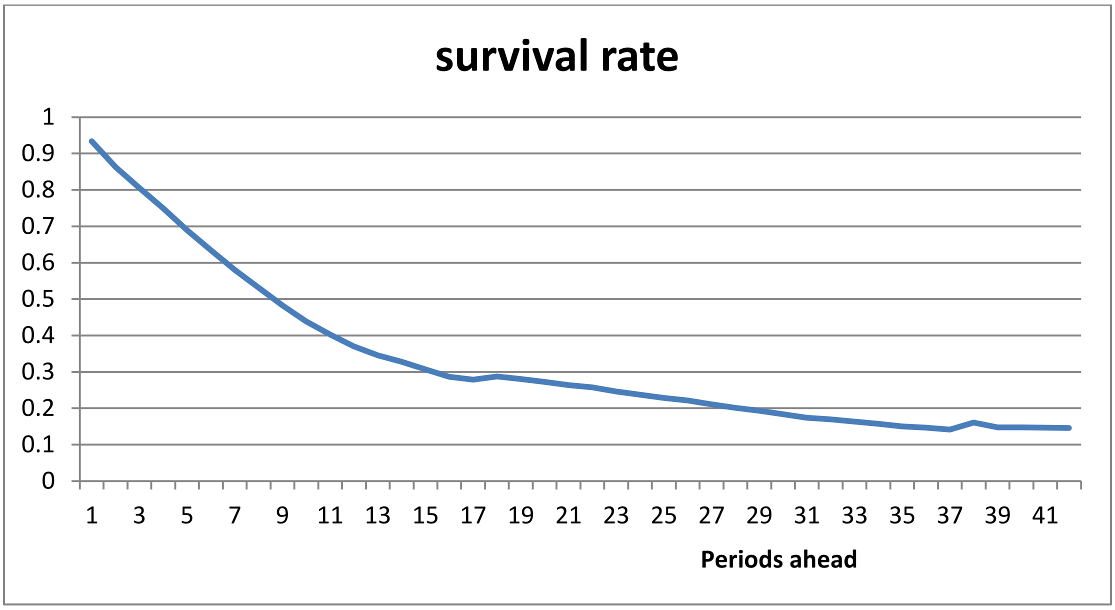


Figure 4.5: The mean cash flow series adjusted for survivorship bias.

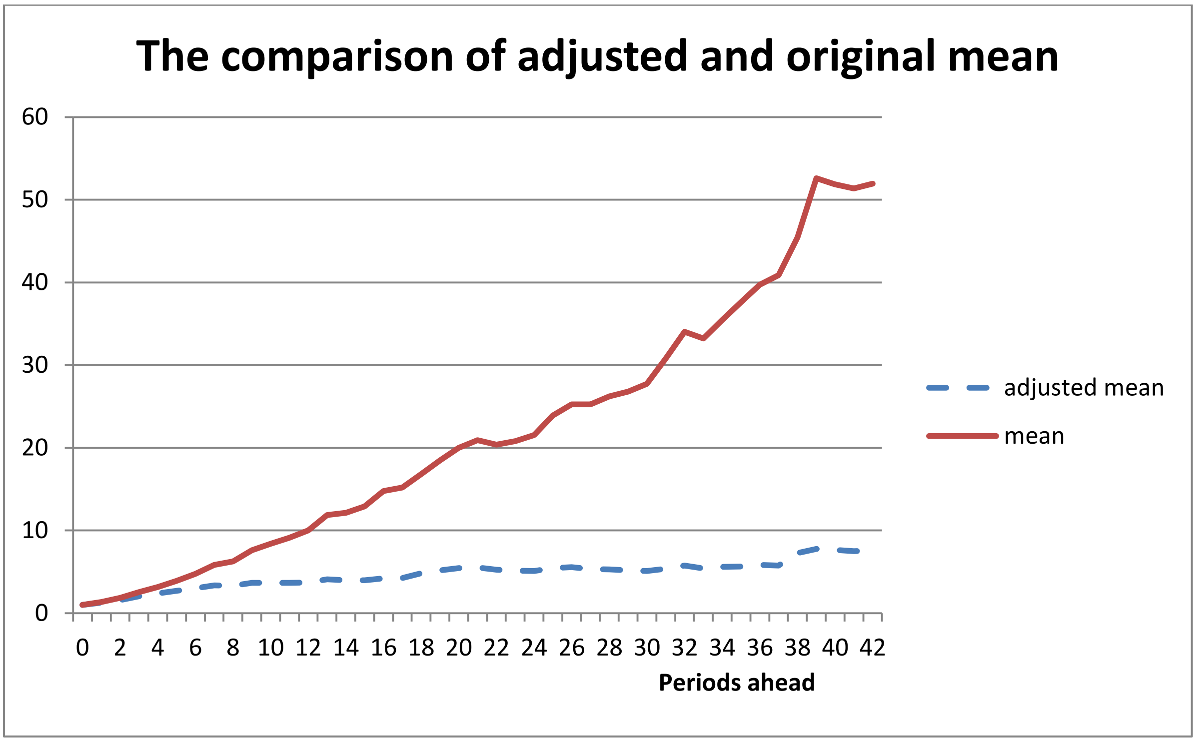


Figure 4.6: The distributions of each parameter, adjusted R squared and residual variance estimated individually for available firms.

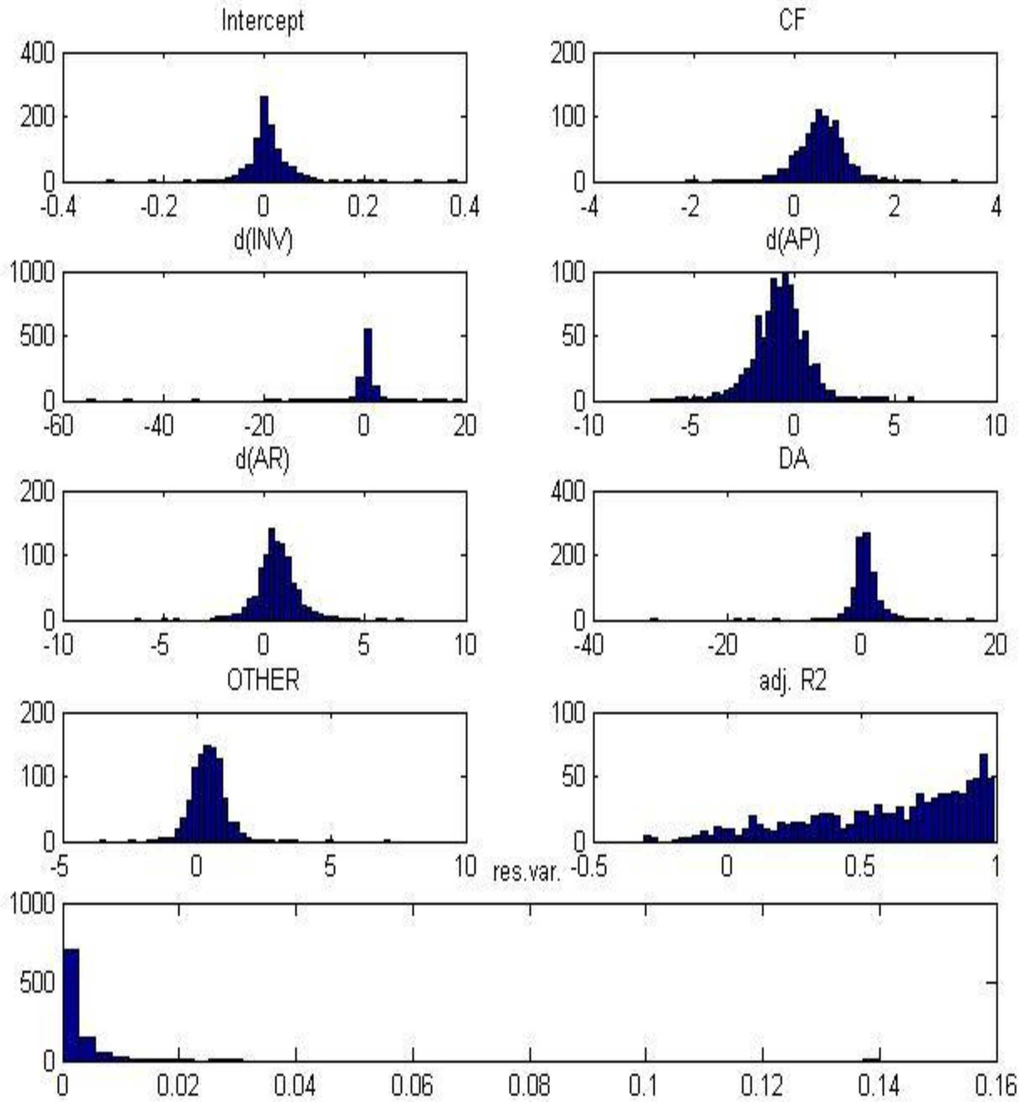


Figure 4.7: Bayesian predictions comparison for AAR Corp.

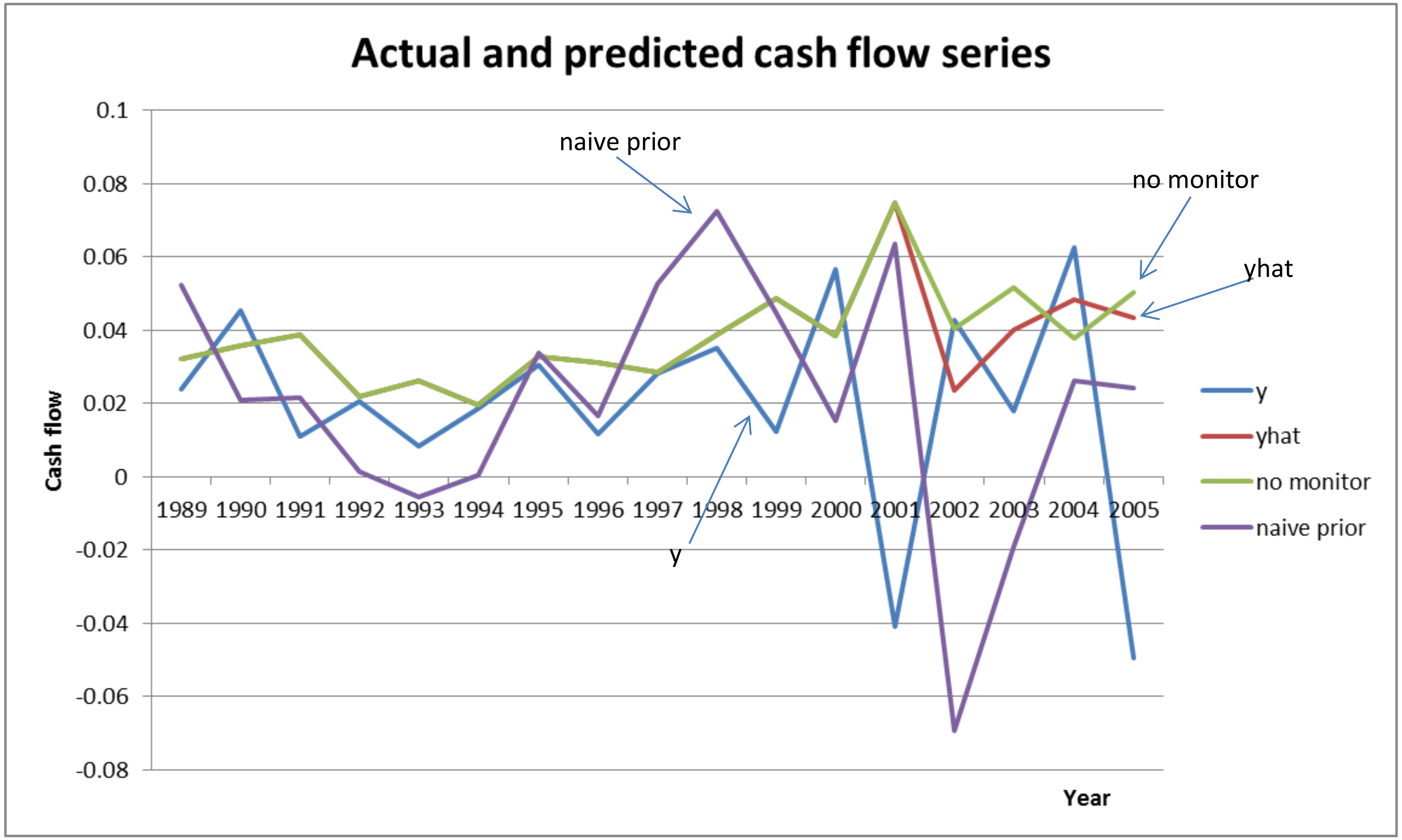


Figure 4.8: the association of parameter values and sales growth rates by grey-box model (U.S. listed firms).

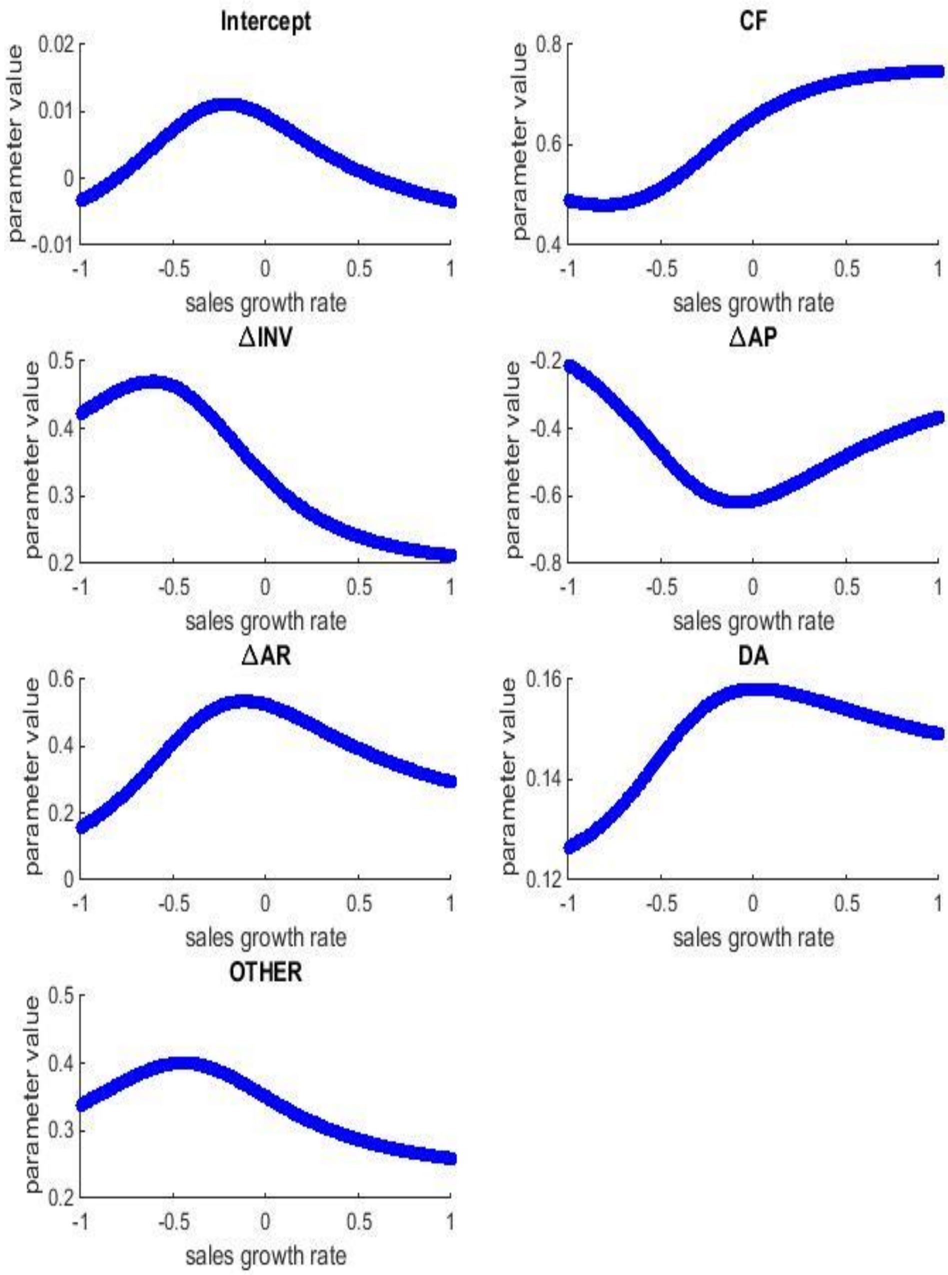


Figure 4.9: The relationship between mean sales growth rate and firm age.

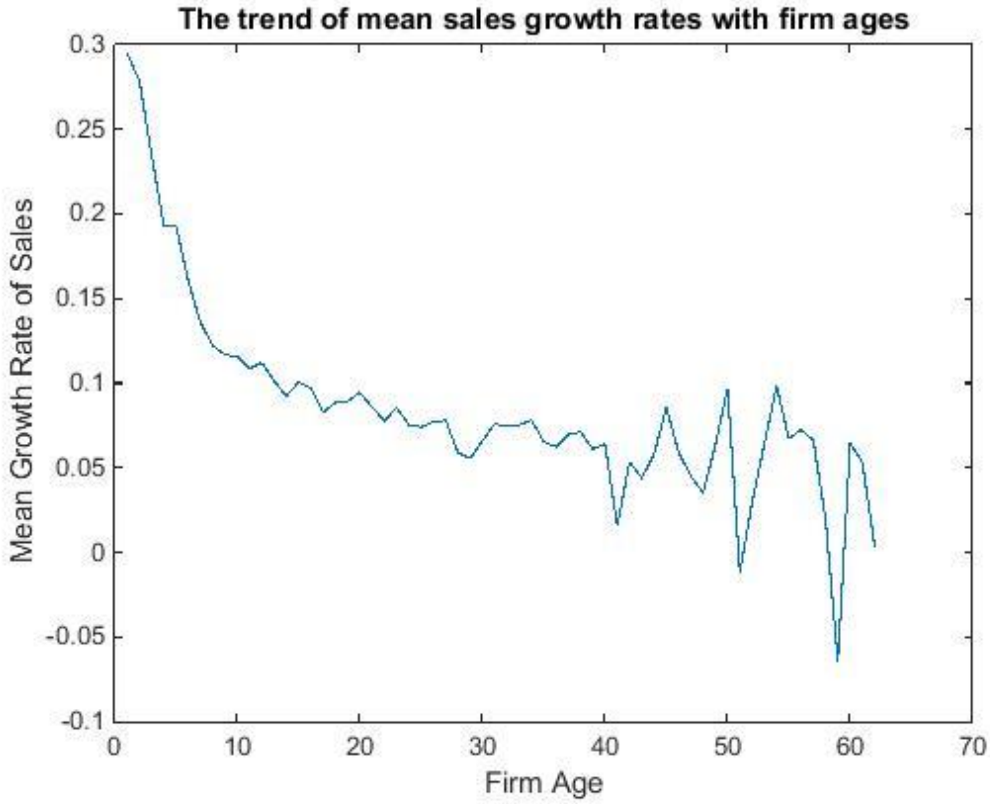


Figure 4.10: The evolution of parameters values with increasing firm age by grey-box model (U.S. listed firms).

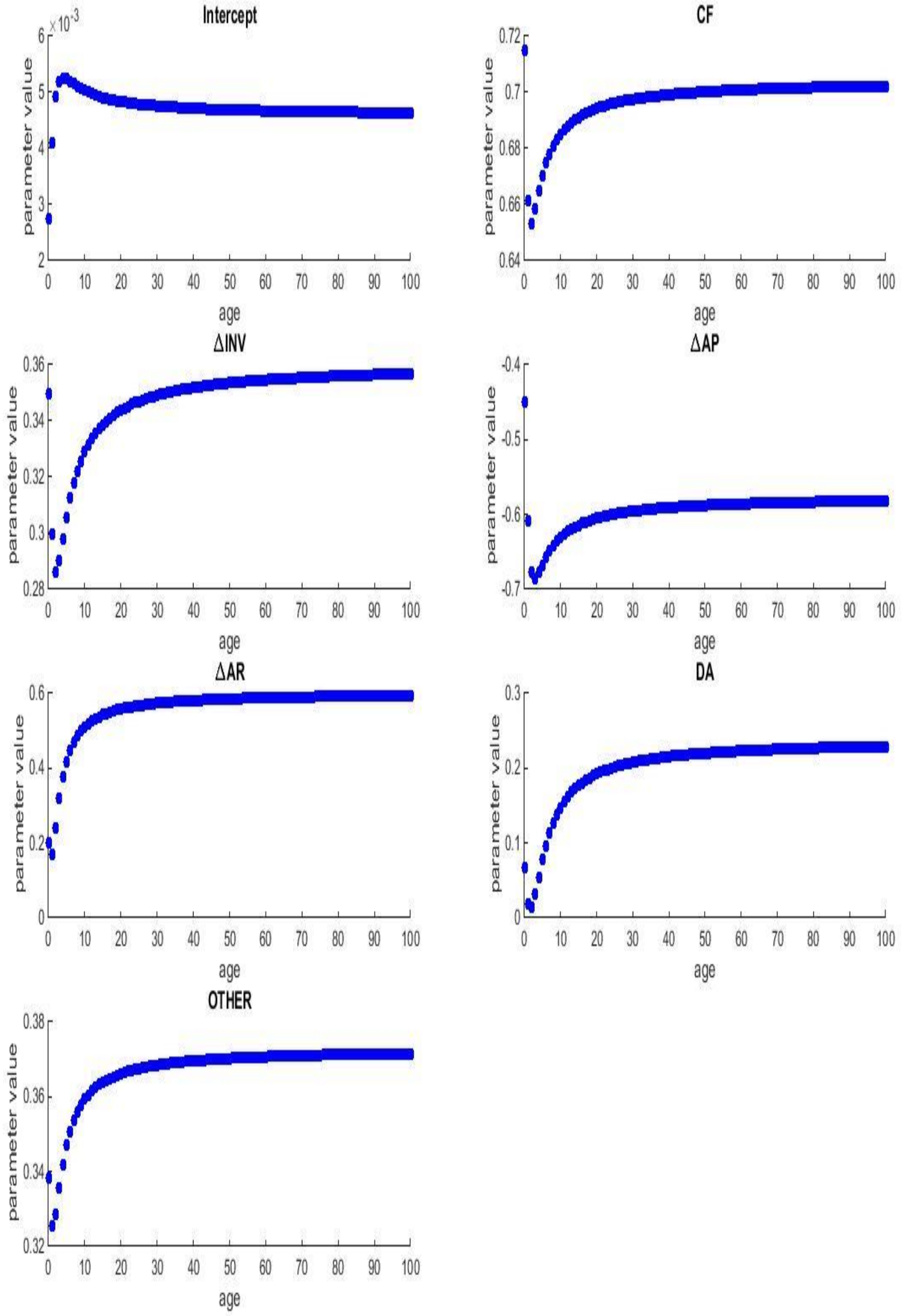
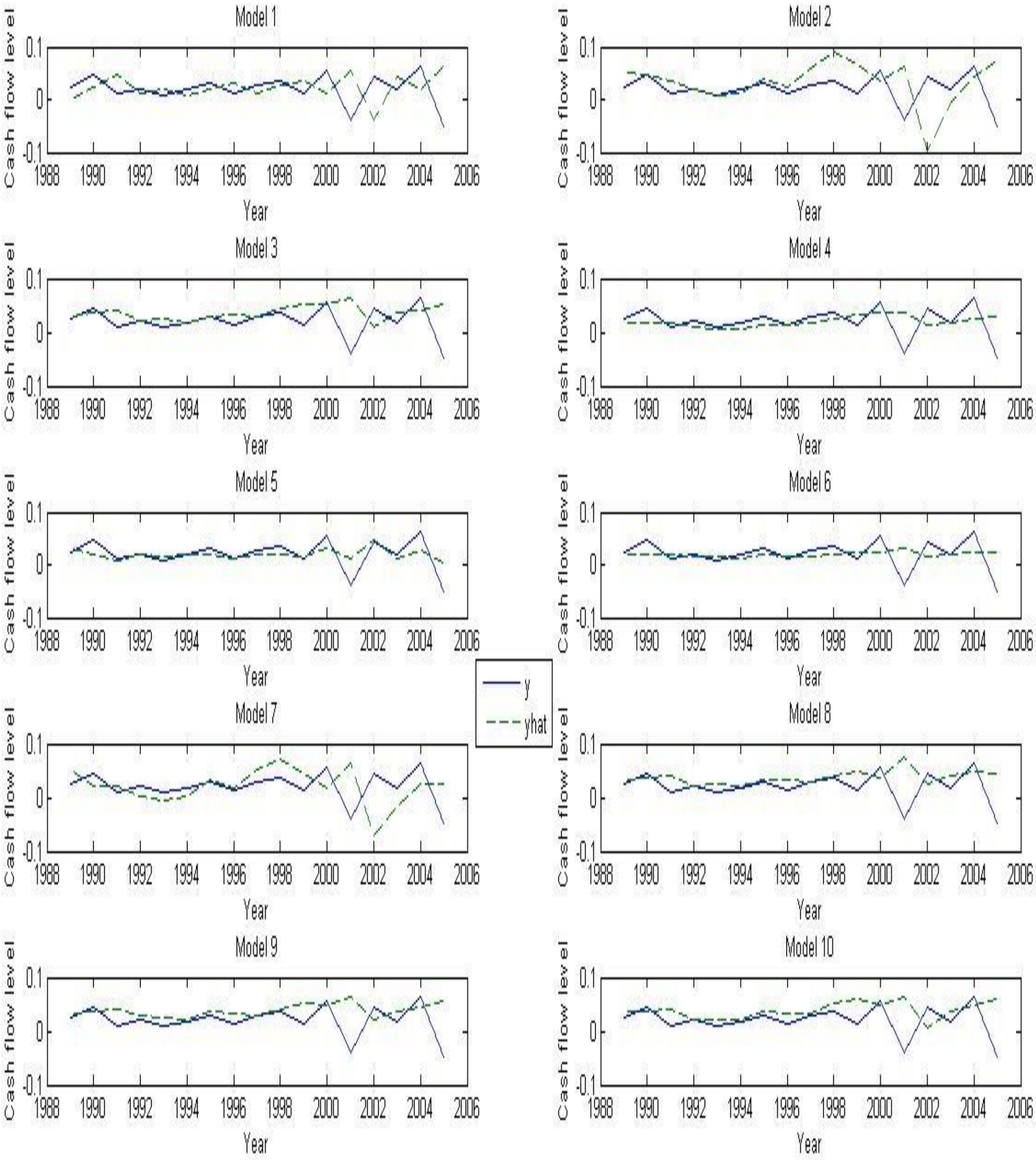
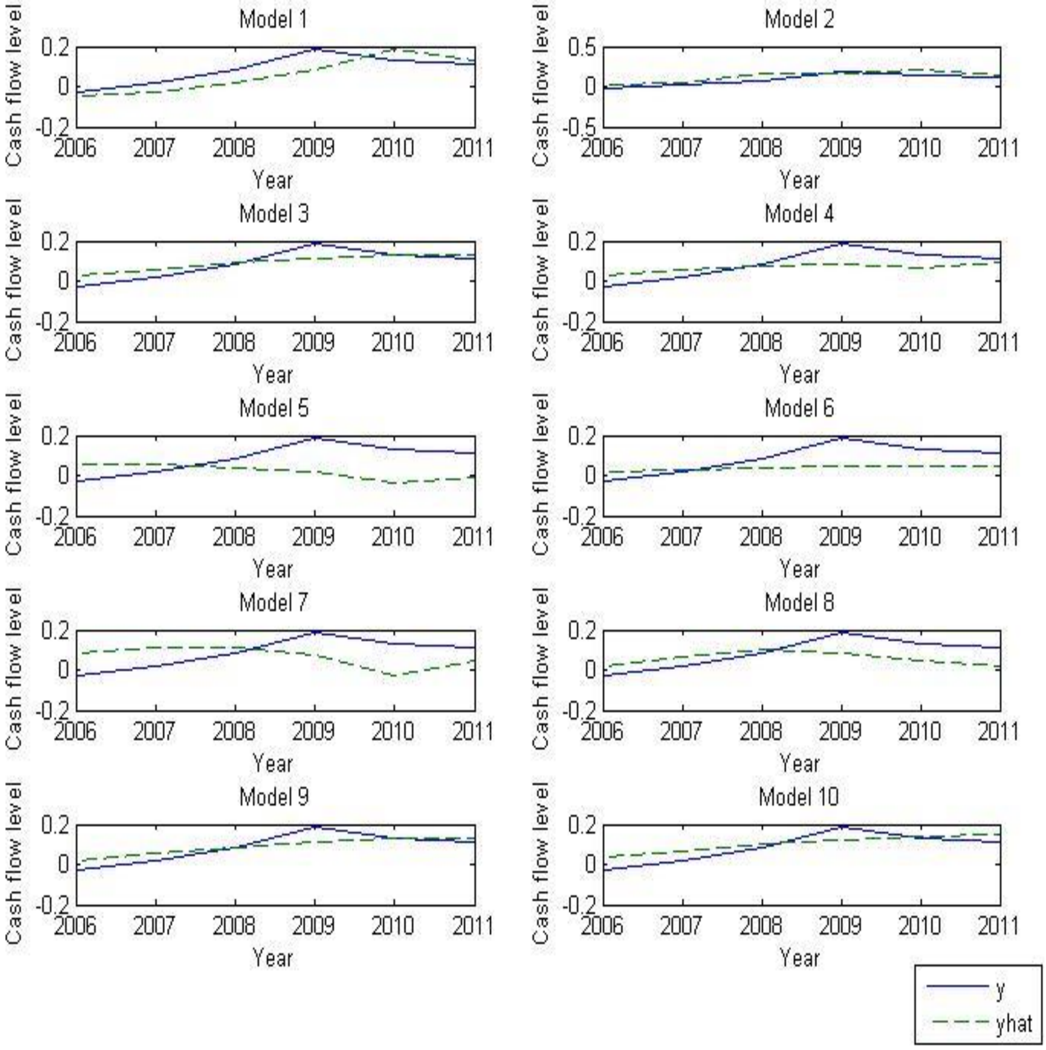


Figure 4.11: The in-sample fitted cash flow series using different model for AAR Corp.



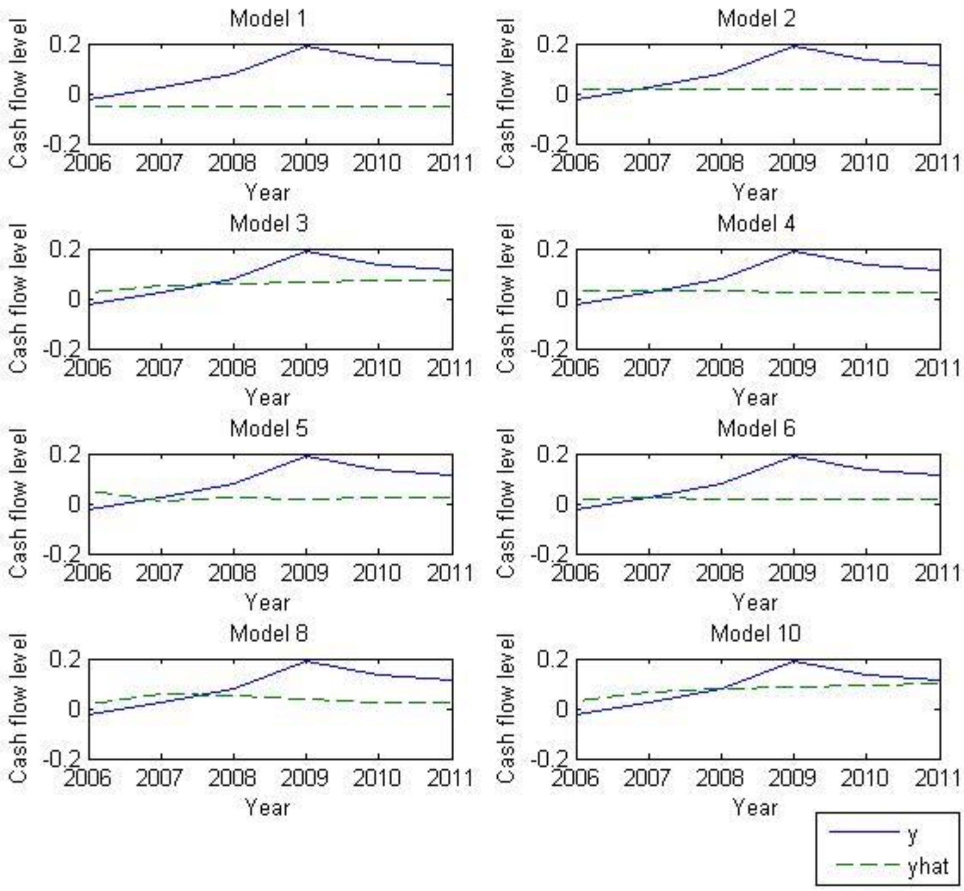
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 7: Bayesian model with naive prior ;
Model 8: Bayesian model with prior from random parameter model;
Model 9: grey-box model using sales growth rates;
Model 10: grey-box model using firm age.

Figure 4.12: The out-of-sample one-period-ahead predicted cash flow series using different model for AAR Corp.



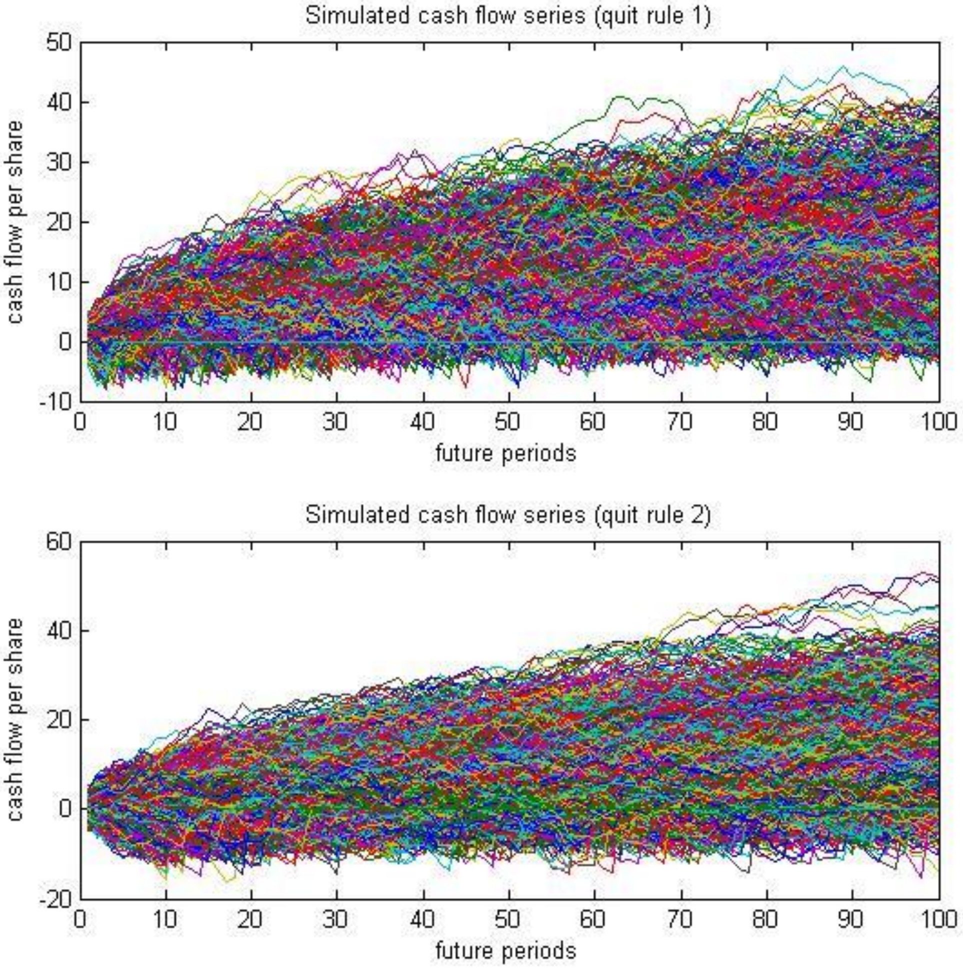
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 7: Bayesian model with naive prior ;
Model 8: Bayesian model with prior from random parameter model;
Model 9: grey-box model using sales growth rates;
Model 10: grey-box model using firm age.

Figure 4.13: The out-of-sample multi-period-ahead predicted cash flow series using different model for AAR Corp.



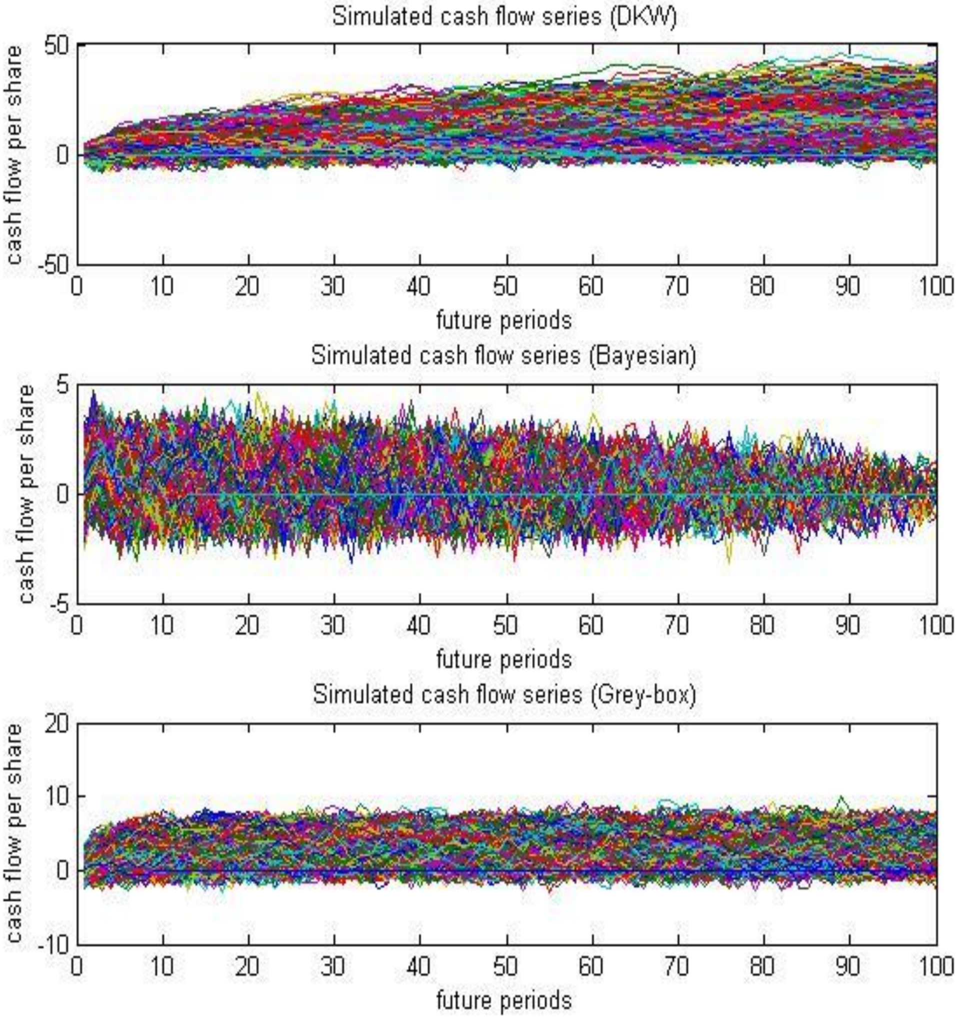
Model 1: random walk model;
Model 2: theoretical DKW model;
Model 3: BCN model (pooled regression),
Model 4: BCN model (demean);
Model 5: BCN model (difference);
Model 6: BCN model (Arellano-Bond estimators);
Model 8: Bayesian model with prior from random parameter model;
Model 10: grey-box model using firm age as additional input.

Figure 4.14: The simulated cash flow series using DKW model for AAR Corp after year 2005 with different quit rules.



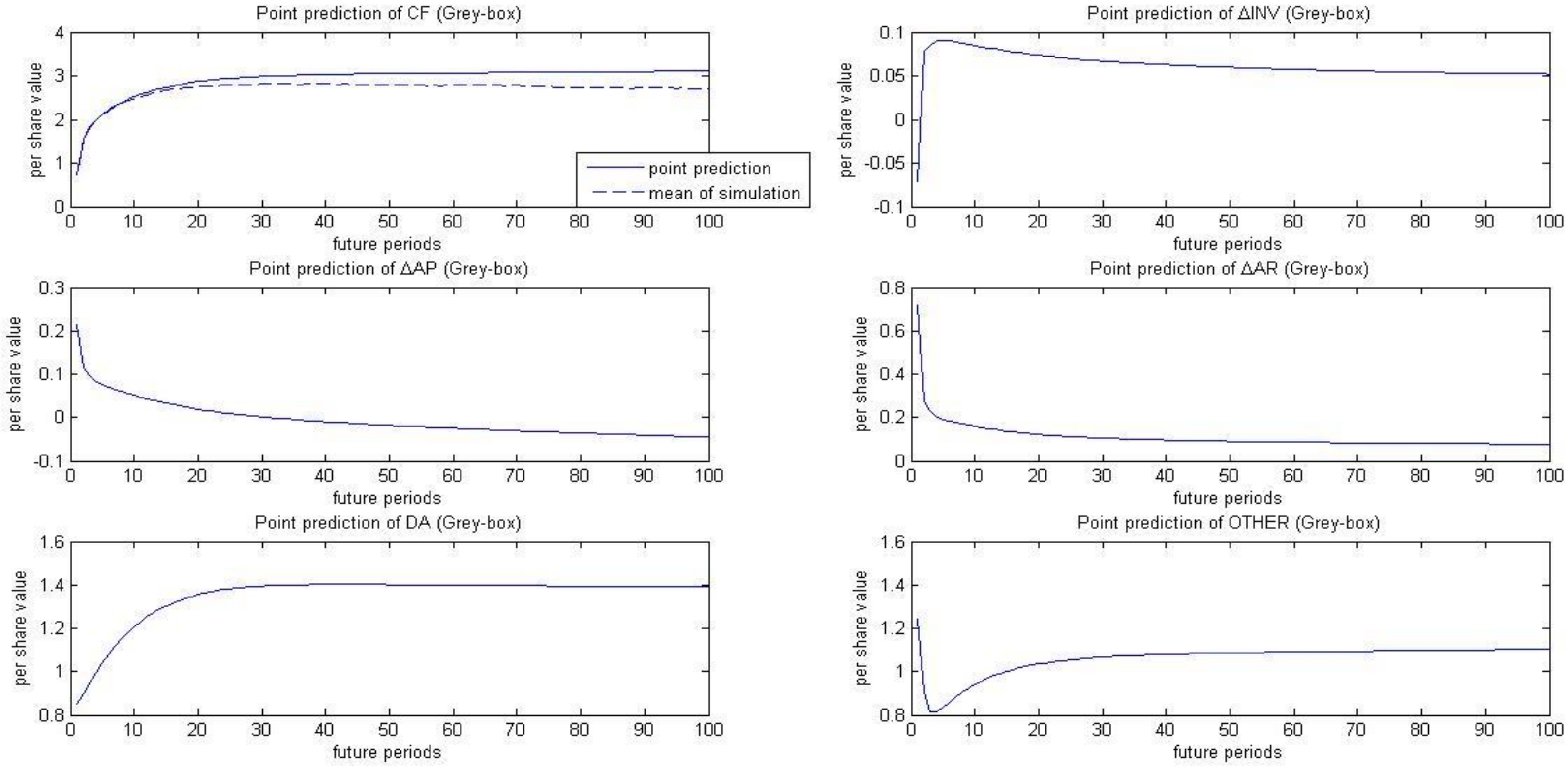
Note: quit rule 1: 3 negative cash flows one after another means firm breaks up;
quit rule 2: 10 negative cash flows one after another means firm breaks up.

Figure 4.15: The simulated cash flow series using DKW model, Bayesian model and grey-box model for AAR Corp after year 2005.



Note: all three models apply the quit rule of 3 negative observations.

Figure 4.16: Point predictions of relevant variables using grey-box model for AAR Corp after year 2005.



Note: the dashed line in the first chart is the mean of the simulated series in figure 15, lower chart.

Figure 4.17: CRRA utility functions when the coefficients take different values.

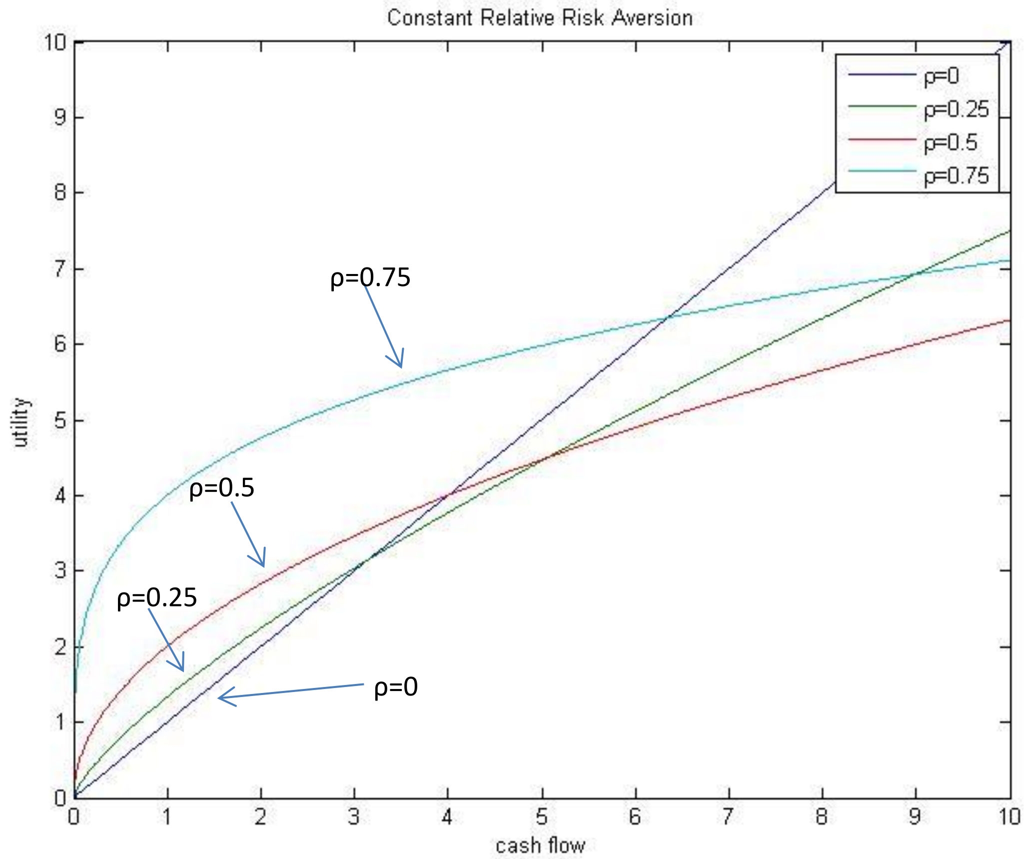


Figure 4.18: The certainty equivalent cash flow of the simulated data by DKW model for AAR Corp.

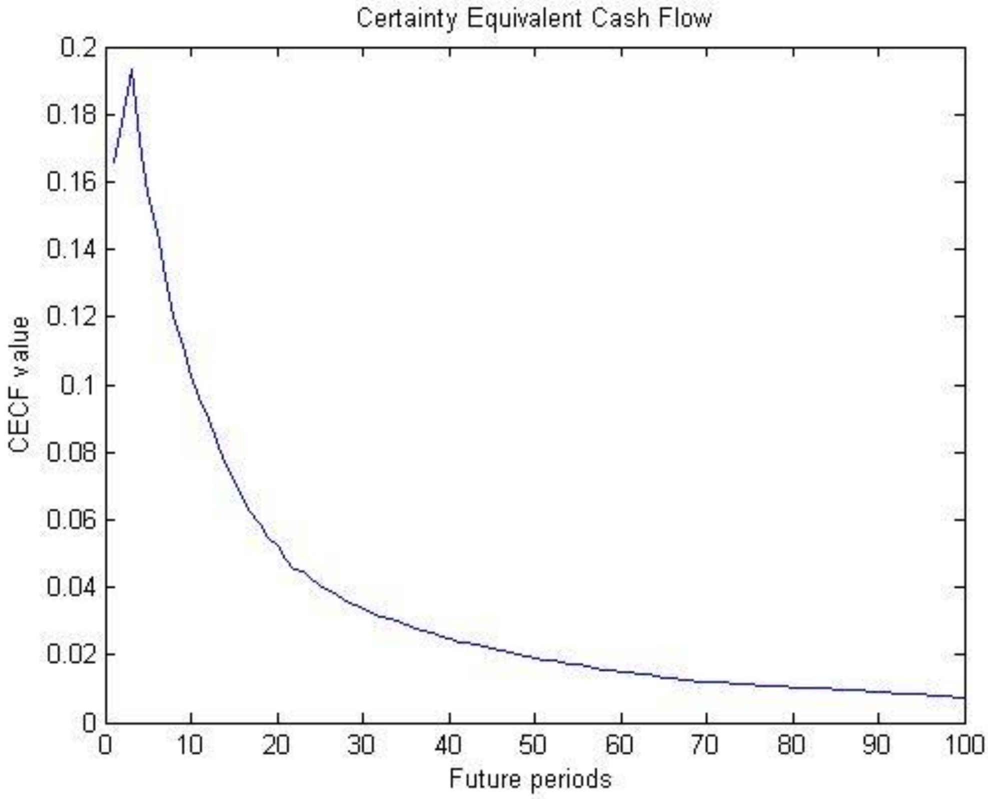


Figure 4.19: The certainty equivalent cash flow of the simulated data by Bayesian model and grey-box model for AAR Corp.

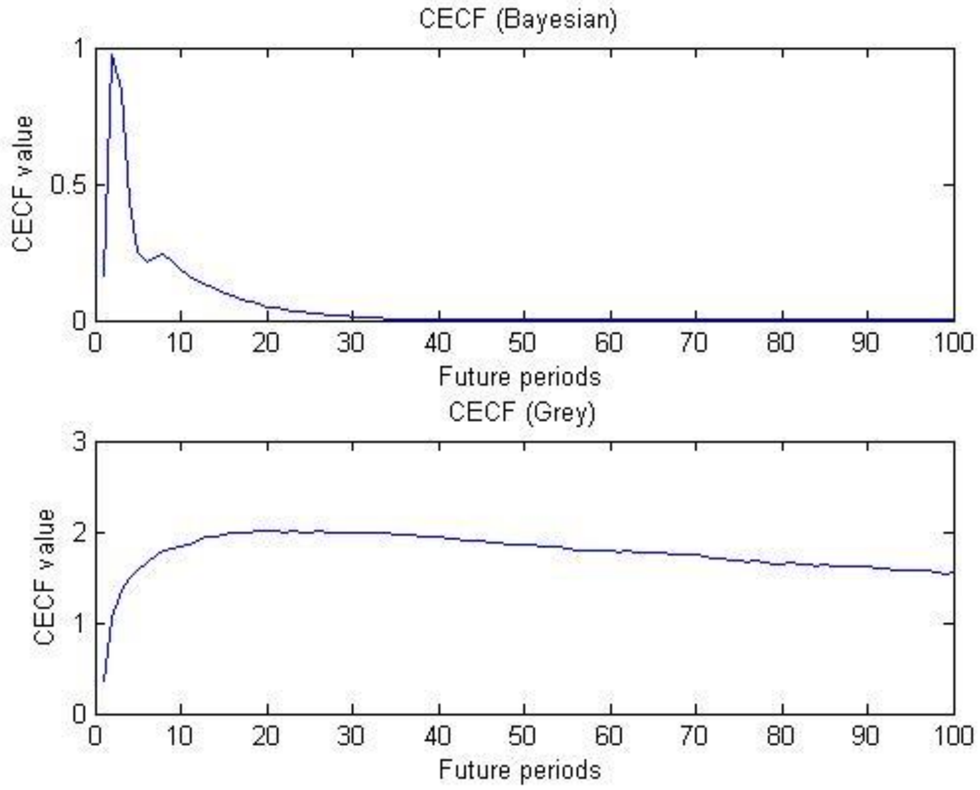


Figure 4.20: The interpolated and extrapolated yield curve in May, 2006.

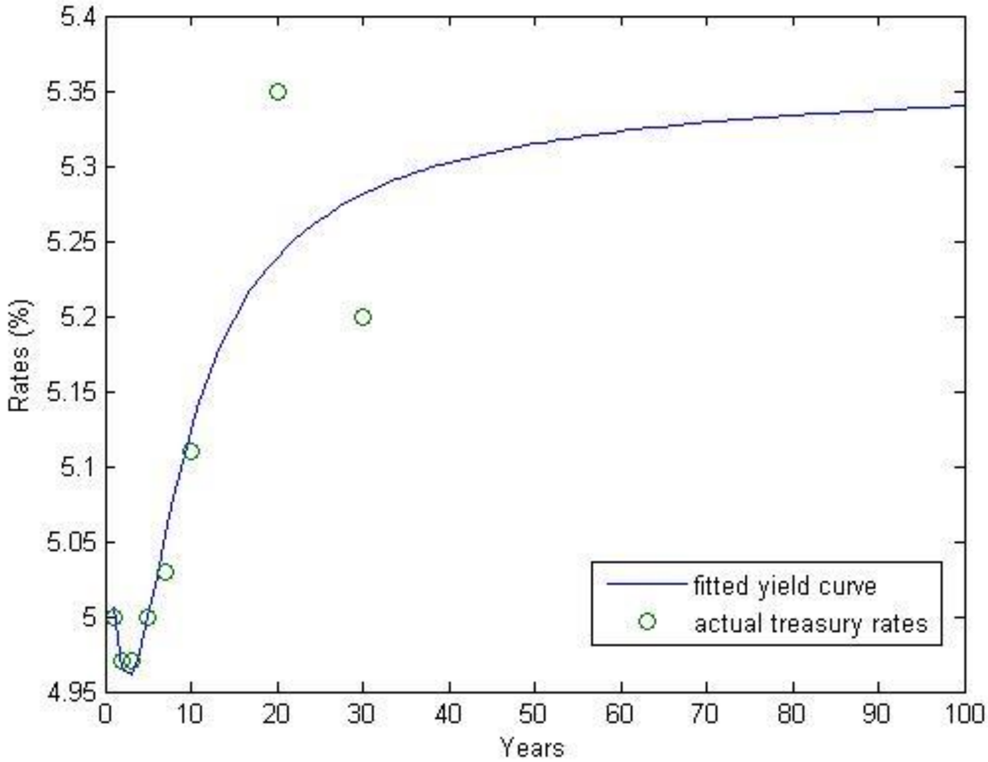


Figure 4.21: The accumulated discounted cash flows of each period by the three model.

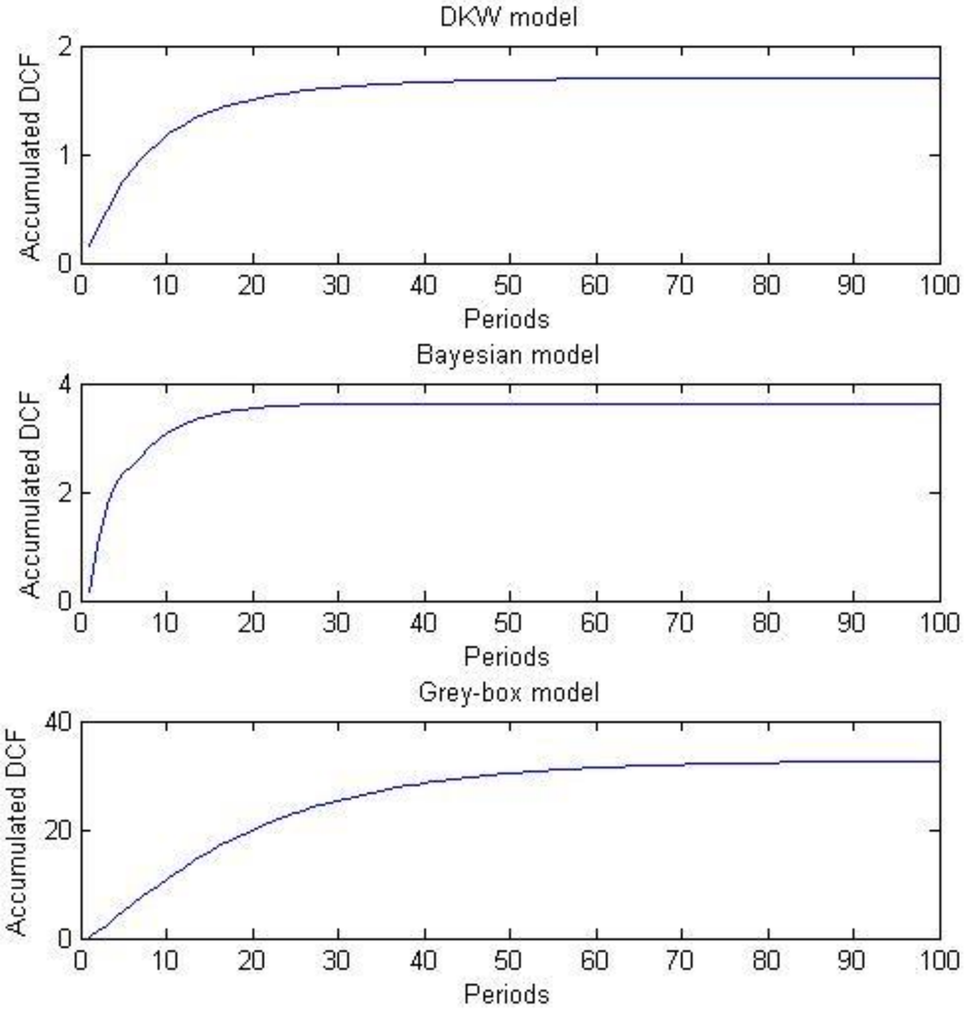


Figure 4.22: The RRA coefficients by DKW model and the contemporary S&P index.

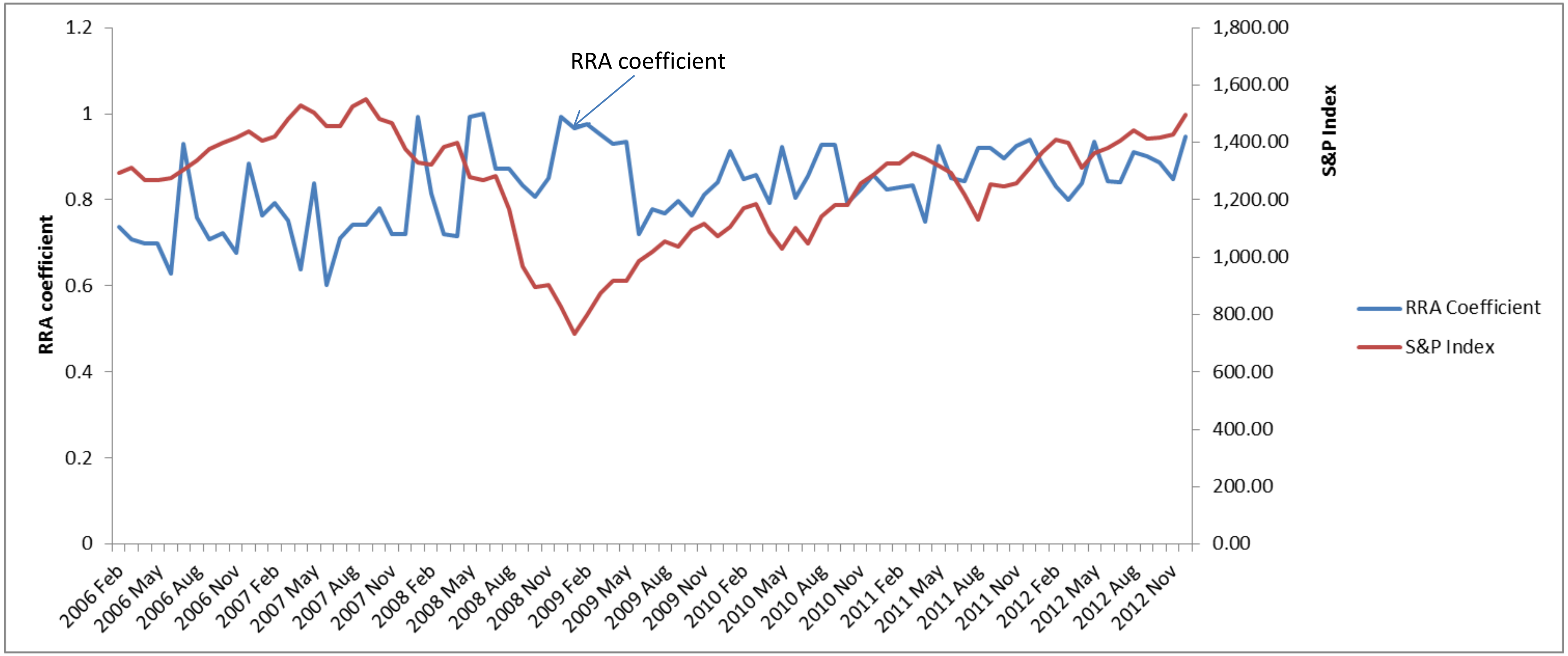


Figure 4.23: The RRA coefficients by grey-box model and the contemporary S&P index.

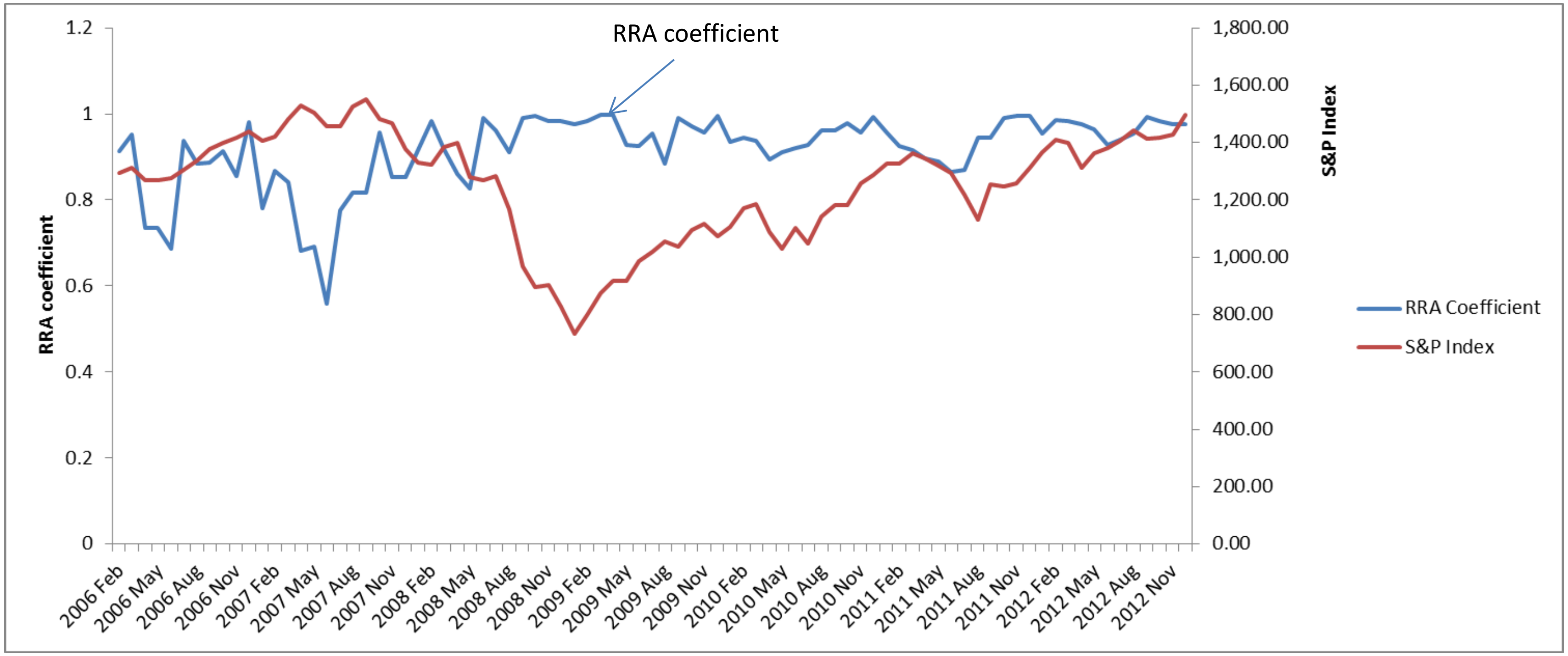


Figure 5.1: The trend of the ratio of net cash flow from operating activities to total assets with firm age increasing (U.K. listed firms).

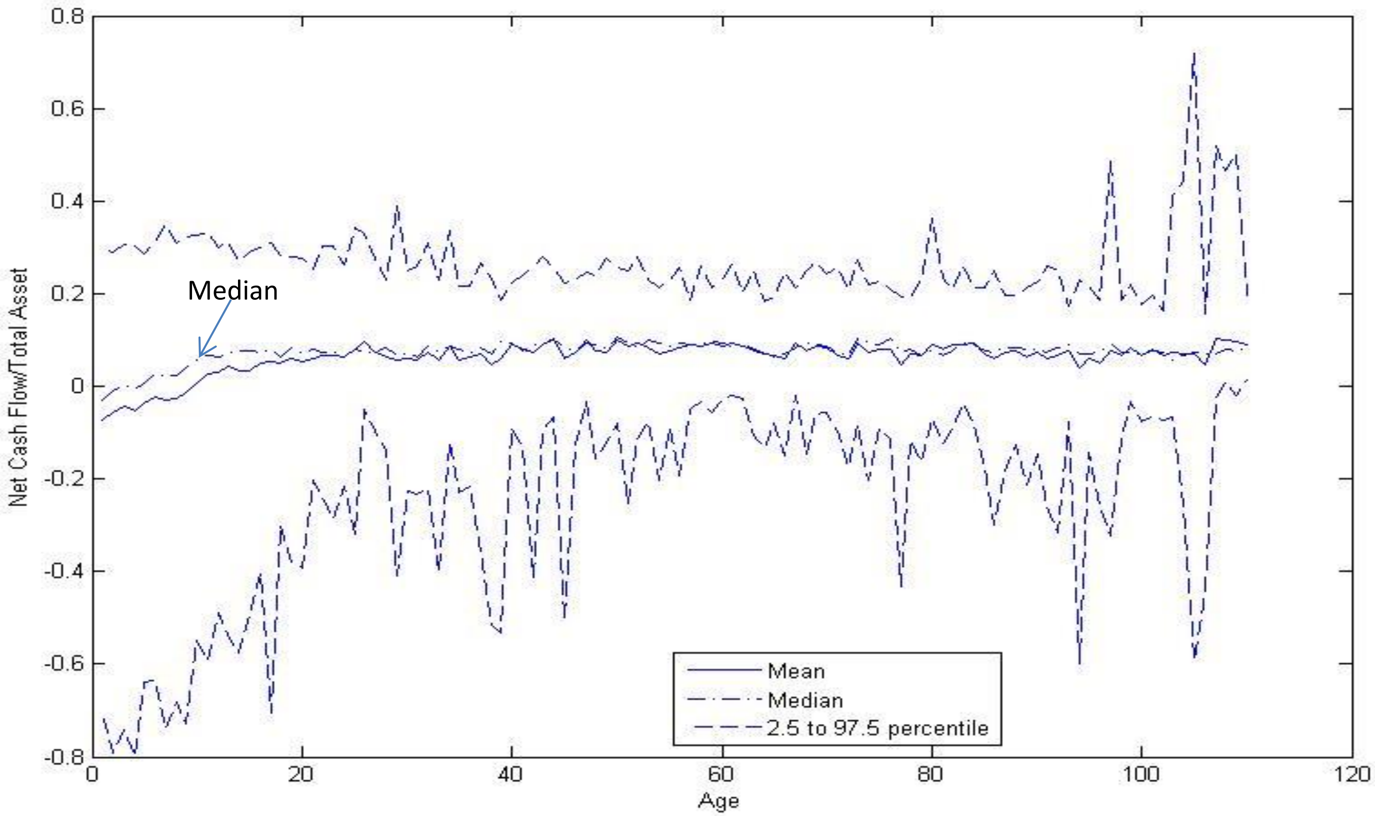


Figure 5.2: The association of parameter values and sales growth rates by grey-box model (U.K. listed firms).

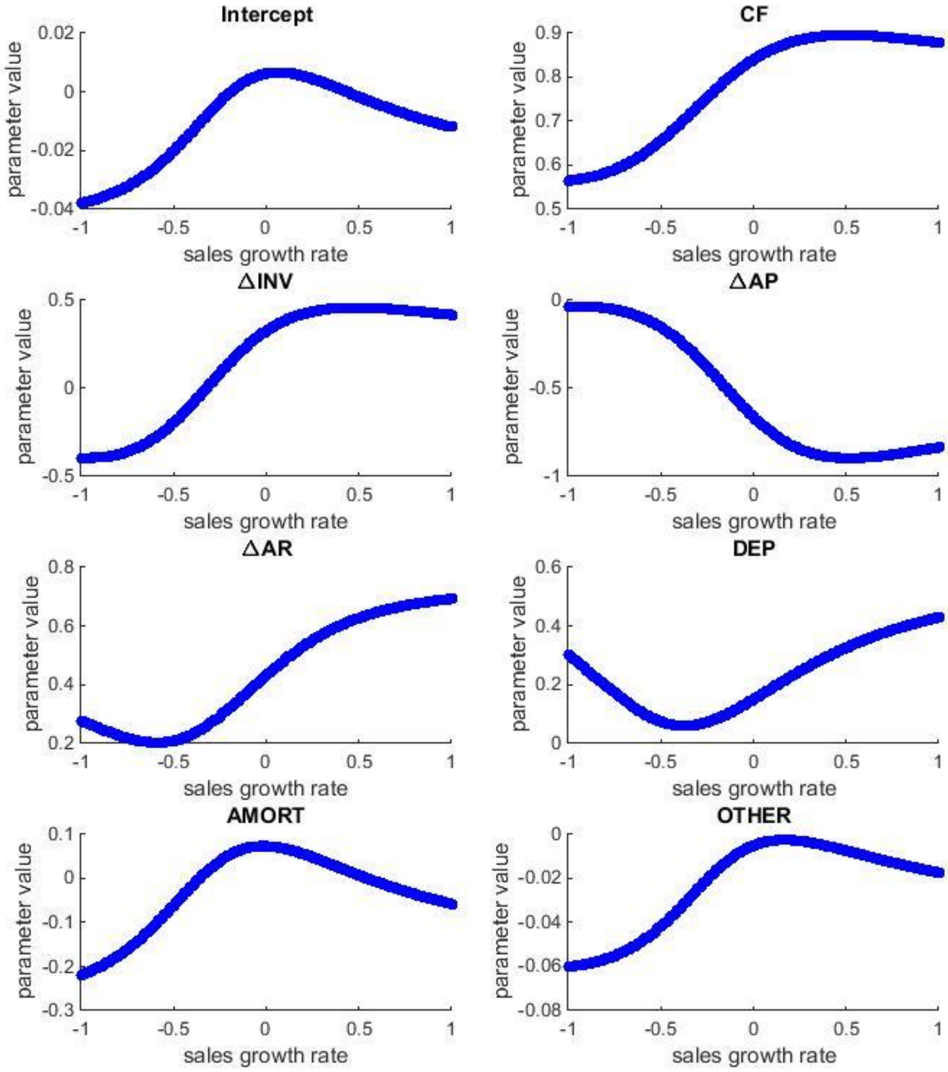


Figure 5.3: The evolution of parameters values with increasing firm age by grey-box model (U.K. listed firms).

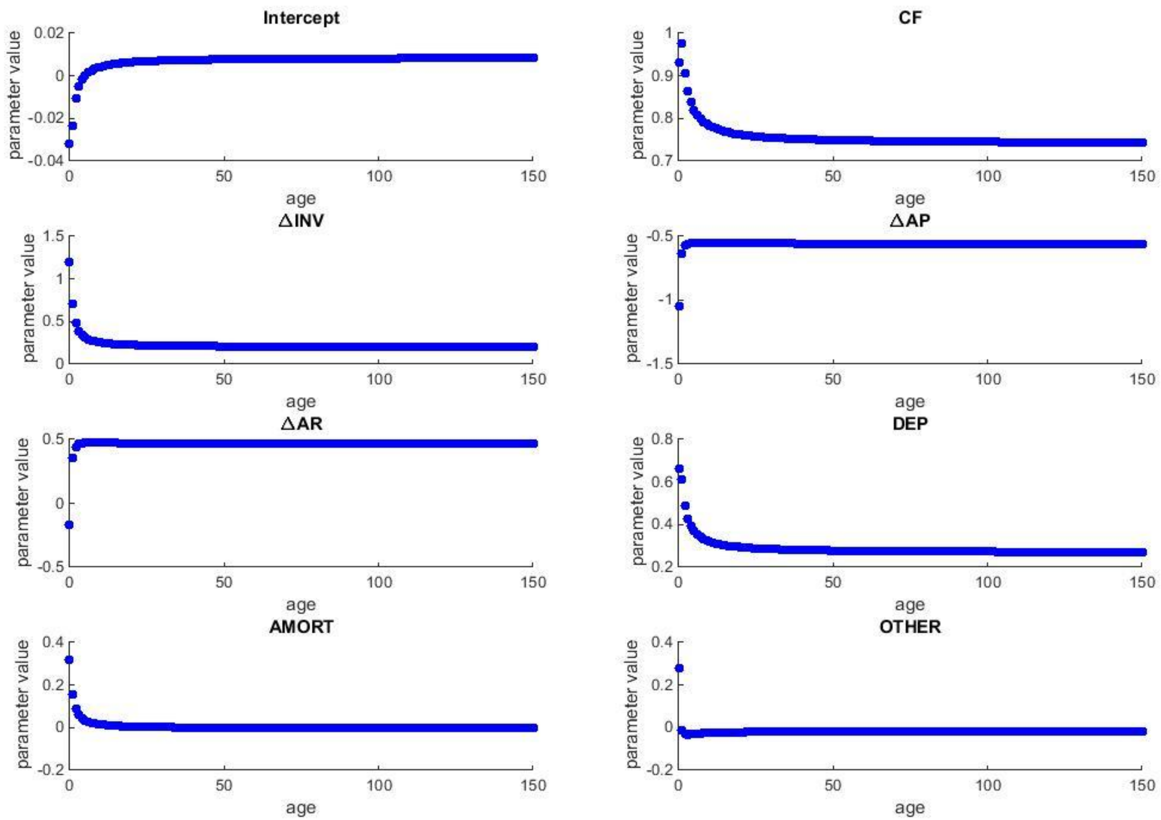


Figure 5.4: Calibrated RRA coefficients by four cash flow prediction models and discounted simulated cash flow model (U.K. listed firms).

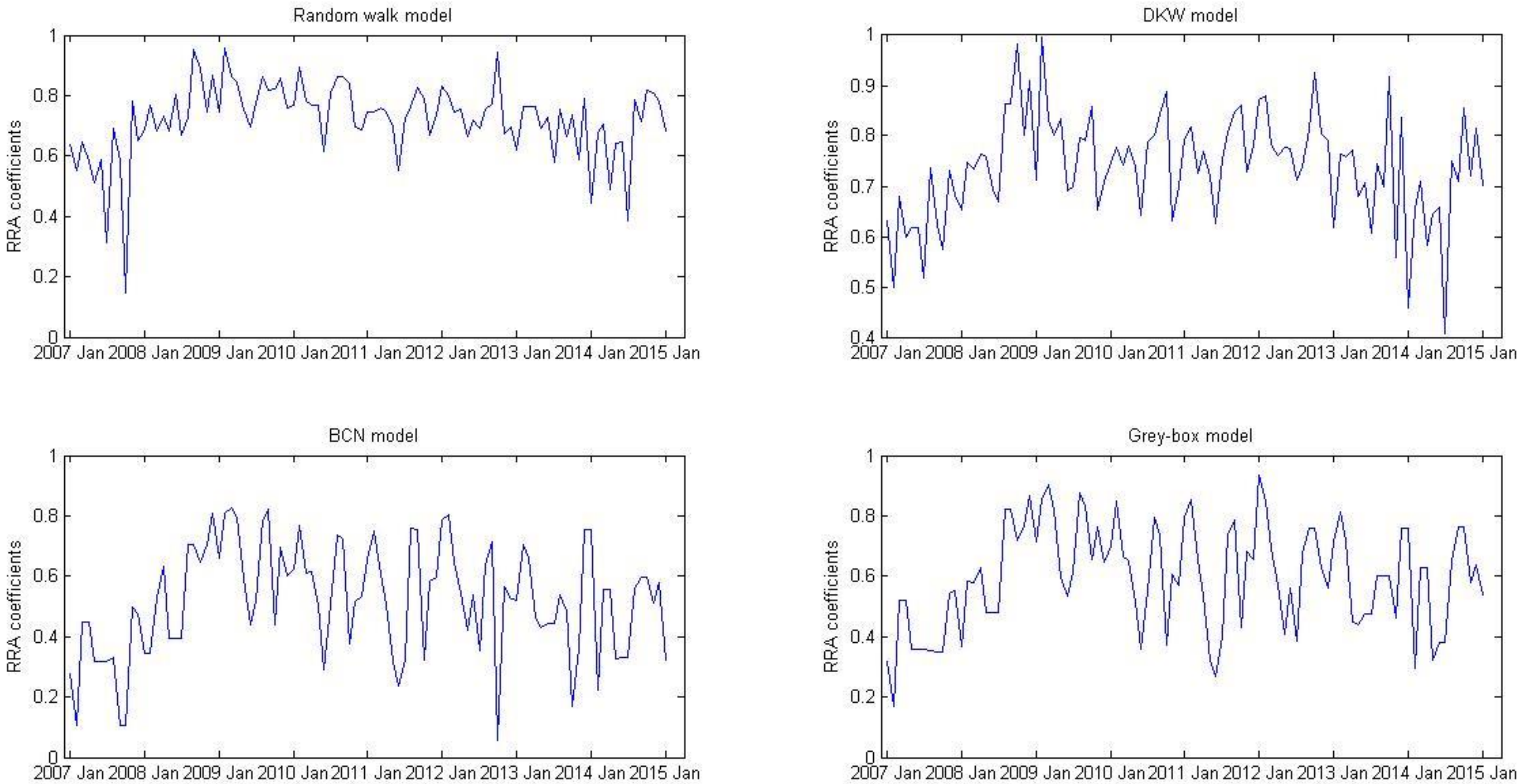


Figure 5.5: RRA coefficients calibrated by random walk model and contemporary FTSE All-Share index.

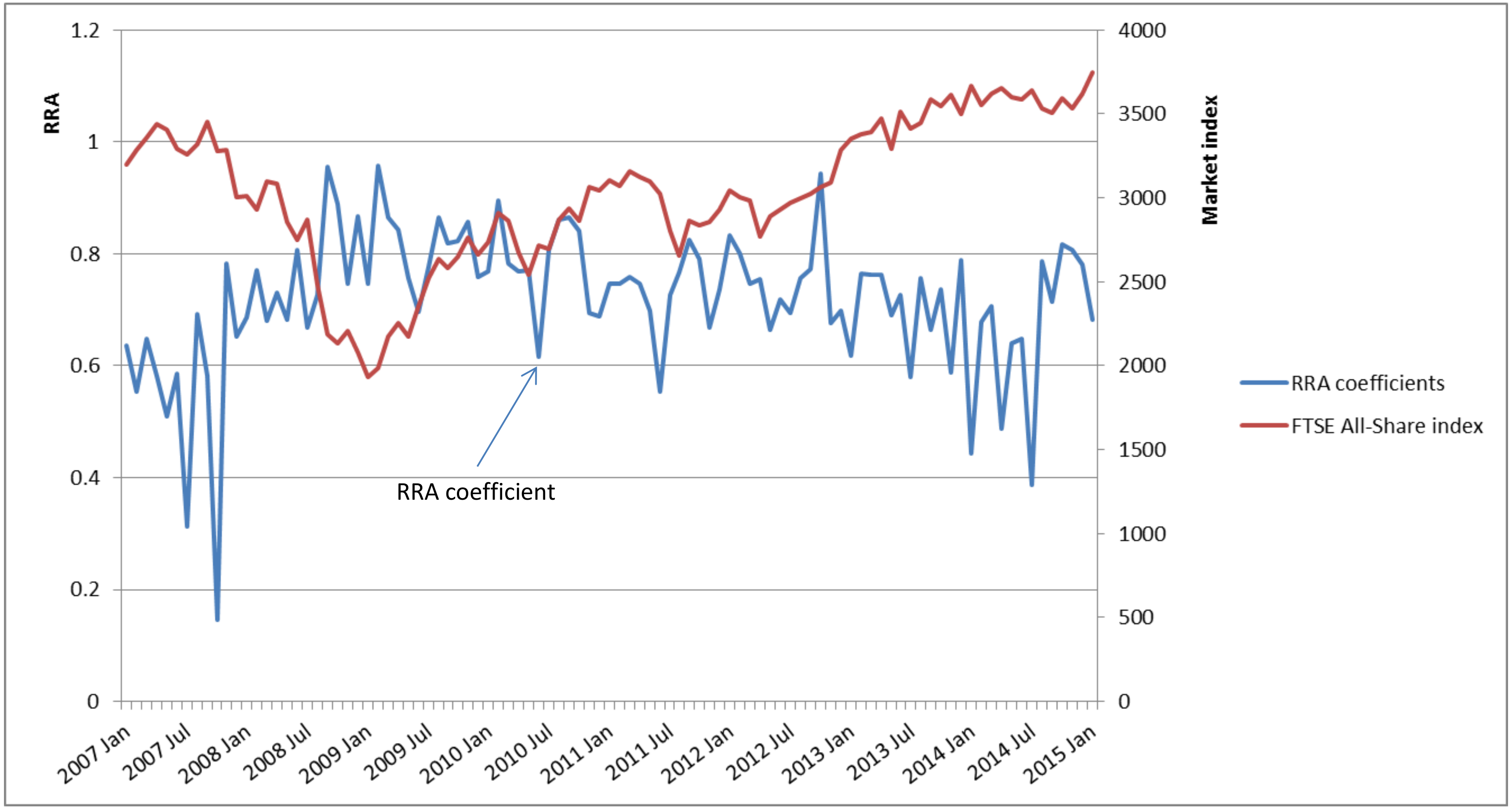


Figure 6.1: The trend of the ratio of net cash flow from operating activities to total assets with firm age increasing (U.K. unlisted firms).

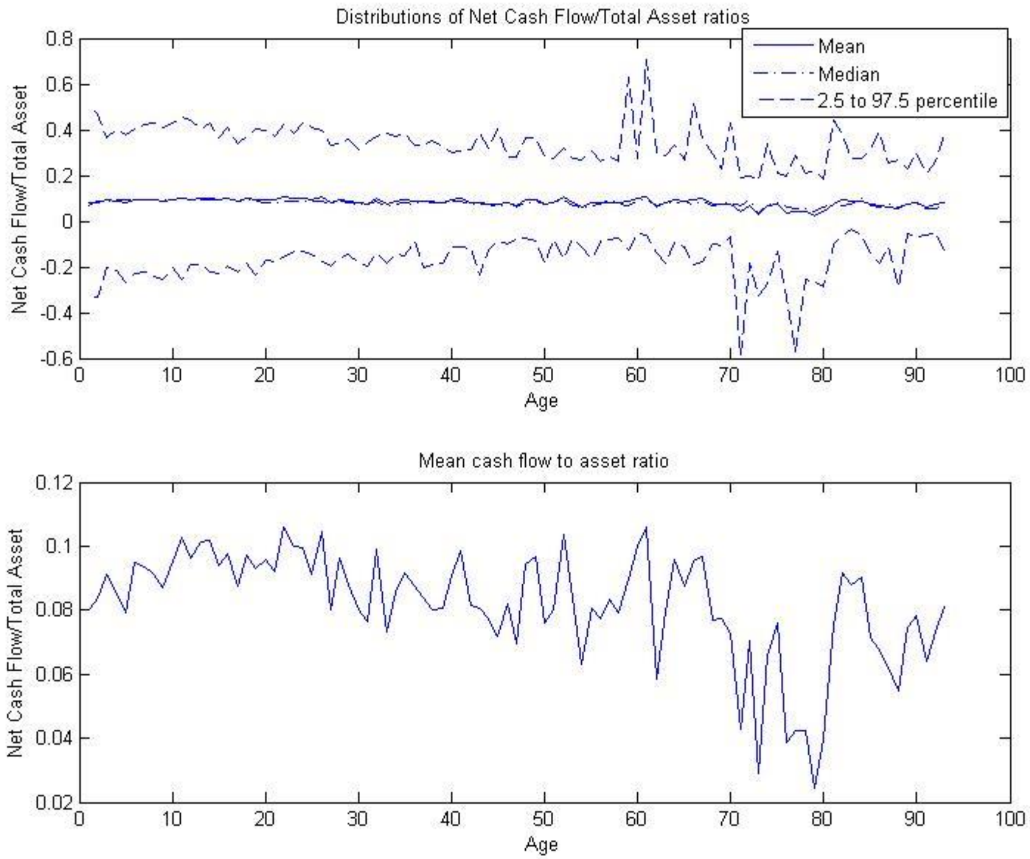


Figure 6.2: The association of parameter values and sales growth rates by grey-box model (U.K. unlisted firms).

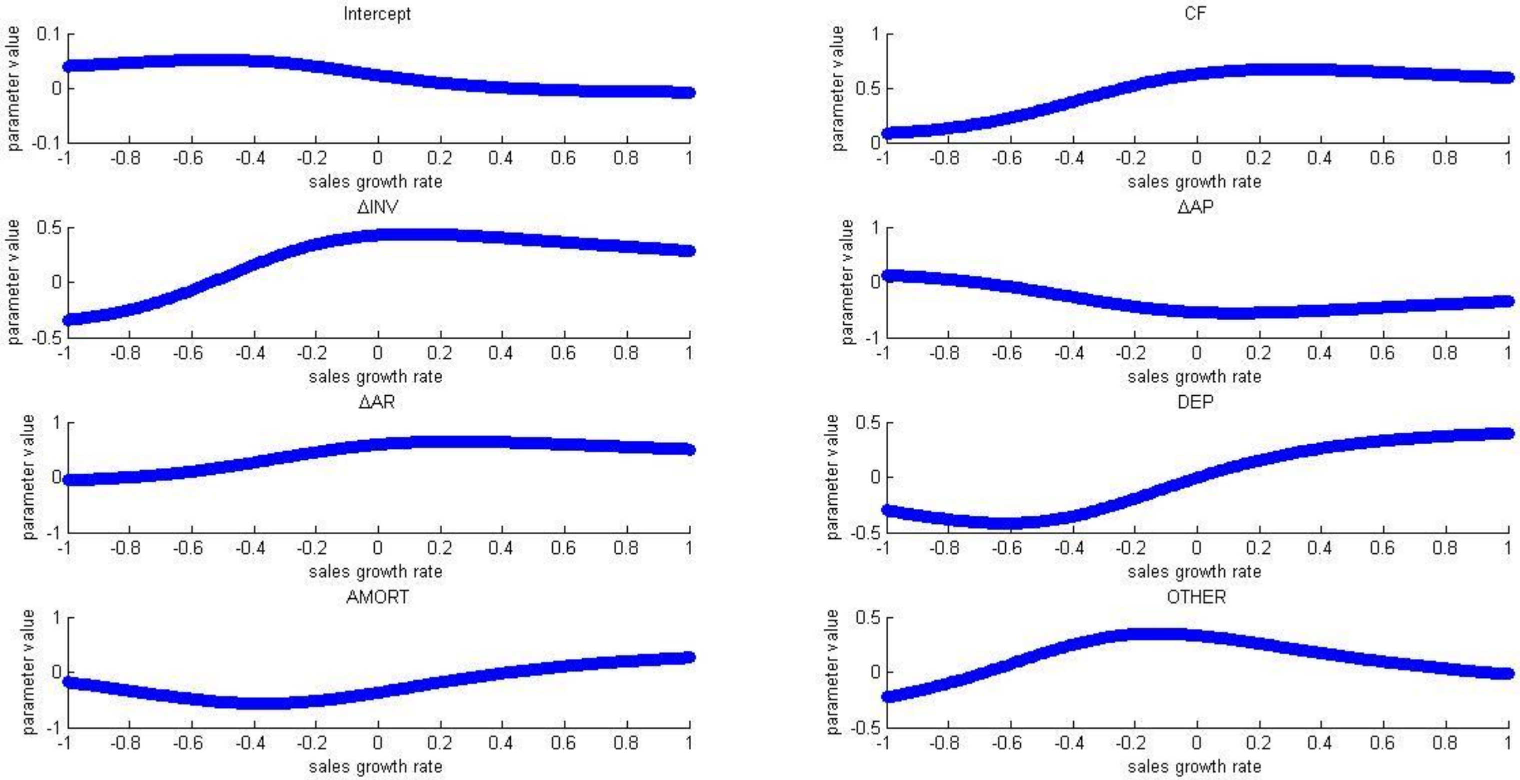


Figure 6.3: The evolvement of parameters values with increasing firm age by grey-box model (U.K. unlisted data).

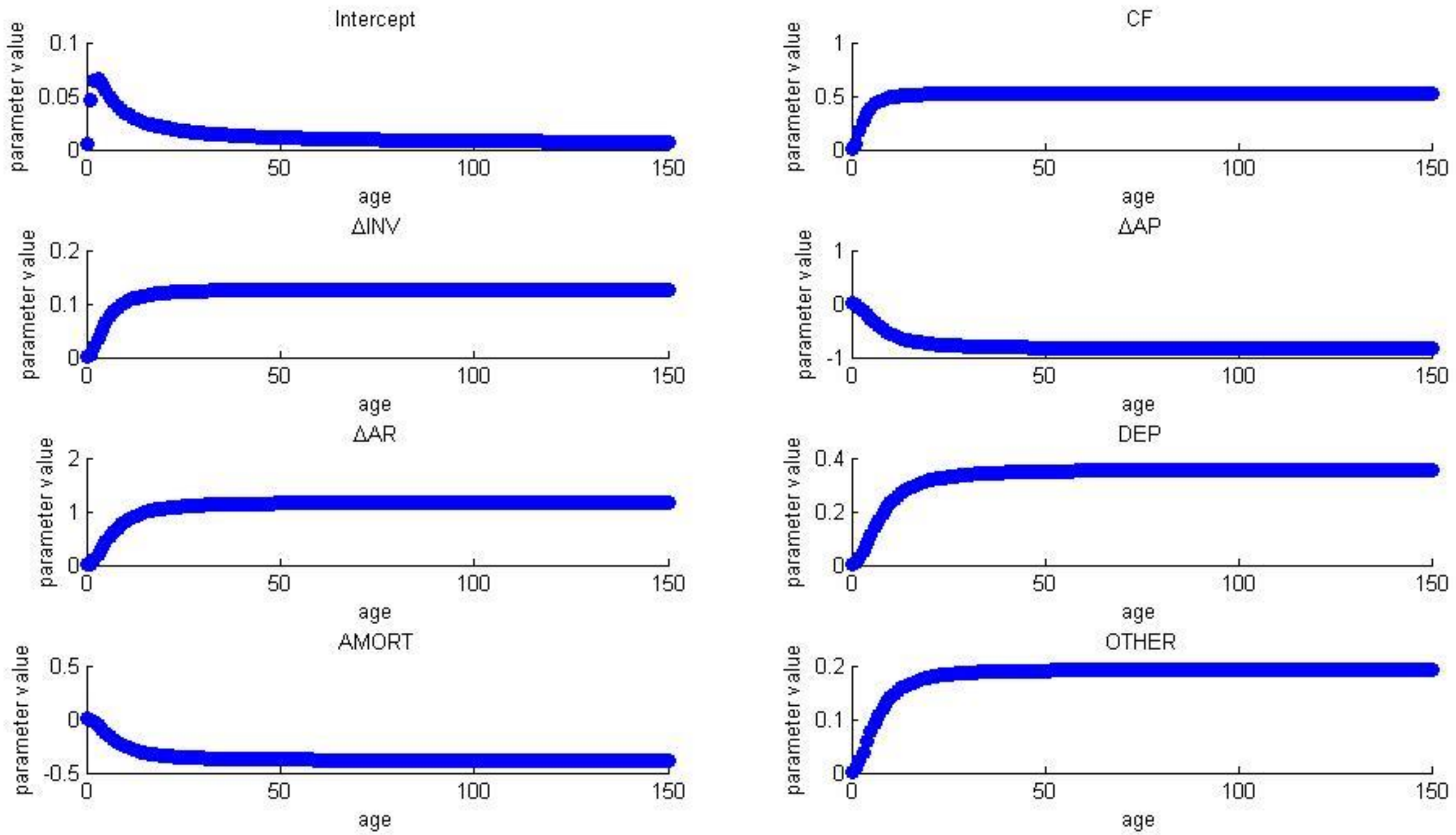


Figure 7.1: The trend of the ratio of net cash flow from operating activities to total assets with firm age increasing (China listed firms).

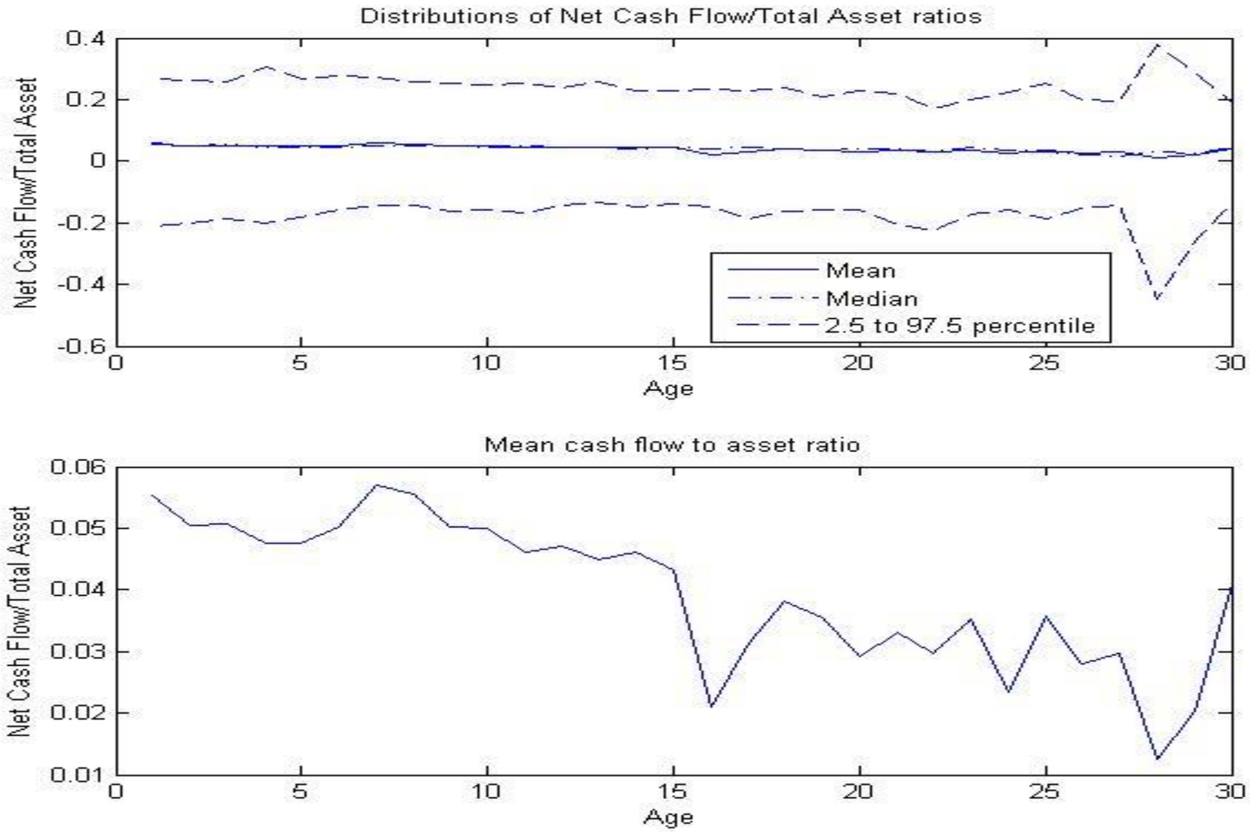


Figure 7.2: The association of parameter values and sales growth rates by grey-box model (China data).

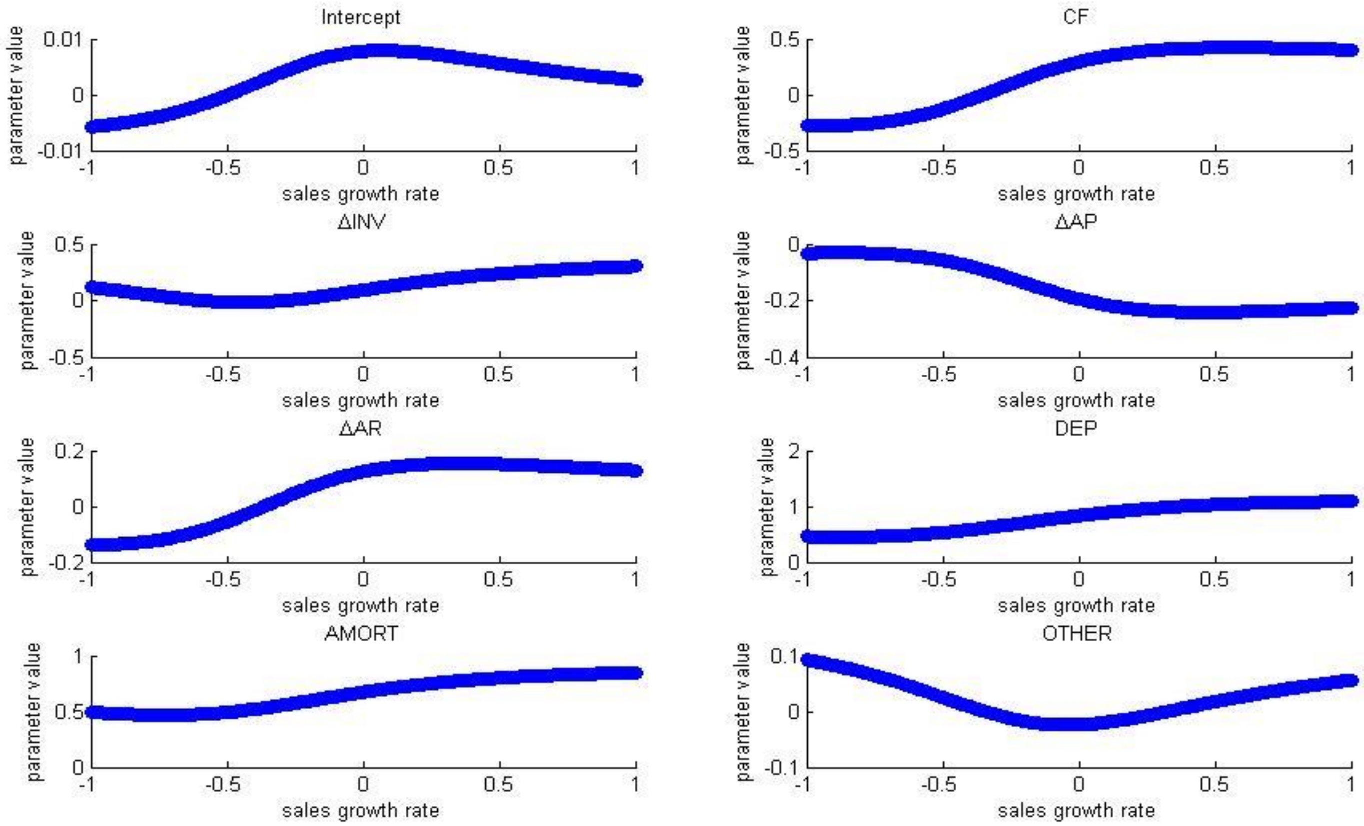


Figure 7.3: The evolution of parameters values with increasing firm age by grey-box model (China data).

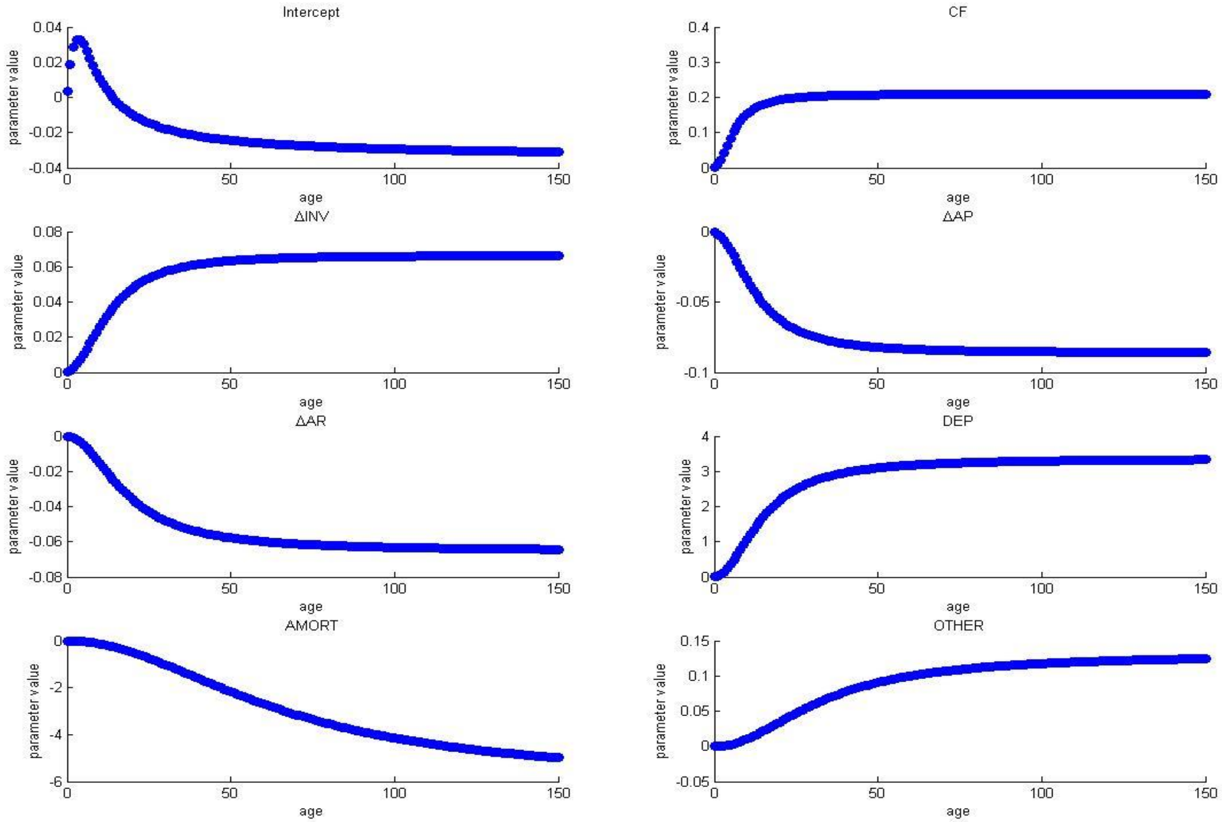


Figure 7.4: RRA coefficients calibrated by random walk model and contemporary Hushen 300 index.