SVERIGES RIKSBANK
OCCASIONAL PAPER SERIES

11

A Dynamic Microeconometric Simulation Model for Incorporated Businesses

Hovick Shahnazarian

AUGUST 2004

OCCASIONAL PAPERS ARE OBTAINABLE FROM
Sveriges Riksbank, Information Riksbank, SE-103 37 Stockholm
Fax international: +46 8 787 05 26
Telephone international: +46 8 787 01 00
E-mail: info@riksbank.se

ISSN 0349-3253
Print: Riksbank Printing Office
Stockholm 2004

A Dynamic Microeconometric Simulation Model for Incorporated Businesses

Hovick Shahnazarian

Sveriges Riksbank Occasional Paper Series
No. 11
August 2004

Sveriges Riksbank
Monetary Policy Department, Forecast Division
SE-103 37 Stockholm

Tel. +46 8 787 0554
E-mail: hovick.shahnazarian@riksbank.se

Fax +46 8 787 0169

Abstract

In this book, we describe the dynamic micro econometric simulation model for incorporated businesses. The purpose is to provide a model that both analyzes the behavior effects induced by changes in the tax code and forecasts the tax revenues. The basic idea is to combine the dynamic behavior of the corporate system with a statistical model that captures the development and the interrelationship between firms' different decision variables. *The dynamic behavior of the corporate system* is captured by several difference equations that identify how different variables in the firms' balance sheets change over time. To be able to do this we use the information in the firms' three basic financial statements: the balance sheet, the income statement, and the statement of changes in financial conditions. Furthermore, the difference equations system also incorporates special features of corporate taxation. *The firms' decisions regarding the flow variables* are modeled in a statistical module. From a dynamic optimization problem we derive the economic relationships between these flow variables and other economic variables. These relationships are then estimated using different robust estimation methods. In the next step, we insert the estimated functions from the statistical module into the difference equations system. This system is finally solved numerically to be able to simulate the future values of the stock variables in the firms' balance sheets.

Contents

	Page
Abstract	ii
Contents	iii
Preface	vi
List of Figures	ix
List of Tables	x
Introduction	1
1 The Different Modules in CIMOD	5
1.1 The Structure of the Modules in CIMOD	5
1.2 The Structure of the Simulation Module	5
1.3 The Structure of the Statistical Module	8
2 The Simulation Module	10
2.1 The Balance Sheet	10
2.2 The Income Statement	12
2.3 The Balance Sheet and its Dynamic Characteristic	16
2.4 The Statement of Changes in Financial Condition	18
2.5 The Constraints on Firms’ Dividend Payments Decision	19
2.6 The Constraints on Firms’ Tax Depreciation of Machinery and Equipment	19
2.7 The Constraint on Firms’ Allocations to the Periodical Reserves	22
2.8 The System Dynamics for the Firm	23
3 The Dynamic Optimization Model	25
3.1 The Value of the Firms and Their Cash Flow Constraint	25
3.2 The Constraints on the Firms’ Financial Decision	30
3.3 The Firms’ Objective and the Optimality Conditions	32
3.4 The Economic Relationships between Different Decision Variables	37
3.5 The Econometric Models	45
3.5.1 The Economic Relationships for Different Income Statement Items	46
3.5.2 The Recursive System Used to Estimate Firms’ Decision Variables	49
4 Description of Data	53
4.1 The Source of the Database	53
4.2 The Sampling Construction	54
4.3 Check for Inconsistencies and Errors	56
5 Different Variables Used in the Statistical Module	58
5.1 Legal and Accounting Constraints on the Firms’ Behavior	58
5.2 Different Macroeconomic Variables	59
5.3 The Maturity of the Firms	60
5.4 The Location of the Firms	60
5.5 The Concentration of Incorporated Firms within the Industry	61
5.6 Controlling for the Decisions Made by the Firms	61
6 The Statistical Module	63
6.1 Influence Diagnostics	63
6.2 Robust Estimation	64
6.3 Logistic Models with Complementary Log-Log Function	72
6.4 Multinomial Models with Complementary Log-Log Function	73

6.5 The Tobit Model with a Logistic Distribution 74

7 The Estimation Results 79

7.1 Estimating Economic Depreciation of Machinery and Equipment 80

7.1.1 The Probability that Economic Depreciation of Machinery and Equipment is Positive 80

7.1.2 The Level of Economic Depreciation of Machinery and Equipment 83

7.2 Estimating Net Sales of Machinery and Equipment 83

7.2.1 The Probabilities that Firms make Net Sales of their Machinery and Equipment 84

7.2.2 The Level of Positive Sales of Machinery and Equipment 85

7.2.3 The Level of Negative Sales of Machinery and Equipment 85

7.3 Estimating Investment in Machinery and Equipment 86

7.3.1 The TOBIT Model for Investment in Machinery and Equipment 86

7.4 Estimating Economic Depreciation of Buildings 87

7.4.1 The Probability that Economic Depreciation of Buildings is Positive 88

7.4.2 The Level of Economic Depreciation of Buildings 89

7.5 Estimating Investment in Buildings 89

7.5.1 The Probability that Investment in Buildings is Either Positive or Negative 90

7.5.2 The Level of Positive and Negative Investment in Buildings 91

7.6 Estimating the Net Change in Other Fixed Assets 91

7.6.1 The Probability that the Net Change in Other Fixed Assets is Positive 92

7.6.2 The Probability that Investment in Buildings is Negative 93

7.6.3 The Level of the Positive Net Change in Other Fixed Assets 94

7.6.4 The Level of the Negative Net Change in Other Fixed Assets 94

7.7 Estimating the Net Change in Current Assets 95

7.7.1 Estimation Results for the Net Change in Current Assets 95

7.8 Estimating the Net Change in Long-Term Liabilities 96

7.8.1 The Probability that the Net Change in Long-Term Liabilities is Either Positive or Negative 96

7.8.2 The Level of the Positive and Negative Net Change in Long-Term Liabilities 97

7.9 Estimating the Net Change in Current Liabilities 98

7.9.1 Estimation Results for the Net Change in Current Liabilities 98

7.10 Estimating the Net Change in Share Capital 99

7.10.1 The Probability that the Net Change in Share Capital is Positive 100

7.10.2 The Probability that Investment in Buildings is Negative 100

7.10.3 The Level of the Positive Net Change in Share Capital 101

7.10.4 The Level of the Negative Net Change in Share Capital 102

7.11 Estimating the Net Changes in Restricted Reserves 103

7.11.1 The Probability that the Net Changes in Restricted Reserves are Either Positive or Negative 103

7.11.2 The Level of the Positive and Negative Net Changes in Restricted Reserves 104

7.12 Estimating Operating Income Before Depreciation 105

7.12.1 Estimation Results for Operating Income Before Depreciation 105

7.13 Estimating Firms' Financial Income 106

7.13.1 The Probability that Firms' Financial Income is Positive 106
7.13.2 The Level of Firms' Financial Income 107
7.14 Estimating Firms' Financial Expenses 108
7.14.1 The Probability that Firms' Financial Expenses are Positive 108
7.14.2 The Level of Firms' Financial Expenses 109
7.15 Estimating the Tax Depreciation of Machinery and Equipment 110
7.15.1 The Tobit Model for Positive Tax Depreciation of Machinery and Equipment 110
7.16 Estimating Reversals from Periodical Reserves 111
7.16.1 The Probability that Reversals from Periodical Reserves are Positive 111
7.16.2 The Level of Reversals from Periodical Reserves 112
7.17 Estimating the Net Change in Other Untaxed Reserves 113
7.17.1 The Probability that Firms Make Net Changes in Other Untaxed Reserves 113
7.17.2 The Level of the Positive Net Change in Other Untaxed Reserves 114
7.17.3 The Level of the Negative Net Change in Other Untaxed Reserves 114
7.18 Estimating the Net Group Contribution 115
7.18.1 The Probability that the Net Group Contribution is Positive 116
7.18.2 The Probability that Investment in Buildings is Negative 116
7.18.3 The Level of the Positive Net Group Contribution 117
7.18.4 The Level of the Negative Net Group Contribution 118
7.19 Estimating Other Allocations 119
7.19.1 The Probability that Firms Make Other Allocations 119
7.19.2 The Level of the Positive Other Allocations 119
7.19.3 The Level of the Negative Other Allocations 120
7.20 Estimating Firms' Tax Liability 121
7.20.1 The Probability that Firms' Tax Liability is Either Positive or Negative 121
7.20.2 The Level of the Positive and Negative Tax Liability 122
7.21 Estimating Other Tax Adjustments 123
7.21.1 The Probability that Firms Make Other Tax Adjustments 123
7.21.2 The Level of the Positive Other Tax Adjustments 124
7.21.3 The Level of the Negative Other Tax Adjustments 124
7.22 Estimating the Tax Depreciation of Buildings 125
7.22.1 The Tobit Model for Positive Tax Depreciation of Buildings 125
7.23 Estimating Allocations to Periodical Reserves 126
7.23.1 Estimation Results for Allocations to Periodical Reserves 126
7.24 Estimating Firms' Reduction of Taxes 127
7.24.1 The Probability that Firms' Reduction of Taxes is Either Positive or Negative 127
7.24.2 The Level of the Positive and Negative Reduction of Taxes 128

8 The Evaluation of the Swedish Tax Model for Incorporated Businesses 130

8.1 Financial Ratio Analysis 130
8.1.1 The Liquidity Ratio 130
8.1.2 Leverage Ratios 131
8.1.2.1 Leverage Ratios Based on the Balance Sheet 131
8.1.2.2 Leverage Ratios Based on the Income Statement 132
8.1.3 Profitability Ratios 132
8.1.4 Market Value Ratios 134

8.2 Simulation Results Using Current Tax Rules 135
8.3 Forecasting Accuracy 137
8.4 Simulation Results for a Proposed Corporate Tax Rate Reduction of Three Percent 138
8.5 Simulation Results for a Hypothetical Change in Macroeconomic Developments 139

9 Future Developments and Improvements 142

Conclusions 144

Appendix A List of Variables 146
Appendix B The Framework of the Simulation Model within a Simple Model with Two Assets 150
Appendix C General Structure and Assumptions Underlying the Classical Regression Models 157

Figures 166

Tables 170

References 261

Preface

My interest in microsimulation started to develop when I took a graduate course in systems and modeling. These modeling approaches promised the possibility to represent the corporate system by simulating the behavior of individual microeconomic units on a computer. I found that the construction of a large-scale microsimulation model for firms could yield much more detailed results than the existing models. I discussed these ideas with my former supervisor, Professor Jan Södersten, in 1992. He found the idea promising (and risky) but unsuitable for a Ph.D. thesis. He advised me to choose a different subject for my thesis. Today, I am grateful to him for his advice, given the time, resources, and efforts that have been necessary to develop such a simulation model.

After my graduate studies, I received an offer from the Ministry of Finance to develop a dynamic simulation model for incorporated firms. It was Anders Kristoffersson, director at the division for taxation policy, who sought a model that both analyzed the behavioral effects induced by changes in the tax code and forecast the tax revenues. I was surprised that someone was willing to invest in such a risky project. I found the project very stimulating and challenging. It gave me the opportunity to explore my old ideas. This book is the result of that project, which was labeled the CIMOD project. CIMOD stands for Corporate and Individual tax policy simulation and forecasting MODel. The project was initiated in 1996 and was organized as follows:

Leading body: Anders Kristoffersson (chairman) and Hovick Shahnazarian (advisor)
Expert panel: Prof. Jan Södersten (expert on corporate taxation) and Prof. Anders Klevmarken (expert on econometrics and dynamic microeconometric simulation models)
Project leader: Hovick Shahnazarian
Research team: Peter Brose (1996-2001) and Altin Vejsiu (2001-2002)
Research Assistants: Claes Tidanå and Hanna Ågren

The simulation model presented in this book is the result of the successful cooperation between all members in the CIMOD project. The project has benefited considerably from Peter Brose's programming skills as well as his modeling experience. Developing the computer implementations has been an integral part of the research. My cooperation with Peter regarding the translation of initially vague ideas into mathematical formulae, and subsequently into a structured computer language, has been very rewarding. In addition to Peter, I have had the privilege to work with another proficient person, namely Altin Vejsiu. His curiosity and information retrieval have improved the statistical module of the simulation model. Peter's and Altin's hard work and ingenuity have been very valuable for the CIMOD project and are gratefully acknowledged. Thank you guys!

I owe my deepest gratitude to Anders Kristoffersson, who has been the chairman of the CIMOD project. His expert comments have greatly improved the quality of the simulation model. His support and faith in both the project and me have been invaluable, as have our theoretical and practical discussions.

This book has benefited considerably from comments by Jan Södersten and Anders Klevmarken from the expert panel.

I have also benefited from stimulating discussions with Bo Lindén. He introduced me to the exciting field of corporate accounting. I enjoyed our conversations, which provided me with a deeper insight into the theory and practice of corporate accounting.

I would also like to thank Claes Tidanå and Hanna Ågren for research assistance. Airi Ekström and Christina Hedenborg are acknowledged for their assistance. Further, I would like to thank seminar participants at the Swedish Ministry of Finance for numerous comments and helpful discussions.

I have also enjoyed my discussions with Rickard Löfqvist and Sten Hanssen. I am grateful to them for reading different parts of my drafts and giving me useful comments and suggestions on how to improve them. Michael Howett at Sveriges Riksbank is acknowledged for checking for linguistic errors in this manuscript.

I am also indebted to Leif Johansson, Klas Lindström, Karin Kristensson, and Elisabeth Eklund, from Statistics Sweden, who helped the project to develop FRIDA (Firm Register and Individual DAtabases). I am also grateful to Rolf Johansson for his assistance during the data construction. Further, I would like to acknowledge the National Tax Board (especially Mats Douhan, Ken Lundberg, and Pia Löfgren) and the Swedish Patent and Registration Office for supplying data and for all their support during the development of FRIDA.

I would never have been able to complete the simulation model without the support of my wife and best friend Gun. The project has involved overtime during both evenings and weekends - time that I should have shared with my wife. Thank you for always being there for me when I needed your understanding, encouragement, support, good advice, and love.

List of Figures

Figure 1: The Structure of the Modules 166
Figure 2: The Structure of the Theoretical Module (cont.) 167
Figure 3: The Structure of the Theoretical Module 168
Figure 4: The Structure of the Statistical Module 169

List of Tables

Table 1: Estimation Results for $EDEP^{MA}$ 170
Table 2: Estimation Results for S^{MA} 174
Table 3: Estimation Results for I^{MA} 179
Table 4: Estimation Results for $EDEP^{BU}$ 181
Table 5: Estimation Results for I^{BU} 185
Table 6: Estimation Results for *dofa* 189
Table 7: Estimation Results for *dca* 195
Table 8: Estimation Results for *dll* 197
Table 9: Estimation Results for *dcl* 201
Table 10: Estimation Results for *dsc* 203
Table 11: Estimation Results for *drr* 209
Table 12: Estimation Results for *OIBD* 212
Table 13: Estimation Results for *FI* 214
Table 14: Estimation Results for *FE* 217
Table 15: Estimation Results for $TDEP^{MA}$ 221
Table 16: Estimation Results for *ZPF* 223
Table 17: Estimation Results for *dour* 226
Table 18: Estimation Results for *GC* 230
Table 19: Estimation Results for *OA* 236
Table 20: Estimation Results for *TL* 240
Table 21: Estimation Results for *OTA* 243
Table 22: Estimation Results for $TDEP^{BU}$ 247
Table 23: Estimation Results for P^{allo} 249
Table 24: Estimation Results for *ROT* 251
Table 25: Simulation Results Using Current Rules, MSEK 255
Table 26: The Forecasting Accuracy in Year 2000 257
Table 27: Simulation Results for a Proposed Corporate Tax Reduction of 3 Percent, MSEK 258
Table 28: Simulation Results for an Alternative Macroeconomic Development, MSEK 259

Introduction

The Ministries of Finance in Sweden and other countries have to a large extent, when it comes to corporations, been restricted to theoretical evaluation of the behavioral effects induced by changes in the tax code. The lack of micro data has made it difficult to conduct a valid empirical investigation. Estimation of behavioral factors plays an important role in assessing the financial implications of proposals for change in the tax code. It is desirable that more qualified assessments are made possible. A change in this direction will imply a better revenue estimate with regard to the corporation tax.

Another main task of the Ministries of Finance in different countries is revenue forecasting. Corporation tax is an area where forecasts have been unreliable. The forecasting failure blurs the assumed connection between the development of the economy and income tax. The aggregate material used has not enabled the necessary provisions in the forecasts for the corporations' allocation of profits over time. The tax code allows for reservation of profits and different deductions, which grant corporations, tax deferral. It may be assumed that such provisions are utilized to a varying degree[1], depending on the individual firms' economic situation. To capture the individual firms' economic behavior, it is necessary to use micro data. The estimation of behavioral factors is thus an important part of improving the methods for forecasting corporation tax revenue.

The Ministries of Finance in different countries are largely dependent upon the development of micro simulation models for this purpose. In this book, we will describe the model for incorporated businesses that has been developed to deal with the needs of the Ministry of Finance in Sweden. We believe that such a model should be useful for other countries as well. The idea behind the dynamic microeconometric simulation approach used in this book can be expressed in the following way.

The simulation module: Initially, the variables in the database are divided into two different groups: stock variables and flow variables. The flow variables in our database are usually the firms' decision variables (most of them can be found in the firms' income statements). The stock variables of the firms are usually collected in the firms' balance sheets. Having done this, we identify the ways in which different flow variables affect different stock variables. After this identification, we specify several difference equations for the stock variables. To sum up, what we do is to specify several difference equations that identify how different variables in the firms' balance sheets change over time. To be able to do this we use the information in the firms' three basic financial statements: the balance sheet, the income statement, and the statement of changes in financial conditions. Furthermore, the difference equations system also incorporates special features of corporate taxation.[2]

This way of modeling has to our knowledge not extensively been used in the field of economics (or at least not in the field of corporate taxation and finance).[3] This could be due to

[1] See Forsling (1998) for a careful study of the utilization of different tax allowances.

[2] The system dynamic approach has more frequently been used in natural and technical sciences. This approach has also been used in the business field. An original reference is Forrester's (1961) industrial dynamic system. See also Kumar and Vrat (1989) and Clarke and Tobias (1995) for a review of system dynamic modeling concerning a corporation.

[3] However, there is an extensive literature in the field of dynamic optimization that indirectly uses a system dynamic approach. In the field of corporate taxation and finance there are many examples of this indirect use of a system approach. In this field the authors are usually interested in examining the impact of taxation on corporate financial policy and the cost of capital. Hall & Jorgenson (1967), Dorfman (1969), King (1974), Stiglitz (1973, 1985), Poterba & Summers (1985), King (1975), Bergström & Södersten (1981a), Summers (1981), Abel (1982), Hayashi (1982), Bergström & Södersten (1982b), Sinn (1987), Atkinson & Stiglitz (1980), Auerback (1983), Auerback (1986), Andersson and Norrman (1987), Osterberg (1989), Sinn (1990) Kanniainen & Södersten

the fact that the implementation of the idea into a functioning model is not an easy task. There is of course one exception. Tongeren (1995) uses this approach in microsimulation modeling of Dutch firms.[4,5] This is done by focusing on the relationship between the firms and the economy as a whole. However, Tongeren's simulation model explores micro-macroeconomic relationships. This is not the case in the model presented in this paper. The behavior of firms is instead decided by using econometric tools. In our model, we use macroeconomic variables as explanatory variables in our estimations of the behavior of the firms. However, we do not have repercussion from the behavior of the firms on the macroeconomic variables. What is common to Tongeren's simulation model and our model is the mathematical modeling of firms' three interrelated sets of accounts: the income statement, the cash flow statement and the balance sheet. In our case, we model these interrelationships in more detail, compared to Tongeren, because we are interested in more details about different variables in these accounts.

The statistical module (the behavior modeling): The specification of difference equations is the first element in a dynamic model. The decision element in the dynamic structure is the other tricky element in such a model. We know that a flow (or decision) variable is usually affected by other flow variables. Theoretically, it is possible to capture the simultaneous effect of different flow variables by analytical functions. This, of course, is not often the case. Thus, we come to the next step in our approach, which is to identify the interrelationship between different flow variables. The behavior of the firms are modeled and estimated in two steps. We use a dynamic optimization model to derive the economic relationships between different decision variables and their relationships with other economic variables.

The relationships between different decision variables that are identified in the economic model are estimated using pooled data[6]. Depending on the nature of the variable, we use one of the following 8 different estimation methods:

- Huber-Schweppe robust estimation method.
- For those variables that only take non-negative values, we use a logistic model with the *cumulative logistic distribution function* to find the probability that the variable is positive. Then, we use the Huber-Schweppe robust estimation method to estimate the positive level of the variable.
- For those variables that only take non-negative values, we use a logistic model with the *complementary log-log distribution function* to find the probability that the variable is positive. Then, we use the Huber-Schweppe robust estimation method to estimate the positive level of the variable.
- For those variables that can be either negative, zero or positive, we use a logistic model with the *cumulative logistic distribution function* to find the probability that the variable is positive. Second, we use another logistic model with the *cumulative logistic distribution function* to find the probability that the variable is negative. Third, we use the Huber-Schweppe robust estimation method to estimate the positive level of the variable. Finally, we

(1994), Kanniainen & Södersten (1995) and Shahnazarian (1996) are examples of such studies. The models used in these papers are often a simplified version of basic system dynamic models. In this paper, we are able to show the symmetric characteristics between dynamic optimization problems used in corporate taxation models and the basic system dynamic approach. See further Chapter 3.

[4] The idea behind system dynamic modeling of corporate firms has also been called financial statement models. A very simple example of such modeling is given in Benninga (1992).

[5] The interested reader is referred to Tongeren (1995) for a detailed description of the connection between the micro simulation model of corporate firms and several sub-disciplines of economic sciences.

[6] The database consists of three years of cross-section accounting and supplementary taxation information. Each cross-section sample is supplemented by the information for the previous accounting year, thus rendering us with a three-year panel for every cross-section sample.

use the Huber-Schweppe robust estimation method to estimate the negative level of the variable.

- For those variables that can be either negative, zero or positive, we use a logistic model with the *complementary log-log distribution function* to find the probability that the variable is positive. Second, we use another logistic model with the *complementary log-log distribution function* to find the probability that the variable is negative. Third, we use the Huber-Schweppe robust estimation method to estimate the positive level of the variable. Finally, we use the Huber-Schweppe robust estimation method to estimate the negative level of the variable.
- For those variables that can be either negative, zero or positive, we use a multinomial model with the *complementary log-log distribution function* to find the probabilities that the variable is positive, equal to zero and positive. Second, we use the Huber-Schweppe robust estimation method to estimate the positive level of the variable. Finally, we use the Huber-Schweppe robust estimation method to estimate the negative level of the variable.
- For those variables that only take non-negative values, we use a *TOBIT model with a logistic distribution function*, which is a combination of a truncated regression model and a probit.
- For those variables that can be either negative, zero or positive, we use two *TOBIT models with a logistic distribution function*.

The regressors are of different kinds: they can include income statement variables, balance sheet variables, macroeconomic variables, dummy variables for the location of the firms, the market concentration variable, a variable for the firms' dominance in the market, the boundaries given by the tax and accounting rules, a dummy variable showing whether firms are small closed companies or not, etc.

The robust estimations of the level of the variables are done in different steps. First, we estimate the level of the decision variables using the method of ordinary least square. This method is applied to a polynomial model. Second, we undertake different checks and tests (test of normally distributed residuals, checking for the existence of multicoliniarity, test of heteroskedastic residuals). Third, we transform the variables to obtain homoskedastic residuals. Fourth, we check for observations with unusually large influence on the least square estimates and predictions. These observations arise because of heavy-tailed distributions. As a result, an examination of the residuals may be misleading. Accordingly, in a final step, we apply the Huber-Schweppes robust estimation method to modify least square procedures so that these observations have much less influence on the final estimates.

The database of the Ministry of Finance consists of both accounting and supplementary taxation information. Today, we have a database consisting of five years (1995-1999) of cross-section information for a sample of firms. The cross-section samples are supplemented by the information for the previous accounting year. The cross-section samples for 1997-1999 are also supplemented by the information for the two years before the accounting year. In order for the Ministry of Finance to apply and work with the model, we ignore the panel features of the model and instead pool the data. The reason for this is our belief that the population of firms is very dynamic because of bankruptcies, new entries, mergers, consolidation, etc.

This book is organized as follows. In Appendix B, we introduce the framework of the simulation model within a simple model with two assets. We strongly recommend the reader to read this appendix before continuing with Chapter 1. In this chapter, we briefly describe different modules in CIMOD. Chapter 2 presents the simulation module. In Chapter 3, we introduce the dynamic optimization model from which we obtain the economic relationships between different decision variables. In Chapter 4, we describe the data used in both the statistical and simulation modules. In Chapter 5, we define several variables, in addition to balance sheet and income statement variables, which we use as explanatory variables in the estimations of the firms' decision variables. In Chapter 6, we present different estimation methods that we use to estimate the decision variables. Appendix C gives an introduction to the general structure and assumptions behind the classical regression. We believe that this makes it easier for the reader to follow Chapter 6. Chapter 7 summarizes the estimation results. Due to considerations of space, we do not provide in-depth comments in sections 7.1-7.24 on the results. However, section 7.1 includes a more extensive explanation of the way the estimated coefficients should be interpreted. We suggest that the reader read this section carefully and skim (or skip) through the other sections. In Chapter 8, we evaluate the simulation model by analyzing the simulation results using both current tax rules and a hypothetical proposed corporate tax rate decrease of three percent. We also examine the forecasting accuracy of the simulation results. We conclude the chapter by simulating firms' behavior in an alternative macroeconomic environment. Chapter 9 outlines future developments and improvements that can be made in the simulation model.

1 The Different Modules in CIMOD

In section 1.1, we present the structure of the modules in CIMOD. Section 1.2 briefly discusses the structure of the simulation module. Finally, in section 1.3, we show the structure of the statistical module.

1.1 The Structure of the Modules in CIMOD

To be complete, a simulation model should consist of five different sub-modules. The structure of these modules is summarized in Figure 1. The simulation is made in different steps. First, the model must be able to create new companies. Second, the simulation model must be able to forecast whether existing firms in the data set will go into liquidation. Third, there is a possibility that firms included in the data set will change their organization form during the simulation. Therefore, a complete simulation model must include a module, in which the probability of changing the organization form is determined. Fourth, having established the demographic structure, the model then estimates the companies' decision variables. These estimated variables are then used in the simulation module where several different difference equations are specified to identify how different variables in the firms' balance sheets change over time. To be able to specify these difference equations the simulation module uses the information in the firms' three basic financial statements: the balance sheet, the income statement and the statement of changes in financial conditions (the cash flow statement).

As mentioned above, the first three modules are used to establish a correct demographic structure. These modules are not part of the CIMOD project. Nevertheless, we are aware of these questions, as they become more important when we have a functioning simulation model for incorporated firms. This is not the case today. The purpose of the CIMOD project has been to deliver such a model (module (4a) and (4b)).[7]

1.2 The Structure of the Simulation Module

CIMOD maintains three interrelated sets of accounts: a balance sheet account, a profit and loss account, and a statement of changes in financial conditions. Altogether, these three accounts give a financial description of the firms at a given moment in time. The three accounts will be formalized in Chapter 2. In this section, we introduce the idea behind the model using a block diagram of the flow of funds. Figure 2 depicts the financial flows of concern within a company.

Operating income before depreciation ($OIBD_t$) is the operating revenue remaining after operating expenses. Operating income before depreciation is split into two elements:

---

[7] We have, however, looked at a method for creating companies that was developed by the Ministry of Finance in Denmark. We believe that a similar method could be used in the first module. In the case of bankruptcy, we believe that the method developed by Altman (1968, 1984) is promising. This method is well-established and has proven to be very successful for forecasting the probability of bankruptcy. However, this method requires a database that includes both firms that have been liquidated and firms that have not. Our current database includes only non-liquidated firms. To be able to use Altman's method, we have to complete our selected database with additional selection, which includes liquidated firms. MacKie-Mason and Gordon (1992 and 1993) analyzed the choice of organization form in two papers. We believe that the methods used in these two papers could be used in the module for the choice of organization. These are questions for future research, and we will not comment on these issues anymore in this book. We hope that other researchers will supplement modules that establish the demographic structure during the simulation lapse.

provision for economic depreciation of machinery and equipment ($EDEP_t^{MA}$) and economic depreciation of buildings ($EDEP_t^{BU}$). These are purely bookkeeping operations: they identify the part of operating income before depreciation that is estimated as necessary to cover the cost of the deterioration of machinery, plant, etc. In themselves, these operations involve no flow of funds into or out of the company.

The reminder, operating income after economic depreciation ($OIAD_t$), is one part of the earnings before allocations (EBA_t). In addition, earnings before allocations include the financial income (FI_t). Financial income is interest income, dividends received on stocks and participations, income from affiliated undertakings, income from participating interests, income from other investments and loans forming part of the fixed assets (with a separate indication of that derived from affiliated undertakings), profits on operations disposed or closed, gains from the sale of business lines or facilitates, profits on sales of capital assets, profits on sales of equipment, facilities, etc., and exchange rate profits. Furthermore, earnings before allocations contains financial expenses (FE_t), which include interest costs (on all types of short- and long-term borrowing), the value adjustments of fixed assets[8], losses on operations disposed or closed, losses on sales of business lines or facilitates, losses on sales of capital assets, losses on sales of equipment, facilities, etc., and exchange rate losses.

By adding net allocations to earnings before allocations we obtain earnings before taxes (EBT_t). Net allocations include:

1. Allocations to accumulated supplementary depreciation (ΔASD), which may be positive or negative. This is the difference between allowances (amortization) for depreciation (depreciation for income tax) and depreciation according to plan.
2. Allocations/reversals to/from periodical reserves, which include the allocation during the current period (ΔPF_t), and reversals from the periodical reserves from period *t-6*, *t-5*, *t-4*, *t-3*, *t-2*, and *t-1*: ΔPF_{t-6}, ΔPF_{t-5}, ΔPF_{t-4}, ΔPF_{t-3}, ΔPF_{t-2}, and ΔPF_{t-1}.
3. Other allocations (OA_t), which can also be both positive and negative. Among other things, other allocations also include received and given group contributions (net group contributions- GC_t).[9]

Allocations to untaxed reserves are purely bookkeeping operations. They identify the part of earnings before allocations that is estimated as necessary (a) to cover the tax cost of the deterioration of machinery, plant, etc. and (b) to allocate for future investment purposes. In themselves, these operations involve no flow of funds into or out of the company.

In practice, firms' actual tax payments in a specific period have no clear connection to earnings before taxes. Therefore, we will denote the tax liability in a specific period (TL_t).[10]

---

[8] This is a purely bookkeeping operation: it identifies the part of earnings before allocations that is estimated as necessary to cover the decrease in value of fixed assets. In itself, this operation involves no flow of funds into or out of the company.

[9] This was at least the case from 1995-1997. The accounting rules were changed in 1998 regarding accounting for GC_t. Since 1998, firms have been allowed to book group contributions either as other untaxed reserves or as an adjustment of income for tax purposes (OTA_t).

[10] The reason for this is that firms usually close their books for one accounting period long before they fill in the tax return form. For this reason, firms must make a good estimate of their tax liability. For this, firms make the following considerations:
- For example, they estimate the different tax adjustments that they believe they will make in the tax return form. However, if the actual tax adjustments differ from the estimated tax adjustments, firms must pay penalty (extra) tax for underpayment of tax or receive a refund for overpayment of tax.

Net income (NI_t) is calculated after deducting the tax liability. However, the cash flow in a company depends on actual tax payments. Typically, these are not identical as tax is normally paid in arrears. By adjusting net income for tax purposes, (TA_t), firms are able to calculate the amount of tax they have to pay. However, firms may obtain a tax reduction (ROT_t), for example for income taxes paid abroad. The reminder, after the actual tax payments, is net business income (NBI_t).

The first financial decision, the proportion of funds to be retained, is then made by dividing net business income into dividends paid to shareholders, (DIV_{t-1})[11], the maximum amount available for dividends in the current period (the so-called cash flow ($cashfl_t$)), allocations to restricted reserves (drr_t), and retained earnings (the change in other unrestricted equity (ΔURE)). The total funds generated internally are then depreciation provisions, different allocations to untaxed reserves, allocations to restricted reserves and retained earnings.[12]

Adding new long-term debt finance (dll_t), new short-term debt finance (dcl_t), and new equity issues (dsc_t) in proportions determined by the firms' gearing decisions, we obtain the total of funds available to the firms. Long-term debt finance includes different items, for example long-term bank loans (at fixed or variable rates of interest), loans from other companies or governments, and allocations for pensions. Gathering all these together under one heading removes differences in marketability of long-term debt finance and differences in the extent to which the interest rate is variable. The item is net of repayment of outstanding loans. Current liabilities include mainly amounts owned to creditors, banks and other short-term loans. New equity issues cover the issues of all types of shares to raise more finance.

- Further, they must also include the different taxes that they pay abroad. There are two different methods of adjusting tax payments for foreign revenues and taxes: the credit method and the exempt method. In the case of the credit method, both taxes paid abroad and taxes paid in Sweden will be brought back among tax adjustments. Later, when the final tax has been calculated, Swedish tax authorities pay back their foreign tax as a credit. If the exempt method is used, firms are taxed on their worldwide income in Sweden, so that their revenues and costs are reported together with their Swedish revenues and costs. Further, the preliminary tax payment includes the foreign taxes paid (which is also the case when the credit method is used). In this case, firms do not bring back the taxes paid abroad among tax adjustments. Instead, they bring back their earnings abroad among tax adjustments. Different ways of adjusting for tax payments abroad cannot be calculated with the database we are currently working with. Instead, we use another approach to solve this problem. This is partly captured as a reduction of taxes.

[11] The dividend payment procedure can be concluded as follows. The board of directors proposes the dividend (usually in a financial statement bulletin) at the beginning of the year *t*. The dividend is paid to all those shareholders registered on the company's books on the holder-of-record date as owners of specific shares of stock. The right to the dividend remains with the stock until x business days before the holder-record-day (the so-called ex-dividend day). In order to receive the dividend, a new investor must purchase the stock before the ex-dividend day. From the ex-dividend day, the stock is said to sell ex-dividend. The stockholders' meeting adopts the balance sheet (normally the adopted balance sheet coincides with the preliminary balance sheet presented by the board of directors). The stockholders' meeting also decides the amount of dividends that should be paid, the allocations to be made to restricted reserves and free reserves, and the remaining retained earnings. However, once a dividend has been decided, it becomes a current liability of the corporation. After the stockholders' meeting, firms mail out the dividends as soon as possible. When this happens, the current liability is eliminated, and the firms' current assets decline.

[12] The sum of these is generally given the rather misleading term, "internal cash flow". The fact that depreciation provisions, allocations to restricted reserves and free reserves, and different allocations to untaxed reserves are sources of funds should not be taken to imply that an increase in them would make more funds available. Unless tax, dividends or allocations to restricted and free reserves are changed, the increase in them will be exactly offset by a fall in retained earnings.

These funds are shown as going towards four uses. The first two are net investment, predominantly in machinery, equipment and buildings (I_t^{MA} and I_t^{BU}). Second, funds can be used to increase other fixed assets (which also includes financial assets). Third, funds may also be used to build up current, i.e. easily realizable, assets (mainly stocks and work-in-progress, financial assets, short-term loans to debtors, and cash balances), (dcl_t).

There are three main types of accounting documents that describe different aspects of the flow of funds diagram. The first, normally required by law to be published annually, is the income statement. This is essentially the statement of the top half of Figure 2. A second document, which some firms provide, is a statement of Sources and Uses of Funds (the cash flow statement in Figure 2).

Both statements presented so far are flow statements, showing various financial flows occurring during the course of one year. The third statement is the Balance Sheet (Figure 3), which is a stock concept. It shows the value of various company assets and liabilities outstanding at the end of each financial year. Like the income statement, the balance sheet is required by law to be published. To obtain this, we use the summarized sources and uses of funds table in Figure 2. As the two columns in Figure 3 give the same total in any given time period, the sum of each for all previous time periods up to the Balance Sheet date will also be equal. The sum of all periods' new long-term net loans is the total outstanding *long-term liabilities* on the date specified. Similarly, the sum of all periods' net equity issues is the total outstanding equity. Moreover, the sum of all periods' net restricted reserves, net free reserves, and untaxed reserves is the total outstanding restricted reserves, free reserves, and different untaxed reserves.

On the left-hand side, the totals are: first, the sum of all periods' net increase in current assets is the total outstanding current assets on the date specified. Second, the total expenditure on machinery and equipment minus the accumulated depreciation and the sale of these assets for all periods is the current book value of firms' machinery and equipment. Third, the total net expenditure on buildings minus the accumulated depreciation is the current book value of firms' buildings. Fourth, the sum of all periods' net increase in other fixed assets is the total outstanding other fixed assets on the date specified. The total of either column is known as capital employed and is the most common measure of the resources available to the firms, over the long term, with which to earn profit.

1.3 The Structure of the Statistical Module

Before we move on to the description of the statistical module, let us have a closer look at what is usually known as the firms' cash flow cycle. Firms begin by issuing various debt and equity claims against future profits in order to receive cash. This cash is used to acquire fixed assets and raw materials, which, together with labor, are turned into finished goods and inventories. As the goods are sold, the inventories are replaced by accounts receivable (from credit sales) or cash (from cash sales). As customers pay their bills, the accounts receivable are converted into cash and so on. In the interim, the firms pay taxes and interest, amortize their debts, and pay dividends. The remaining cash is reinvested in the firm as retained earnings. A rapidly growing firm may find that its cash requirements have outstripped the firm's ability to generate cash internally, thus leading it to issue additional claims against future income. In a going concern, there is no start or end point. Cash is constantly flowing in and out of cash reservoirs, and the components of working capital (current assets minus current liabilities) are continually changing.

One way to solve the simulation model in section 1.2 is to write out the equations for the various balance sheet items explicitly and to solve them as a system of simultaneous linear equations. This is the approach taken by Warren and Shelton (1971).[13] Another way to solve the model is to use recursion. Each unknown in a recursion model is written as a function of other unknowns, and the unknowns are substituted one into other. At any stage, the current value of an unknown depends on the previous values of the other unknowns on which it is dependent. As an example of recursion, let us look at the firms' share capital (contributed capital). Share capital in any given year is the sum of share capital in the previous year and new issues. New issues can be defined in terms of other balance sheet items (retained earnings for the current year, long-term liabilities for the current year, and so on). However, these items, in turn, can be functions of balance sheet and income statement items. Ultimately, our model involves extensive circularity of argument: share capital depends on new issues, which depends on retained earnings, which depends on net income, which depends on financial income, which depends on long-term liabilities, which depends on share capital. This is only one example of the kinds of circularities involved.

The approach we use to identify the recursion follows the traditional approach in economic theory. We use the solution (total differentiation of the first order condition) of a dynamic optimization model to find out the economic relationships between the changes of different balance sheet items. These relationships are then estimated using different estimation methods. Different items in the income statement are related to different balance sheet items. This implies that the income statement items are estimated after estimating the changes of different balance sheet items. Finally, the estimated relationships are inserted into the difference equations system (described in section 1.2), which is then solved numerically to be able to simulate the future values of the stock variables in the firms' balance sheets.

The identified recursive method (from the dynamic optimization problem) used in the statistical module follows the structure in Figure 4: first, we decide the economic depreciation of machinery and equipment, the sale of machinery and equipment, and the investment in these assets, (sections 7.1-7.3). In sections 7.4-7.5, we estimate the economic depreciation of buildings, and the net investment in these assets. The net changes in other fixed assets and current assets are estimated in sections 7.6-7.7. Second, having established the net changes in different assets, we proceed to investigate the funds available to undertake such investment. In sections 7.8-7.11, we estimate the net change in long-term liabilities, the net change in current liabilities, the net change in share capital, and the net change in restricted reserves. Finally, in sections 7.12-7.24, we estimate operating income before depreciation, financial income, financial expenses, the change in other untaxed reserves, net group contributions, other allocations, tax liabilities, other tax adjustments, the tax depreciation of buildings, and the reduction of taxes.

[13] For different approaches to simulate firms' financial planning, the interested reader is referred to Gentry and Pyhrr (1973), Downes (1973), Weston (1974), Warren (1974), Lyneis (1975), Francis and Rowell (1978), Gentry (1979), Francis (1983), Kumar and Vrat (1989), Benninga (1992), Clarke and Tobias (1995), and Tongeren (1995).

2 The Simulation Module

When looking at the economic development of firms, the most important source of data is contained in the firms' three basic financial statements: the balance sheet, the income statement, and the statement of changes in financial conditions (the cash flow statement). In what follows, we will formalize these three financial statements in order to develop a model for describing the firms' economic situation (section 2.1-2.3). Section 2.4 contains the dynamic characteristic of the balance sheet. Further, section 2.5 defines the constraints on the firms' financial decisions. Finally, in section 2.6, we conclude the dynamic characteristic of these statements and the interrelationship between them.

2.1 The Balance Sheet

The balance sheet presents firms' assets, liabilities, and equity, at the moment when the books are closed. Assets represent firms' investments. Liabilities and equity indicate how these investments are financed.[14]

The asset side of the balance sheet shows the assets owned by the firms. The assets are divided into two major components: current assets and fixed assets. Current assets are assets with a maturity of less than one year (which means that they are likely to be converted into cash within one year). The liability and equity side of the balance sheet shows what the firms owe their lenders, business partners, the tax authorities and their shareholders. This side is divided into four major components: current liabilities, long-term liabilities, untaxed reserves, and shareholders' equity (contributed capital). Current liabilities are liabilities that must be paid within one year. Long-term liabilities are borrowings mainly used to expand the firms' income-producing base.[15] The shareholders' equity represents the total common shareholders' ownership in a firm at the end of the year. Moreover, in Sweden financial reports are strongly linked to tax reports, and allocations made only for tax purposes should be reported as untaxed reserves in the balance sheet.

Let us now set up the closing balance sheet for firms at the end of year *t* as is shown below. The current assets are kept in cash or bank accounts for normal business use, or are invested in short-term notes. They also include accounts receivable, inventories[16], shares and other participations, bonds and other securities, notes receivable, claims for income tax refunds, prepaid expenses and accrued income, other current receivables and advances to suppliers for items to be rendered in the future.

Fixed assets are assets that firms intend to hold for more than one year, such as shares in other companies, machinery, equipment, intangible assets, etc. These assets are booked at the original cost less depreciation and revaluation. Fixed assets are defined as below

(2.1) $$FA_t = MA_t + BU_t + OFA_t$$

[14] Asset and liability accounts do not represent current values. Because accountants have adopted historical cost as the basis for valuation, balance sheet figures reflect the value in effect at the time the asset and/or the liability was acquired. Inflation compounds the problem associated with using accounting principles based on historical cost by increasing the discrepancy between historical and current values. Because the value of the firms' assets measured at replacement cost is higher than their book value, the annual change for depreciation and the cost of goods sold must also be higher when using replacement cost accounting. Thus, excluding other considerations, earnings measured on a replacement cost basis turn out to be smaller than officially reported. However, one should also remember that the replacement cost is not the same as the market value.

[15] Long-term liabilities are borrowings payable after one year.

[16] The value of these inventories is based on the FIFO (First-In and First-Out) accounting method, which reflects the most recent value of goods.

where MA_t consists of machinery and equipment (which also includes goodwill and other intangible assets), BU_t includes buildings (and fixed assets other than machinery and equipment), and OFA_t includes shares and other participations in domestic and foreign companies, bonds and other securities, and other fixed assets[17].

The closing balance sheet for incorporated firms in period *t*

Assets		**Liabilities**	
Fixed assets	FA_t	*Equity (contributed) capital*	EC_t
Current assets	CA_t	*Untaxed reserves*	UR_t
		Long-term liabilities	LL_t
		Current liabilities	CL_t

We define the accounting value of assets as

$$K_t = CA_t + FA_t \tag{2.2}$$

Firms' liabilities include their current liabilities, their long-term liabilities, their untaxed reserves and their equity capital, thus

$$B_t = CL_t + LL_t + UR_t + EC_t \tag{2.3}$$

Current liabilities are liabilities payable within one year to suppliers for goods and services provided, and for services and products paid by customers that the firm will provide in the near future.[18] Long-term liabilities include long-term debt and allocations for pensions. Untaxed reserves is defined as

$$UR_t = ASD_t + \hat{PF}_t + OUR_t \tag{2.4}$$

where ASD_t contains the accumulated supplementary depreciation (in excess of plan) and OUR_t contains other untaxed reserves.[19] One of the main untaxed reserves is the periodical reserve. Corporations are allowed to have six different periodical reserves. However, each periodical reserve must be brought back as an income at the latest six years after the

[17] This includes loans to partners or related persons, other long term receivables, advances to suppliers, deferred charges and other non-depreciable assets.

[18] More precisely, current liabilities include accounts payable, accrued expenses and prepaid revenues, advance payments from customers, and other current liabilities. It is worth mentioning that the payments of different taxes during a year do not need to coincide with the actual taxes firms must pay. Some part of the payment can be made during the subsequent year. In such cases, other current liabilities include the tax paid in the subsequent year.

[19] Which includes national investment reserves, the compensation fund, the tax equalization reserve (SURV), the foreign exchange reserve, and the inventory reserve.

allocation was made. This forces us to monitor seven different reserves at the same time in our theoretical model. In the balance sheet, we define the remaining periodical reserves from *t-5, t-4, t-3, t-2* and *t-1* as PF_t^{t-5}, PF_t^{t-4}, PF_t^{t-3}, PF_t^{t-2},and PF_t^{t-1}. Further, we define the periodical reserve in the current period as PF_t^t. This means that

$$\text{(2.5)} \qquad \hat{PF}_t = PF_t^{t-5} + PF_t^{t-4} + PF_t^{t-3} + PF_t^{t-2} + PF_t^{t-1} + PF_t$$

Firms' equity is defined as follows

$$\text{(2.6)} \qquad EC_t = SC_t + RR_t + URE_t$$

where share capital (SC_t) and restricted reserves (RR_t) are different parts of restricted equity, while URE_t characterizes unrestricted equity.

In the balance sheet the value of the firms' assets must be equal to that of their liabilities, so that

$$\text{(2.7)} \qquad K_t = B_t$$

Inserting equations (2.1)-(2.6) into equation (2.7) gives us the following condition

$$\text{(2.8)} \qquad \begin{aligned} &CA_t + MA_t + BU_t + OFA_t = SC_t + RR_t + URE_t + ASD_t + PF_t^{t-5} + PF_t^{t-4} \\ &+ PF_t^{t-3} + PF_t^{t-2} + PF_t^{t-1} + PF_t^t + OUR_t + LL_t + CL_t \end{aligned}$$

2.2 The Income Statement

The income statement depicts the change in shareholders wealth during a period of time. It is important to distinguish between reported profit and cash flow.

Gross income or the operating income before depreciation ($OIBD_t$**)** equals the difference between the period's revenues and the expenses the firm incurred to generate those revenues. Operating revenues include (net) sales and other operating revenues, while operating expenses include the cost of raw materials and consumables, other external charges, and staff costs.[20]

Operating income after economic depreciation can be calculated as follows

$$\text{(2.9)} \qquad OIAD_t = OIBD_t - EDEP_t^{MA} - EDEP_t^{BU}$$

[20] More precisely, operating expenses include the *cost of sales* directly related to operating levels (wages, salaries, raw materials, energy, transportation, pensions, supplies, services, etc.), *taxes* (other than income taxes) that are not closely related to operating levels (social security, fees, allowances for expenses, real state taxes, etc.), and *selling, research, general and administrative expenses* (pensions and other general expenses). Further, operating income before depreciation also includes variations in stocks of finished goods and work in progress.

where $EDEP_t^{MA}$ is an estimate of the actual economic depreciation of machinery and equipment and $EDEP_t^{BU}$ is the economic depreciation of buildings. By adding financial income and subtracting financial expenses from operating income before depreciation, we acquire earnings before allocations and taxes

(2.10) $$EBA_t = OIAD_t + FI_t - FE_t$$

Moreover, by adding net allocations[21] to earnings before allocations we obtain earnings before taxes as follows

(2.11) $$EBT_t = EBA_t - \Delta ASD - \Delta PF_t - OA_t$$

The change in untaxed reserves is the result of three different changes: allocations to accumulated supplementary depreciation ($\Delta ASD = ASD_t - ASD_{t-1}$), allocations to the periodical reserve ($\Delta PF_t = \hat{PF}_t - \hat{PF}_{t-1}$), and other allocations (OA_t)[22].

The first allocation in our model is supplementary depreciation [in excess of plan] which is the difference between depreciation for income tax purposes and depreciation according to plan. In our case the supplementary depreciation in excess of plan is[23]

(2.12) $$\Delta ASD = TDEP_t^{MA} - EDEP_t^{MA}$$

The second allocation/reversal is the allocation/reversal to/from the periodical reserves. According to the tax legislation regarding periodical reserves, firms may allocate a maximum of 25 percent (in 1998 the rate of allocation was 20 percent) of their taxable income each year to a special reserve. This reserve appears as an entry in the balance sheet under untaxed reserves. Firms are allowed to have six different reserves. Hence, the allocated amount in a reserve, *at the latest* six years after the fiscal year when the allocation was made, recurs as taxable income. However, firms have the opportunity to choose, within these six years, when to bring the allocated amount back to taxation. Furthermore, since taxable income is the base for allocations to a periodical reserve, the firm can in principal reallocate 25 percent of the old reserves to a new reserve. Thus, in order to model the periodical reserves correctly, we must monitor each of the six reserves separately.

Let us begin by defining $\hat{PF}_{t-1}$

(2.13) $$\hat{PF}_{t-1} = PF_{t-1}^{t-5} + PF_{t-1}^{t-4} + PF_{t-1}^{t-3} + PF_{t-1}^{t-2} + PF_{t-1}^{t-1} + PF_{t-1}$$

[21] Kanniainen & Södersten (1995) pointed out that with a uniform reporting convention the tax balance sheet of firms must coincide with that drawn up for their shareholders. A description of the two different accounting regimes that govern reporting practice in most developed countries and the tax effects of these reporting conventions can be found in Cummins, Harris & Hasset (1994) and Kanniainen & Södersten (1995).

[22] Other allocations include (net) group contributions (GC_t).

[23] The tax code only specifies the maximum amount of tax depreciation which firms may deduct from their taxable income. See section 2.5 for a detailed explanation of the supplementary depreciation in excess of plan.

Using equation (2.5) and (2.13), we can derive the allocations/reversals to/from periodical reserves (ΔPF_t) as below

(2.14) $\Delta PF_t = p_t^{allo} - zpf_t^{t-5} - zpf_t^{t-4} - zpf_t^{t-3} - zpf_t^{t-2} - zpf_t^{t-1} - PF_{t-1}^{t-5}$

where $-zpf_t^{t-5} = PF_t^{t-5} - PF_{t-1}^{t-4}$, $-zpf_t^{t-4} = PF_t^{t-4} - PF_{t-1}^{t-3}$, $-zpf_t^{t-3} = PF_t^{t-3} - PF_{t-1}^{t-2}$, $-zpf_t^{t-2} = PF_t^{t-2} - PF_{t-1}^{t-1}$, $-zpf_t^{t-1} = PF_t^{t-1} - PF_{t-1}$, and $p_t^{allo} = PF_t$.[24]

Finally, the last allocation/reversal is denoted other allocations OA_t .

Using the definitions above and equations (2.12) and (2.14), we can rewrite equation (2.11) as

(2.15) $EBT_t = EBA_t - TDEP_t^{MA} + EDEP_t^{MA} - p_t^{allo} + zpf_t + OA_t$

where $zpf_t = zpf_t^{t-5} + zpf_t^{t-4} + zpf_t^{t-3} + zpf_t^{t-2} + zpf_t^{t-1} + PF_{t-1}^{t-5}$. Net income (NI_t) can thus be derived as

(2.16) $NI_t = EBT_t - TL_t$

where TL_t is the tax liability. Moreover, EBT_t can be misleading because firms may enter costs in their accounts that are not deductible for tax purposes. Another example is earnings that firms do not book in their accounts even though these earnings are classified as taxable income according to the tax law.[25] By adjusting EBT_t, we are able to derive the firms' tax payments as

(2.17) $TAX_t = \tau \max[0, (EBT_t - TL_t + TA_t)]$

where τ is the corporate tax rate, TA_t (which can be positive, negative or equal to zero) is the firms' tax adjustments. It is worth mentioning that firms pay tax on their income if and only if $EBT_t - TL_t + TA_t > 0$. Tax adjustments are derived as below

(2.18) $TA_t = OTA_t - TDEP_t^{BU} - OL_{t-1}$

where OTA_t is other tax adjustments and includes: the taxable revenues which are not booked in the income statement, deductible expenses which are not accounted for in the income statement, accounted revenues which are not taxable,[26] accounted expenses which are not

---

[24] See section 2.6 for a detailed explanation of the allocations to a periodical reserve fund.

[25] Since 1998, firms are allowed to book group contributions either as other allocations or as an adjustment of their income for tax purposes.

[26] Such revenues are profits that firms receive because of composition arrangements with creditors, stockholders (shareholders) contributions, profits on the Swedish lottery or premium (price, lottery) bonds, and so on.

deductible,[27] the tax adjustment because of firms' sales of shares, the tax adjustment because of firms' sales of buildings, and the tax adjustments if firms are partners in a partnership[28]. Further, firms that sell (dispose of) forest products are allowed to make deductions for such sales. The deduction is 50 percent of the calculated revenue. Finally, other tax adjustments include deductions for depletion.[29] Tax adjustments are also made for the tax depreciation of buildings ($TDEP_t^{BU}$) and losses from previous years[30] (OL_{t-1} - which are fully deductible for firms).

However, to be able to derive firms' final tax payments ($FTAX_t$), we adjust firms' tax payments for their reduction of taxes (ROT_t) so that

$$FTAX_t = TAX_t - ROT_t \tag{2.19}$$

We can thus write net business income, when $EBT_t - TL_t + TA_t \geq 0$, as[31]

$$NBI_t = EBT_t - FTAX_t \tag{2.20}$$

However, when $EBT_t - TL_t + TA_t < 0$, firms increase their stocks of old losses by the same amount

$$OL_t = \min[0,(EBT_t - TL_t + TA_t)] \tag{2.21}$$

Net business income, NBI_t, increases unrestricted equity. However, unrestricted equity also decreases because of cash dividend payments during period *t*, which are attributable to net business income in the previous period (DIV_{t-1}). Further, unrestricted equity changes due to allocations/removals to/from restricted reserves (ΔRR_t). Moreover, unrestricted equity decreases also because of the maximum amount available for dividends in the current period (the so-called net cash flow, $cashfl_t$). It is important to keep in mind that income determines tax payments, whereas cash is what the firm can use to service debt, fund capital expenditures, and pay dividends. Thus, unrestricted equity in period t can be derived from

$$URE_t = URE_{t-1} + NBI_t - DIV_{t-1} - \Delta RR_t - cashfl_t \tag{2.22}$$

The dividend policy is the scheme for allocation of profits between shareholders and reinvestment in firms. As such, it is an important component of the firms' long-run financing strategies. Earnings that are retained in the firms can be used to fund additional investment or

[27] For instance *the Swedish general taxes*, costs for entertainment that exceed SEK 90/person, cooperation (association, union or society) fees, gifts and fines.
[28] The firms' earnings must be adjusted for the reported earnings from the partnership.
[29] Deductions for depletion can be made because of extraction of natural resources.
[30] This is not the case for losses on sales of shares held as a portfolio investment.
[31] τTA can be positive, negative or equal to zero. If $\tau TA > 0$ this can be interpreted as a tax reduction resulting from the tax adjustments, while $\tau TA < 0$ should be interpreted as a tax increase resulting from the tax adjustments. When $\tau TA = 0$, firms get neither a tax reduction nor a tax increase.

to decrease debt. Alternatively, firms that pay higher dividends have less internally generated cash available for investment purposes. The firms' boards of directors must propose dividend payments. Dividends are normally related to firms' earnings and set at a level that management believes is sustainable in the long run. There is a vast literature in this area. However, it seems like this decision is based on expected future earnings, anticipated investment opportunities, and the proportion of those opportunities to be financed by internal funds.

2.3 The Balance Sheet and its Dynamic Characteristic

The stock (state) variables of this model are LL_t, CL_t, SC_t, RR_t, URE_t, ASD_t, PF_t^{t-5}, PF_t^{t-4}, PF_t^{t-3}, PF_t^{t-2}, PF_t^{t-1}, PF_t^{t}, OUR_t, CA_t, MA_t, BU_t, and OFA_t. In what follows we will specify equations of motions that hold for these state variables.

The level of long-term liabilities at the end of time *t* equals the level of long-term liabilities at the end of time *t-1* increased by new long-term liabilities (dll_t^{if}) and decreased by repayments of loans (dll_t^{of}). Thus, a system dynamics model of firms' financial system might contain the following equation: $LL_t = LL_{t-1} + dll_t^{if} - dll_t^{of}$. By defining $dll_t = dll_t^{if} - dll_t^{of}$ as the net change in long-term liabilities, we can rewrite the equation of motion as

(2.23) $$LL_t = LL_{t-1} + dll_t$$

The level of current liabilities at the end of time *t* equals the level of current liabilities at the end of time *t-1 plus* the net change in current liabilities

(2.24) $$CL_t = CL_{t-1} + dcl_t$$

The level of share capital at the end of time *t* equals the level of share capital at the end of time *t-1*, plus the net change in share capital. The net change in share capital includes new share issues**,** stock dividend issues and share splits.[32]

(2.25) $$SC_t = SC_{t-1} + dsc_t$$

The level of restricted reserves at the end of time *t* equals the level of restricted reserves at the end of time *t-1*, plus the net change in restricted reserves.

[32] In addition to cash dividends, companies may pay stock dividends or split their stock. A *stock dividend issue* is the payment of additional shares of stock to common shareholders. For example, when a firm declares a 25 percent stock dividend issue, it means that for every four shares owned, a shareholder on the record date will receive an additional share. A *share split* is a proportionate increase in the number of common shares. For example, if a company splits its share three for one, it means that the shareholders receive three shares for each one held on the record date. Although there is no real financial difference between stock dividend issues and share splits- shareholders simply receive more paper- both the typical motives behind them and their accounting treatment differ. The technical distinction between the two is that a stock dividend issues appears as a transfer of retained earnings to the capital stock account, whereas a stock split is shown as a reduction in the value of each share.

(2.26) $RR_t = RR_{t-1} + drr_t$

The level of current assets at the end of time *t* equals the level of current assets at the end of time *t-1 plus* the net change in current assets

(2.27) $CA_t = CA_{t-1} + dca_t$

The level of firms' machinery and equipment at the end of time *t* equals the level of machinery and equipment at the end of time *t-1 plus* net investment (gross investment minus sales of old machinery and equipment) in new machinery and equipment *minus* economic depreciation of machinery and equipment

(2.28) $MA_t = MA_{t-1} + I_t^{MA} - S_t^{MA} - EDEP_t^{MA}$

The level of buildings at the end of time *t* equals the level of buildings at the end of time *t-1 plus* net investment in new buildings *minus* economic depreciation of buildings.

(2.29) $BU_t = BU_{t-1} + I_t^{BU} - EDEP_t^{BU}$

The level of other fixed assets at the end of time *t* equals the level of other fixed assets at the end of time *t-1 plus* net investment in new fixed assets

(2.30) $OFA_t = OFA_{t-1} + dofa_t$

The level of accumulated supplementary depreciation at the end of time *t* equals the level of accumulated supplementary depreciation at the end of time *t-1 plus* supplementary depreciation during period t (which is the difference between tax depreciation of machinery and equipment and economic depreciation).

(2.31) $ASD_t = ASD_{t-1} + (TDEP_t^{MA} - EDEP_t^{MA})$

The level of other untaxed reserves at the end of time *t* equals the level of other untaxed reserves at the end of time *t-1*, plus the net increase in other untaxed reserves

(2.32) $OUR_t = OUR_{t-1} + dour_t$

The levels of PF_t^{t-5}, PF_t^{t-4}, PF_t^{t-3}, PF_t^{t-2}, PF_t^{t-1}, and PF_t^{t} are defined as

(2.33) $PF_t^{t-5} = PF_{t-1}^{t-4} - zpf_t^{t-5}$

(2.34) $PF_t^{t-4} = PF_{t-1}^{t-3} - zpf_t^{t-4}$

(2.35) $PF_t^{t-3} = PF_{t-1}^{t-2} - zpf_t^{t-3}$

(2.36) $PF_t^{t-2} = PF_{t-1}^{t-1} - zpf_t^{t-2}$

(2.37) $PF_t^{t-1} = PF_{t-1} - zpf_t^{t-1}$

(2.38) $PF_t = p_t^{allo}$

2.4 The Statement of Changes in Financial Condition

The financial statement that is closest in reporting cash flow is formally referred to as the statement of changes in financial condition, more commonly known as the cash flow statement. This statement attempts to account for the sources of funds and all the uses of funds during an accounting period. A cash flow statement can either be based on the working capital concept of funds or define funds as cash[33]. However, it is possible to go from a "sources and uses of working capital statement" to a "sources and uses of cash statement". In our simulation model, we are not interested in the change in all the components of working capital. Thus, we have chosen to base our statement of changes in financial condition (the sources and uses of funds statement) on the working capital concept of funds.[34]

Inserting the difference equations for LL_t, CL_t, SC_t, RR_t, URE_t, ASD_t, PF_t^{t-5}, PF_t^{t-4}, PF_t^{t-3}, PF_t^{t-2}, PF_t^{t-1}, PF_t^{t}, OUR_t, CA_t, MA_t, BU_t, and OFA_t (from (2.22)-(2.38)) into equation (2.8) we obtain the cash flow constraint

(2.39)
$$\begin{aligned} cashfl_t = {} & OIBD_t + FI_t - FE_t + OA_t - FTAX_t - DIV_{t-1} + dsc_t + dcl_t + dll_t + dour_t \\ & - I_t^{MA} + S_t^{MA} - I_t^{BU} - dofa_t - dca_t \end{aligned}$$

This is the maximum amount available for dividends in the current period. In the next section, we will derive firms' dividend payments for the current period (DIV_t).

[33] The statement of sources and uses of cash attempts to explain the changes in the firms' cash between two points in time. It enables the analyst to answer three key questions: (1) How much cash did firms generate during the accounting period? (2) Where did the cash come from? (3) What did firms do with their cash?

[34] Working capital equals current assets minus current liabilities. Financial analysts often prepare *a sources and uses of working capital statement* as an indication of firms' liquidity: The more working capital firms can provide, the more liquid firms are considered to be. This statement is similar to a cash flow statement except that it records only the changes in working capital; it omits changes in individual current assets and current liabilities. In order to prepare a sources and uses of working capital statement, we must (1) Determine the change in each non-working capital item on the balance sheet during the period and the effect that change has on the cash position. (2) Examine the income statement to identify the effects on cash of operations, as well as the division of income between retained earnings and dividends. These items are then consolidated on the sources and uses statement. Sources of working capital include increases in non-current liabilities (via for instance new borrowing), decreases in non-current assets (such as property, plant, and equipment), and increases in net worth because of additions to retained earnings or the sale of additional shares of stock. Uses of working capital include decreases in non-current liabilities (such as the repayment of a term loan), increases in non-current assets (including investment in fixed assets), decreases in net worth by paying dividends and/or having an operating loss.

2.5 The Constraints on Firms' Dividend Payments Decision

In what follows we will specify various constraints on firms' dividend payments decision.

1) "Dividends" cannot exceed unrestricted equity in period t-1 (URE_{t-1}) plus the current period's net business income minus allocations to restricted reserves. If dividends exceed this amount the equity base of the firm would fall. The conditions can be summarized mathematically as

(2.40) $$cashfl_t \leq mcash_t$$

where

(2.41) $$mcash_t = URE_{t-1} + NBI_t - drr_t$$

This constraint is of course only valid for $URE_{t-1} > 0$.

2) Firms are not allowed to pay negative "dividends": Negative dividends would mean subsidies from government in the form of tax refunds

(2.42) $$cashfl_t \geq 0$$

Combining (2.39) and (2.42), we end up with the following expressions for the firms' dividend payments

(2.43) $$DIV_t = \max[0, \min(cashfl_t, mcash_t)]$$

2.6 The Constraints on Firms' Tax Depreciation of Machinery and Equipment

The rules for tax depreciation of machinery and equipment can be summarized as follows. For machinery and equipment, there are three kinds of depreciation rules available
1a. The main rule for depreciation according to Swedish tax legislation: *the declining balance method* (at 30 percent).
1b. The supplementary rule (to 1a) for depreciation according to Swedish tax legislation: *the straight line method* (at 20 percent).
2. *The rest value method* (at 25 percent).

To be able to explain the implication of the three different methods for depreciation, we need to define the taxable residual value of machinery and equipment and also the equation of motion for such assets.

The accounting database does not contain the taxable residual value of machinery and equipment. However, the taxable residual value of these assets can be calculated by the following equation

(2.44) $CMA_t = MA_t - ASD_t$

Moreover, we know that

(2.45) $MA_t = MA_{t-1} + I_t^{MA} - S_t^{MA} - EDEP_t^{MA}$

(2.46) $ASD_t = ASD_{t-1} + (TDEP_t^{MA} - EDEP_t^{MA})$

where S_t^{MA} is the sale of machinery and equipment (it also includes the revaluation of machinery and equipment during the accounting year). Using equations (2.45) and (2.46) and (2.26) we can rewrite equation (2.44) as

(2.47) $CMA_t = MA_{t-1} - ASD_{t-1} + I_t^{MA} - S_t^{MA} - TDEP_t^{MA}$

However, we know that $CMA_{t-1} = MA_{t-1} - ASD_{t-1}$ so that equation (2.47) can be rewritten to define the equation of motion for the taxable residual value of machinery and equipment as follows

(2.48) $CMA_t = CMA_{t-1} + I_t^{MA} - S_t^{MA} - TDEP_t^{MA}$

Equation (2.48) states that the level of the firms' taxable residual value of machinery and equipment at the end of period *t* (CMA_t) equals the level of the taxable residual value of machinery and equipment at the end of period *t-1* plus the firms' net investment (gross investment minus sales of these assets) during period *t* minus depreciation for income tax during that period.

The declining balance method: The tax code in Sweden specifies the maximum amount of tax depreciation that firms may deduct from their taxable income. The declining balance method allows a maximum deduction of 30 percent of the remaining taxable residual value and the gross investment made in period *t*. However, the taxable residual value must be adjusted for different investment grants and the sale price of machinery and equipment that has been disposed of during period *t*. Further, the deduction rate must also be adjusted if the income year is longer than 12 months. For example, if the income year is 18 months the depreciation rate must be multiplied by 18/12, which gives a depreciation rate of 45%. The declining balance method implies that the tax depreciation $TDEP_t^{MA}$ is constrained by

(2.49) $TDDB_t^{MA} = (M/12)\delta^{db}[CMA_{t-1} + I_t^{MA} - S_t^{MA} - IG_t]$

where *M* is the number of months in the firms' income year, IG_t is the investment grants, and $\delta^{db} = 0.3$ is the maximum rate of depreciation allowed for tax purposes. The allowed tax depreciation rate is assumed to be higher than the economic depreciation rate.

The straight line method: The calculation of tax depreciation according to the straight line method necessitates knowledge about firms' investment during the last three years before the

income year. This method allows a 20% depreciation of inventories and equipment. Further, the deduction rate must be adjusted if the income year is longer than 12 months.
The straight line method starts from the taxable residual value at period *t-1*. However, this value is reduced by 80% of the firms' investment in year *t-3*, 60% of the firms' investment in year *t-2*, 40% of the firms' investment in year *t-1*, and 20% of the firms' investment in year *t*.[35] The resulting difference is the maximum tax depreciation according to the straight line method. The straight line method implies that the tax depreciation $TDEP_t^{MA}$ is constrained by

$$\text{(2.50)} \quad \begin{aligned} TDSL_t^{MA} &= CMA_{t-1} + I_t^{MA} - S_t^{MA} - IG_t \\ &\quad - [\delta_t^s (I_t^{MA} - IG_t) + \delta_{t-1}^s I_{t-1}^{MA} + \delta_{t-2}^s I_{t-2}^{MA} + \delta_{t-3}^s I_{t-3}^{MA}] \end{aligned}$$

where $\delta_t^s = [1-(M/12)\delta^s]$, $\delta_{t-1}^s = \delta_t^s - \delta^s$, $\delta_{t-2}^s = \delta_{t-1}^s - \delta^s$, and $\delta_{t-3}^s = \delta_{t-2}^s - \delta^s$. δ^s is the allowed tax depreciation rate according to straight line method. For example, if we assume that $\delta^s = 0.2$ and $M = 12$ then $\delta_t^s = 0.8$, $\delta_{t-1}^s = 0.6$, $\delta_{t-2}^s = 0.4$, and $\delta_{t-3}^s = 0.2$. The firm chooses the depreciation rule that gives the highest tax depreciation. The choice can be defined mathematically by a maximum function: $\max(TDDB_t^{MA}, TDSL_t^{MA})$.

The rest value method: This method follows the same rules as the declining balance method. However, the rest value method is simpler. The method is mainly used by unincorporated businesses that do not prepare annual accounts. However, incorporated firms are also allowed to use this method. This will give them lower tax depreciation (compared to the declining balance method). The rest value method gives these firms the opportunity to make deductions for tax depreciation even if they do not have annual accounts. It is also important to note that these firms are not allowed to use the supplementary rule for depreciation according to Swedish tax legislation (the straight line method).

Mathematically, the rest value method will be handled just like the declining balance method as follows

$$\text{(2.51)} \quad TDRV_t^{MA} = (M/12)\delta^{rv}[CMA_{t-1} + I_t^{MA} + S_t^{MA} - IG_t]$$

where $\delta^{db} = 0.25$ is the maximum rate of depreciation allowed for tax purposes according to the rest value method.

Let us now define a variable that captures the constraints that the tax code in Sweden puts on the amount of tax depreciation that firms may deduct from their taxable income as $MTDM_t$.[36] The tax depreciation of machinery and equipment is constrained by $MTDM_t$ so that

[35] It is worth mentioning that investment in year *t-3*, *t-2*, *t-1*, and *t* must be decreased by the purchasing value of the part of this investment that has been disposed of during year *t*. However, our database does not contain such detailed information. Thus, we exclude these adjustments as a simplifying assumption. The implication of this assumption is that we will systematically underestimate the maximum allowed depreciation according to the straight line method.
[36] This variable will be used as an explanatory variable in our estimations. See further Chapter 6.

(2.52) $TDEP_t^{MA} \leq MTDM_t$

However, as we mentioned above, $MTDM_t$ depends on following conditions:[37]

1) If $TDDB_t^{MA} > TDRV_t^{MA}$ and $TDSL_t^{MA} > TDRV_t^{MA}$ then

(2.53) $MTDM_t = \max(TDDB_t^{MA}, TDSL_t^{MA})$

2) If $TDDB_t^{MA} \leq TDRV_t^{MA}$ and $TDSL_t^{MA} \leq TDRV_t^{MA}$ then

(2.54) $MTDM_t = TDRV_t^{MA}$

2.7 The Constraint on Firms' Allocations to the Periodical Reserves

The tax code in Sweden specifies the maximum amount of allocations firms can allocate to periodical reserves each year. As mentioned above, firms can deduct up to 25 percent (the allocation was decreased to 20 percent in1998) of its taxable income each year (adjusted for different items). In the balance sheet this deduction is booked as a reserve (under untaxed reserves). The maximum base for the allocation to this reserve during the income year is calculated according to the following equation:

(2.55) $pbase_t = OIBD_t - EDEP_t^{BU} + FI_t - FE_t - TDEP_t^{MA} + zpf_t + OA_t - TL_t + TA_t$

The allocation to the periodical reserve fund is constrained to $\eta = 0.25\%$ of this base[38]

(2.56) $p_t^{allo} \leq MPA_t$

where

(2.57) $MPA_t = \max[0, (\eta \times pbase_t)]$

Thus, the constraint (2.57) can be rewritten as follows

[37] Another way of modeling this is: If $CMA_t^{DB} < CMA_t^{RV}$ and $CMA_t^{SL} < CMA_t^{RV}$ then $CMAM_t = \min(CMA_t^{DB}, CMA_t^{RV})$ and $MTDM_t = CMA_{t-1} + I_t^{MA} + S_t^{MA} - IG_t - CMAM_t$, where $CMA_t^{DB} = CMA_{t-1}^{DB} + I_t^{MA} - S_t^{MA} - TDDB_t^{MA}$, $CMA_t^{RV} = CMA_{t-1}^{RV} + I_t^{MA} - S_t^{MA} - TDRV_t^{MA}$, and CMA_t^{SL} are the taxable residual value of machinery and equipment for the declining balance method, the rest value method, and the straight line method.

[38] MPA_t will be used as an explanatory variable in our estimations. See further Chapter 6.

(2.58) $P_t^{allo} = \max[0, \min(P_t^{allo}, (\eta \times pbase_t))]$

2.8 The System Dynamics for the Firm

A system dynamics model consists of a set of equations defining various causal relationships between variables and/or integrations of variables that can be written as a set of first-order difference equations. A model's behavior is determined by recursively simulating the set of equations comprising the model. Alternative policies are evaluated by observing the impact each policy (in combination with the parameters (inputs) and structure (equations) of the model) has on the simulated behavior of model variables (outputs). From such policy experiments, inferences can be made concerning decision-making in the real system.

The three basic financial statements modeled in previous sections constitute the base module in our dynamic system. In sections 2.1-2.6, we formalized the dynamic characteristics of these statements and the interrelationship between them. This was done by specifying a set of first-order difference equations defining various causal relationships between different items on the balance sheet and income statement and integrations of some of these items.

By summarizing the findings in sections 2.1-2.6, we end up with the following system dynamic model for firms

(2.59) $MA_t = MA_{t-1} + I_t^{MA} - S_t^{MA} - EDEP_t^{MA}$

(2.60) $BU_t = BU_{t-1} + I_t^{BU} - EDEP_t^{BU}$

(2.61) $OFA_t = OFA_{t-1} + dofa_t$

(2.62) $CA_t = CA_{t-1} + dca_t$

(2.63) $SC_t = SC_{t-1} + dsc_t$

(2.64) $RR_t = RR_{t-1} + drr_t$

(2.65) $OUR_t = OUR_{t-1} + dour_t$

(2.66) $CMA_t = CMA_{t-1} + I_t^{MA} - S_t^{MA} - TDEP_t^{MA}$

(2.67) $ASD_t = ASD_{t-1} + [TDEP_t^{MA} - EDEP_t^{MA}]$

(2.68) $PF_t^{t-5} = PF_{t-1}^{t-4} - zpf_t^{t-5}$

(2.69) $PF_t^{t-4} = PF_{t-1}^{t-3} - zpf_t^{t-4}$

(2.70) $PF_t^{t-3} = PF_{t-1}^{t-2} - zpf_t^{t-3}$

(2.71) $PF_t^{t-2} = PF_{t-1}^{t-1} - zpf_t^{t-2}$

(2.72) $PF_t^{t-1} = PF_{t-1} - zpf_t^{t-1}$

(2.73) $PF_t = p_t^{allo}$

(2.74) $LL_t = LL_{t-1} + dll_t$

(2.75) $CL_t = CL_{t-1} + dcl_t$

(2.76) $EBT_t = OIBD_t - EDEP_t^{BU} + FI_t - FE_t - TDEP_t^{MA} - p_t^{allo} + zpf_t + OA_t$

(2.77) $TA_t = OTA_t - TDEP_t^{BU} - OL_{t-1}$

(2.78) $TAX_t = \tau \max[0, (EBT_t - TL_t + TA_t)]$

(2.79) $FTAX_t = TAX_t - ROT_t$

(2.80) $NBI_t = EBT_t - FTAX_t$

(2.81) $OL_t = abs\{\min[0, (EBT_t - TL_t + TA_t)]\}$

(2.82) $$cashfl_t = OIBD_t + FI_t - FE_t + OA_t - FTAX_t - DIV_{t-1} + dsc_t + dcl_t + dll_t + dour_t - I_t^{MA} + S_t^{MA} - I_t^{BU} - dofa_t - dca_t$$

(2.83) $URE_t = URE_{t-1} + NBI_t - DIV_{t-1} - drr_t - cashfl_t$

(2.84) $mcash_t = URE_{t-1} + NBI_t - drr_t$

(2.85) $DIV_t = \max[0, \min(cashfl_t, mcash_t)]$

where $P_t^{allo} = \max[0, \min(P_t^{allo}, (\eta \times pbase_t))]$.

In Chapter 6, we will describe the estimation methods used to capture firms' behavior regarding the following flow variables: $EDEP_t^{MA}$, S_t^{MA}, I_t^{MA}, $EDEP_t^{BU}$, I_t^{BU}, $dofa_t$, dca_t, dll_t, dcl_t, dsc_t, drr_t, $OIBD_t$, FI_t, FE_t, $TDEP_t^{MA}$, zpf_t, $dour_t$, GC_t, OA_t, TL_t, OTA_t, $TDEP_t^{BU}$, p_t^{allo}, and ROT_t.

3 The Dynamic Optimization Model

This chapter is organized as follows: In section 3.1, we set up an expression for the value of the firms and their cash flow constraint. We also present the necessary assumptions for the optimization model. Firms' optimization problem and the optimality conditions are shown in section 3.2. In section 3.3, we show how we derive the econometric model by differentiating through the optimality conditions.

3.1 The Value of the Firms and Their Cash Flow Constraint

The fundamental condition for equilibrium in the capital market is that the yield on investing in any particular asset will be the same as the yield on all other assets, taking into account taxes and transactions costs:[39]

$$(3.1)\quad (1-\tau_p)iV(u) = (1-\tau_d)\,cashfl(u) + (1-\tau_c)(\dot{V} - dsc)$$

where $V(u)$ is the value of the firms at time u, $\dot{V}$ is the derivative of the firms' value with respect to time, τ_d is the dividend tax, τ_c is the tax on accruing capital gains, i is the market interest rate, τ_p is the marginal rate of income tax on bond income. Hence, $(1-\tau_p)i$ is the shareholders' post-tax return on holding bonds and $(1-\tau_d)\,cashfl(u)$ is the net "dividend" payment paid at time u. Expression (3.1) represents the non-arbitrage condition. By integrating (3.1) we can then obtain the value of the firms at time t:

$$(3.2)\quad \mathrm{V}(t) = \int_{u=t}^{\infty} \left[\theta\, cashfl - dsc\right] e^{-r(u-t)} du$$

where $\theta = \dfrac{1-\tau_d}{1-\tau_c}$ and $r = \dfrac{1-\tau_p}{1-\tau_c} i$. Firms' cash flow, net of profit tax, is defined from the firms' budget constraint. Before we present the budget constraint, we need to make a few assumptions.

As we showed in Chapter 2, earnings before taxes is defined as follows

$$(3.3)\quad EBT = EBA - \dot{ASD} - \dot{PF} + OA$$

where

$$(3.4)\quad EBA = OIBD - EDEP^{MA} - EDEP^{BU} + FI - FE$$

$$(3.5)\quad \dot{PF} = p^{allo} - zPF$$

$$(3.6)\quad \dot{ASD} = TDEP^{MA} - EDEP^{MA}$$

Inserting (3.4)-(3.6) into (3.3) leads us to the following expression

[39] See King (1977) section 4.1, pp. 90-91, Dixon (1992) chapter 5, Poterba & Summers (1985), pp. 232-233, and Sinn (1987), pp. 63-65.

(3.7) $EBT = OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} - p^{allo} + zPF + OA$

where

(3.8) $OA = -dour + GC$

We define operating income before depreciation as

(3.9) $OIBD = f(MA, BU, WC, L) - WL$

where $f(MA, BU, WC, L)$ is the production function[40] with $f_{MA} > 0$, $f_{BU} > 0$, $f_{WC} > 0$, $f_L > 0$, $f_{MAMA} < 0$, $f_{BUBU} < 0$, $f_{WCWC} < 0$, and $f_{LL} < 0$. WL is the firms' cost for factor labor, and $WC = CA - CL$ is the working capital. The reason for letting the working capital of the firms affect their production is our belief that firms need physical, human and working capital to be able to produce different goods. The best way of understanding this (which we also noticed in section 1.3) is to take a closer look at what usually is known as the firms' cash flow cycle. Firms begin by issuing various debt and equity claims against future profits in order to receive cash. This cash is used to acquire fixed assets and raw materials, which, together with labor, are turned into finished goods and inventories. Firms can of course sell goods from inventories. In such a case, they either receive cash (from cash sales) or accounts receivable (from credit sales). As customers pay their bills, the accounts receivable are converted into cash and so on. In the interim, firms pay taxes, interest, amortize their debts, and pay dividends. Remaining cash is reinvested in the firms as retained earnings. Rapidly growing firms may find that their cash requirements outstrip their ability to generate cash internally, thus leading them to issue additional claims against future income. In a going concern, there is no start or end point. Cash is constantly flowing in and out of cash reservoirs, and the components of working capital (current assets minus current liabilities) are continually changing. Therefore, working capital is just as important as physical and human capital in the production process.

Let us now define financial income

(3.10) $FI = h(OFA + MA + BU + CA)$

As we mentioned in Chapter 1, financial income includes interest income, dividends received, income from fixed assets, profits on operations disposed, gains from the sale of assets, etc. We believe that equation (3.10) is a good approximation of the main parts of financial income. Financial expenses are defined as follows

(3.11) $FE = i_B(CL + LL) + n(MA + BU + CA + OFA)$

where i_B is debt interest[41], and $i_B(CL + LL)$ is the firms' interest costs on their debt. As we mentioned in Chapter 1, financial expenses include interest on all types of short- and long-

[40] The firms' production function is equal to the firms' revenue because the product price is assumed to be equal to one.

[41] The debt interest rate is equal to the market interest rate. It is nevertheless convenient to distinguish between these in notation, since in the derived formulae it is often that only one of the two is present. By separating the notation it becomes easier to identify the impact of one or the other.

term borrowing, the value adjustment of fixed assets, losses on operations disposed, losses from the sale of assets, etc. We believe that $n(MA + BU + CA + OFA)$ is a good approximation of these other parts of financial expenses. Economic depreciation of machinery and equipment and economic depreciation of buildings are defined as follows

$$(3.12) \quad EDEP^{MA} = \delta^{MA} MA$$
$$(3.13) \quad EDEP^{BU} = \delta^{BU} BU$$

where δ^{MA} is the economic depreciation rate of machinery and equipment and δ^{BU} is the economic depreciation rate of buildings. Let us now define group contributions. We assume that group contributions are given based on the firms' net income before group contributions ($NIBGC$) so that

$$(3.14) \quad GC = g\, NIBGC = g[OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} + zPF - dour]$$

where $-1 \leq g \leq 1$. (3.8) and (3.14) inserted into (3.7) gives us

$$(3.15) \quad EBT = (1 + g)[OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} + zPF - dour] - p^{allo}$$

From Chapter 2, we know that

$$(3.16) \quad NI = EBT - TL$$
$$(3.17) \quad NBI = EBT - FTAX$$

where

$$(3.18) \quad FTAX = TAX - ROT$$

where

$$(3.19) \quad TAX = \tau(NI + TA)$$

where

$$(3.20) \quad TA = OTA - TDEP^{BU}$$

From Chapter 1, we know that tax liabilities (TL) are firms' approximation of the tax payment, which of course is based on earnings before taxes.

$$(3.21) \quad TL = \alpha\, EBT$$

where $0 \leq \alpha \leq 1$. From Chapter 1, we also know that other tax adjustments are based on firms' net income so that

$$(3.22) \quad OTA = \gamma^{OT} NI$$

We also know that reductions of taxes are based on firms' tax calculation so that

$$(3.23) \quad ROT = \beta^{RO} TAX$$

Finally, we assume that the fiscal depreciation of buildings is also a fraction of the building stock as follows

$$(3.24) \quad TDEP^{BU} = \delta_T^{BU} BU$$

where δ_T^{BU} is the fiscal depreciation rate of buildings.

Equation (3.21) inserted into equation (3.16) gives the following expression for net income

$$(3.25) \quad NI = (1-\alpha)EBT$$

Inserting equations (3.22) and (3.24) into (3.20) gives us the following expression for tax adjustments

$$(3.26) \quad TA = \gamma^{OT}(1-\alpha)EBT - \delta_T^{BU} BU$$

By inserting equations (3.25) and (3.26) into (3.19) we are able to rewrite the equation for the firms' tax payments as

$$(3.27) \quad TAX = \tau[(1+\gamma^{OT})(1-\alpha)EBT - \delta_T^{BU} BU]$$

By inserting equations (3.23) and (3.27) into (3.18) we are able to rewrite the equation for the firms' final tax payments as

$$(3.28) \quad FTAX = (1-\beta^{RO})\tau[(1+\gamma^{OT})(1-\alpha)EBT - \delta_T^{BU} BU]$$

Inserting equations (3.28) and (3.15) into equation (3.17) gives us the following equation for net business income

$$(3.29) \quad \begin{aligned} NBI = {} & (1-\xi\tau)\{(1+g)[OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} + zPF - dour] - p^{allo}\} \\ & + (1-\beta^{RO})\tau\delta_T^{BU} BU \end{aligned}$$

where

$$(3.30) \quad \xi = (1-\beta^{RO})(1+\gamma^{OT})(1-\alpha)$$

Thus, the change in other restricted equity is

$$(3.31) \quad U\dot{R}E = NBI - cashfl - drr$$

Further, from Chapter 2, we know that

$$(3.32) \quad \dot{MA} = I^{MA} - EDEP^{MA} - SMA$$

(3.33) $\dot{BU} = I^{BU} - EDEP^{BU}$

(3.34) $\dot{OFA} = dofa$

(3.35) $\dot{CL} = dcl$

(3.36) $\dot{CA} = dca$

(3.37) $\dot{SC} = dsc$

(3.38) $\dot{RR} = drr$

(3.39) $\dot{LL} = dll$

(3.40) $\dot{OUR} = dour$

(3.41) $\dot{MA} + \dot{BU} + \dot{OFA} + \dot{CA} = \dot{SC} + \dot{RR} + \dot{URE} + \dot{ASD} + \dot{PF} + \dot{OUR} + \dot{LL} + \dot{CL}$

where we assume that firms' sales of machinery and equipment are a fraction of the machinery and equipment stock

(3.42) $SMA = s\,MA$

Inserting equations (3.5)-(3.6), (3.31)-(3.40), and (3.42) into equation (3.41) gives us the budget constraint

$$\begin{aligned}
cashfl &= (1+g)(1-\xi\tau)[OIBD + FI - FE] \\
&\quad + [1-(1+g)(1-\xi\tau)][EDEP^{BU} + TDEP^{MA} - zPF + dour] \\
&\quad + \xi\,\tau\,p^{allo} + (1-\beta^{RO})\tau\delta_T^{BU}BU \\
&\quad - I^{MA} + s\,MA - I^{BU} - dofa - dca + dsc + dll + dcl
\end{aligned} \tag{3.43}$$

Corporate taxation is given by τ, ξ, g, and δ_T^{BU}; the corporate income tax rate, the rate of reduction/increase [which is dependent on the rate for other tax adjustment (γ^{OT}), the rate of reduction of taxes (β^{RO}), and the rate of tax liabilities (α)], the rate of group contributions, and the rate of fiscal depreciation of buildings, respectively. $(1+g)\xi\tau[EDEP^{BU} + TDEP^{MA} - zPF + dour]$ is the tax reduction/increase resulting from accelerated depreciation of machinery and equipment, economic depreciation of buildings, reversals from periodical reserves, and changes in other untaxed reserves. Further, $-\xi\,\tau\,p^{allo} + (1-\beta^{RO})\tau\delta_T^{BU}BU$ is the tax reduction resulting from allocations to periodical reserves and the tax reduction resulting from fiscal depreciation of buildings.

3.2 The Constraints on the Firms' Financial Decision

Various institutional constraints must be taken into consideration when modeling firms' decision-making processes. In Chapter 2 and Chapter 4, we discussed different financial and market-based constraints. Let us now include these constraints in firms' optimization model.

1) The dividends cannot exceed the accounting profit.[42] If dividends exceed accounting profits, then the equity base of the firm would fall. Södersten & Kanniainen (1995) begin by pointing out that with a uniform reporting convention the tax balance sheet of the firm must coincide with that drawn up for the shareholders. Combining this fact with the condition that the dividends paid must not exceed post-tax book profits leads to the following condition:

$$(3.44) \qquad cashfl \leq NBI + URE - drr$$

If we substitute the expression for $cashfl$ from (3.43) into (3.44) we obtain the following

$$(3.45) \qquad \begin{aligned} dsc + dll + dcl \leq\ & I^{MA} + I^{BU} + dofa + dca + URE - EDEP^{BU} - drr \\ & + zPF - p^{allo} - dour - SMA - TDEP^{MA} \end{aligned}$$

This inequality is a constraint on the use of external funds showing that the part of investment in different assets that is not covered by tax depreciation of machinery and equipment, economic depreciation of buildings, sale of machinery and equipment, allocations to restricted reserves, allocations to other untaxed reserves, through the current allocation/removal to/from periodical reserve, and the stock of unrestricted equity, can at most be financed by issuing new long-term liabilities, new current liabilities, or changes in share capital.

2) Firms are not allowed to pay negative "dividends": Negative dividends would mean subsidies from government in the form of tax refunds.

$$(3.46) \qquad cashfl \geq 0$$

If we substitute the expression for $cashfl$ from (3.43) into (3.46) we obtain the following

$$(3.47) \qquad \begin{aligned} dsc + dll + dcl \geq\ & I^{MA} - SMA + I^{BU} + dofa + dca \\ & - EDEP^{BU} - TDEP^{MA} + zPF - dour - p^{allo} - NBI \end{aligned}$$

where NBI is the retainable net profit. Inequality (3.47) states that the sum of long-term liabilities, current liabilities and changes in share capital must at least be large enough to finance the part of investment in different assets that cannot be financed through accelerated depreciation of machinery and equipment, economic depreciation of buildings, allocations/removal to/from periodical reserves, allocations to other untaxed reserves, and retention.

In addition to these constraints, there are also some constraints in the Swedish tax legislation that may influence firms' investment behavior and financial behavior.

3) *The tax code in Sweden specifies the maximum amount of tax depreciation of machinery and equipment.* In section 2.2.1, we defined a variable that captures the constraint that the tax code in Sweden puts on the amount of tax depreciation of machinery and equipment that firms may deduct from their taxable income as $MTDM$. $MTDM$ encompasses both the declining balance method, the straight line method, and the rest value method. However, for the sake of simplicity, we let $MTDM$ include only the declining balance method in the

[42] This constraint makes sure that the equity base of firms is kept intact.

optimization model. The tax depreciation of machinery and equipment is constrained by $MTDM$, so that

$$(3.48) \quad TDEP^{MA} \leq MTDM$$

where

$$(3.49) \quad MTDM = \delta_T^{MA} CMA$$

where

$$(3.50) \quad CMA = MA - ASD$$

$$(3.51) \quad \dot{CMA} = I^{MA} - TDEP^{MA} - SMA$$

The tax depreciation rate is assumed to be higher than the economic depreciation rate ($\delta_T^{MA} > \delta^{MA}$). $\tau\delta_T^{MA} CMA$ is the maximum tax reduction resulting from accelerated depreciation.

4) The tax code in Sweden specifies the maximum amount of allocations firms can allocate each year to periodical reserves. In section 2.2.2, we showed that the allocations to the periodical reserves are constrained to $\eta = 0.25$ of the base MPA so that

$$(3.52) \quad p^{allo} \leq MPA$$

where

$$(3.53) \quad \begin{aligned} MPA = \eta[&OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} + zPF \\ &- dour + GC - TL + TA] \end{aligned}$$

Using equations (3.14), (3.20), and (3.21) we can rewrite MPA as follows

$$(3.54) \quad \begin{aligned} MPA = \eta^* \{&(1+g)(1+\gamma^{OT})(1-\alpha)[OIBD - EDEP^{BU} + FI - FE - TDEP^{MA} + zPF \\ &- dour] - TDEP^{BU}\} \end{aligned}$$

where

$$(3.55) \quad \eta^* = \frac{\eta}{1 + \eta[\gamma^{OT} - \alpha(1+\gamma^{OT})]}$$

The periodical reserves work as follows. Each year firms can deduct up to 25 percent of their taxable income by allocating an equivalent amount to a special fund, which appears as an entry in the balance sheet. This means that firms' tax payments are reduced by an amount equal to the allocation multiplied by the statutory tax rate. The allocated amount must be included with taxable income, no later than six years after the fiscal year when the allocation was made.

5) Firms' share capital must be higher or equal to $\overline{SC} = 100000\ SEK$ so that

(3.56) $SC \geq \overline{SC}$

6) We also impose non-negativity constraints on each and every balance sheet item

(3.57) $MA \geq 0$
(3.58) $CMA \geq 0$
(3.59) $BU \geq 0$
(3.60) $OFA \geq 0$
(3.61) $CA \geq 0$
(3.62) $RR \geq 0$
(3.63) $ASD \geq 0$
(3.64) $PF \geq 0$
(3.65) $OUR \geq 0$
(3.66) $LL \geq 0$
(3.67) $CL \geq 0$

3.3 The Firms' Objective and the Optimality Conditions

We assume that the interest rate and the wage rate are exogenously given. Firms will choose the time path of employment (L), net increase in current liabilities (dcl), net increase in long-term liabilities (dll), net change in share capital (dsc), net change in restricted reserves (drr), net change in other untaxed reserves ($dour$), allocations/removals to/from periodical reserves (p^{allo} / zPF), tax depreciation of machinery and equipment ($TDEP^{MA}$), investment in machinery and equipment (I^{MA}), net investment in buildings (I^{BU}), net investment in other fixed assets ($dofa$), and net change of current assets (dca) so that the market value of their shares is maximized. The state variables of the optimization problem are the stock of machinery and equipment (MA), the stock of buildings (BU), the stock of current assets (CA), the stock of other fixed assets (OFA), the stock of current liabilities (CL), the stock of long-term liabilities (LL), the stock of share capital (SC), the stock of restricted reserves (RR), periodical reserves (PF), accumulated supplementary depreciation (ASD), and the stock of other untaxed reserves (OUR). The equations of motion for these state variables were specified in equations (3.5), (3.6), and (3.32)-(3.40).

The firms' objective is then to maximize (3.2) subject to (3.29), equations (3.5), (3.6), (3.32)-(3.40), (3.45), (3.47), (3.48), (3.52), (3.56)-(3.67), and the initial conditions $MA = MA_0$, $BU = BU_0$, $CA = CA_0$, $OFA = OFA_0$, $CL = CL_0$, $LL = LL_0$, $SC = SC_0$, $RR = RR_0$, $PF = PF_0$, $ASD = ASD_0$, and $OUR = OUR_0$. The current-value Hamiltonian for this problem is:

$$
\begin{aligned}
F &= \theta\{(1-\tau^{e})[OIBD+FI-FE]+\tau^{e}[\delta^{BU}BU+TDEP^{MA}-zPF+dour]+\tau * p^{allo} \\
&+\tau^{\beta}\delta_{T}^{BU}BU-I^{MA}+s\,MA-I^{BU}-dofa-dca+dsc+dll+dcl\}-dsc \\
&+\mu_{MA}[I^{MA}-\delta^{MA}MA-sMA]+\mu_{CMA}[I^{MA}-TDEP^{MA}-sMA] \\
&+\mu_{BU}[I^{BU}-\delta^{BU}BU]+\mu_{OFA}\,dofa+\mu_{CA}\,dca+\mu_{SC}\,dsc+\mu_{RR}\,drr \\
&+\mu_{ASD}[TDEP^{MA}-\delta^{MA}MA]+\mu_{PF}[p^{allo}-zPF]+\mu_{OUR}\,dour+\mu_{LL}\,dll+\mu_{CL}\,dcl \\
&+n_{r}\{I^{MA}+I^{BU}+dofa+dca+URE-\delta^{BU}BU-drr+zPF-p^{allo}-dour-s\,MA \\
&-TDEP^{MA}-dsc-dll-dcl\} \\
&+n_{f}\{dsc+dll+dcl-I^{MA}-s\,MA-I^{BU}-dofa-dca+(1-\tau^{e})[OIBD+FI-FE \\
&+\tau^{e}[\delta^{BU}BU+TDEP^{MA}-zPF+dour]+\tau * p^{allo}+\tau^{\beta}\delta_{T}^{BU}BU\} \\
&+n_{u}^{MA}\{\delta_{T}^{MA}CMA-TDEP^{MA}\}+n_{u}^{PF}\{\eta * a(1+g)[OIBD-\delta^{BU}BU+FI-FE \\
&-TDEP^{MA}+zPF-dour]-\eta * \delta_{T}^{BU}BU-p^{allo}\} \\
&+n_{MA}MA+n_{CMA}CMA+n_{BU}BU+n_{OFA}OFA+n_{CA}CA+n_{SC}(SC-\overline{SC})+n_{RR}RR+n_{PF}PF \\
&+n_{OUR}OUR+n_{ASD}ASD+n_{LL}LL+n_{CL}CL
\end{aligned}
$$

where

(3.68) $\tau^{*}=\xi\tau$

(3.69) $a=(1+\gamma^{OT})(1-\alpha)$

(3.70) $\xi=(1-\beta^{RO})a$

(3.71) $\tau^{e}=\tau^{*}(1+g)-g$

(3.72) $\tau^{\beta}=(1-\beta^{RO})\tau$

(3.73) $\eta^{*}=\dfrac{\eta}{1-\eta(1-a)}$

Equations for $OIBD$, FI, and FE were defined in (3.9)-(3.11). μ_{MA}, μ_{CMA}, μ_{BU}, μ_{OFA}, μ_{CA}, μ_{SC}, μ_{RR}, μ_{ASD}, μ_{PF}, μ_{OUR}, μ_{LL}, and μ_{CL} are the shadow prices or co-state variables of the stock of machinery and equipment (MA), the stock of book value of machinery and equipment (CMA), the stock of buildings (BU),the stock of other fixed assets (OFA), the stock of current assets (CA), the stock of share capital (SC), the stock of restricted reserves (RR), accumulated supplementary depreciation (ASD), periodical reserves (PF), the stock of other untaxed reserves (OUR), the stock of long-term liabilities (LL), and the stock of current liabilities (CL). n_r, n_f, n_u^{MA}, n_u^{PF}, n_{MA}, n_{CMA}, n_{BU}, n_{OFA}, n_{CA}, n_{SC}, n_{RR}, n_{PF}, n_{OUR}, n_{ASD}, n_{LL}, and n_{CL} are the Khun-Tucker shadow-price of constraints (3.47), (3.48), (3.52), and (3.55)-(3.67).

The first-order necessary conditions are:

(3.74) $I^{MA}:\ -\theta+\mu_{MA}+\mu_{CMA}+n_{r}-n_{f}=0$

(3.75) $I^{BU}:\ -\theta+\mu_{BU}+n_{r}-n_{f}=0$

(3.76) $dofa:\ -\theta+\mu_{OFA}+n_{r}-n_{f}=0$

(3.77) $dca:\ -\theta+\mu_{CA}+n_r-n_f=0$

(3.78) $dsc:\ \theta-\mu_{SC}-n_r+n_f=0$

(3.79) $drr:\ \mu_{RR}-n_r=0$

(3.80) $dour:\ \theta\tau^e+\mu_{OUR}-n_r+\tau^e n_f-\eta*a(1+g)n_u^{PF}=0$

(3.81) $p^{allo}:\ \theta\tau*+\mu_{PF}-n_r+\tau*n_f-n_u^{PF}=0$

(3.82) $TDEP^{MA}:\ \theta\tau^e-\mu_{CMA}+\mu_{ASD}-n_r+\tau^e n_f-\eta*a(1+g)n_u^{PF}-n_u^{MA}=0$

(3.83) $zPF:\ -\theta\tau^e-\mu_{PF}+n_r-\tau^e n_f+\eta*a(1+g)n_u^{PF}=0$

(3.84) $dll:\ \theta+\mu_{LL}-n_r+n_f=0$

(3.85) $dcl:\ \theta+\mu_{CL}-n_r+n_f=0$

(3.86) $L:\ (f_L-W)\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}=0$

(3.87) $$MA:\ \dot{\mu}_{MA}=(r+s+\delta^{MA})\mu_{MA}-\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}[f_{MA}+h-n] \\ +s\mu_{CMA}+\delta^{MA}\mu_{ASD}+s(n_r-\theta-n_f)-n_{MA}$$

(3.88) $CMA:\ \dot{\mu}_{CMA}=r\mu_{CMA}-\delta_T^{MA}n_u^{MA}-n_{CMA}$

(3.89) $$BU:\ \dot{\mu}_{BU}=(r+\delta^{BU})\mu_{BU}-\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}[f_{BU}+h-n] \\ +\tau^{\beta}\delta_T^{BU}+\{n_r-(\theta+n_f)\tau^e+\eta*a(1+g)n_u^{PF}\}\delta^{BU}-n_{BU}+\eta*\delta_T^{BU}n_u^{PF}$$

(3.90) $OFA:\ \dot{\mu}_{OFA}=r\mu_{OFA}-\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}(h-n)-n_{OFA}$

(3.91) $CA:\ \dot{\mu}_{CA}=r\mu_{CA}-\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}(f_{CA}+h-n)-n_{CA}$

(3.92) $SC:\ \dot{\mu}_{SC}=r\mu_{SC}-n_{SC}$

(3.93) $RR:\ \dot{\mu}_{SC}=r\mu_{SC}-n_{SC}$

(3.94) $OUR:\ \dot{\mu}_{OUR}=r\mu_{OUR}-n_{OUR}$

(3.95) $PF:\ \dot{\mu}_{PF}=r\mu_{PF}-n_{PF}$

(3.96) $ASD:\ \dot{\mu}_{ASD}=r\mu_{ASD}-n_{ASD}$

(3.97) $LL:\ \dot{\mu}_{LL}=r\mu_{LL}+\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}i_B-n_{LL}$

(3.98) $CL:\ \dot{\mu}_{CL}=r\mu_{CL}+\{(1-\tau^e)(\theta+n_f)+\eta*a(1+g)n_u^{PF}\}(i_B+f_{CL})-n_{CL}$

(3.99) $n_r:\ \dfrac{\delta F}{\delta n_r}\geq 0,\ n_r\geq 0,\ n_r\dfrac{\delta F}{\delta n_r}=0$

(3.100) $n_f:\ \dfrac{\delta F}{\delta n_f}\geq 0,\ n_f\geq 0,\ n_f\dfrac{\delta F}{\delta n_f}=0$

(3.101) $n_u^{MA}:\ \dfrac{\delta F}{\delta n_u^{MA}}\geq 0,\ n_u^{MA}\geq 0,\ n_u^{MA}\dfrac{\delta F}{\delta n_u^{MA}}=0$

(3.102) $n_u^{PF}:\ \dfrac{\delta F}{\delta n_u^{PF}}\geq 0,\ n_u^{PF}\geq 0,\ n_u^{PF}\dfrac{\delta F}{\delta n_u^{PF}}=0$

(3.103) $n_{CMA}: \dfrac{\delta F}{\delta n_{CMA}} \geq 0,\ n_{CMA} \geq 0,\ n_{CMA}\dfrac{\delta F}{\delta n_{CMA}} = 0$

(3.104) $n_{MA}: \dfrac{\delta F}{\delta n_{MA}} \geq 0,\ n_{MA} \geq 0,\ n_{MA}\dfrac{\delta F}{\delta n_{MA}} = 0$

(3.105) $n_{BU}: \dfrac{\delta F}{\delta n_{BU}} \geq 0,\ n_{BU} \geq 0,\ n_{BU}\dfrac{\delta F}{\delta n_{BU}} = 0$

(3.106) $n_{OFA}: \dfrac{\delta F}{\delta n_{OFA}} \geq 0,\ n_{OFA} \geq 0,\ n_{OFA}\dfrac{\delta F}{\delta n_{OFA}} = 0$

(3.107) $n_{CA}: \dfrac{\delta F}{\delta n_{CA}} \geq 0,\ n_{CA} \geq 0,\ n_{CA}\dfrac{\delta F}{\delta n_{CA}} = 0$

(3.108) $n_{SC}: \dfrac{\delta F}{\delta n_{SC}} \geq 0,\ n_{SC} \geq 0,\ n_{SC}\dfrac{\delta F}{\delta n_{SC}} = 0$

(3.109) $n_{RR}: \dfrac{\delta F}{\delta n_{RR}} \geq 0,\ n_{RR} \geq 0,\ n_{RR}\dfrac{\delta F}{\delta n_{RR}} = 0$

(3.110) $n_{PF}: \dfrac{\delta F}{\delta n_{PF}} \geq 0,\ n_{PF} \geq 0,\ n_{PF}\dfrac{\delta F}{\delta n_{PF}} = 0$

(3.111) $n_{OUR}: \dfrac{\delta F}{\delta n_{OUR}} \geq 0,\ n_{OUR} \geq 0,\ n_{OUR}\dfrac{\delta F}{\delta n_{OUR}} = 0$

(3.112) $n_{ASD}: \dfrac{\delta F}{\delta n_{ASD}} \geq 0,\ n_{ASD} \geq 0,\ n_{ASD}\dfrac{\delta F}{\delta n_{ASD}} = 0$

(3.113) $n_{LL}: \dfrac{\delta F}{\delta n_{LL}} \geq 0,\ n_{LL} \geq 0,\ n_{LL}\dfrac{\delta F}{\delta n_{LL}} = 0$

(3.114) $n_{CL}: \dfrac{\delta F}{\delta n_{CL}} \geq 0,\ n_{CL} \geq 0,\ n_{CL}\dfrac{\delta F}{\delta n_{CL}} = 0$

We impose stationary constraints on (3.87)-(3.98), i.e.

(3.115) $$\begin{aligned}&\dot{\mu}_{MA} = \dot{\mu}_{CMA} = \dot{\mu}_{BU} = \dot{\mu}_{OFA} = \dot{\mu}_{CA} = \dot{\mu}_{SC} = \dot{\mu}_{RR}\\ &= \dot{\mu}_{OUR} = \dot{\mu}_{PF} = \dot{\mu}_{ASD} = \dot{\mu}_{LL} = \dot{\mu}_{CL} = 0\end{aligned}$$

Equation (3.115) together with equations (3.87)-(3.98) give the values of μ_{MA}, μ_{CMA}, μ_{BU}, μ_{OFA}, μ_{CA}, μ_{SC}, μ_{RR}, μ_{ASD}, μ_{PF}, μ_{OUR}, μ_{LL}, and μ_{CL}

(3.116) $$\begin{aligned}\mu_{MA} = \frac{1}{r+s+\delta^{MA}}\{&[(1-\tau^{e})(\theta+n_{f})+\eta * a(1+g)n_{u}^{PF}][f_{MA}+h-n]\\ &-s\mu_{CMA}-\delta^{MA}\mu_{ASD}-s(n_{r}-\theta-n_{f})+n_{MA}\}\end{aligned}$$

(3.117) $\mu_{CMA} = \dfrac{1}{r}[\delta_{T}^{MA}n_{u}^{MA}+n_{CMA}]$

(3.118) $\mu_{ASD} = \dfrac{n_{ASD}}{r}$

(3.119) $$\mu_{BU} = \frac{1}{r+\delta^{BU}}\{[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}][f_{BU}+h-n]+(\theta+n_f)\tau^{\beta}\delta_T^{BU} -[n_r-(\theta+n_f)\tau^e+\eta^* a(1+g)n_u^{PF}]\delta^{BU}+n_{BU}-\eta^*\delta_T^{BU}n_u^{PF}\}$$

(3.120) $$\mu_{OFA} = \frac{1}{r}\{[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}](h-n)+n_{OFA}\}$$

(3.121) $$\mu_{CA} = \frac{1}{r}\{[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}](f_{CA}+h-n)+n_{CA}\}$$

(3.122) $$\mu_{SC} = \frac{n_{SC}}{r}$$

(3.123) $$\mu_{RR} = \frac{n_{RR}}{r}$$

(3.124) $$\mu_{OUR} = \frac{n_{OUR}}{r}$$

(3.125) $$\mu_{PF} = \frac{n_{PF}}{r}$$

(3.126) $$\mu_{ASD} = \frac{n_{ASD}}{r}$$

(3.127) $$\mu_{LL} = \frac{1}{r}\{-[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}]i_B+n_{LL}\}$$

(3.128) $$\mu_{CL} = \frac{1}{r}\{-[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}](i_B+f_{CL})+n_{CL}\}$$

Using the steady state solutions for μ_{MA}, μ_{CMA}, μ_{BU}, μ_{OFA}, μ_{CA}, μ_{SC}, μ_{RR}, μ_{ASD}, μ_{PF}, μ_{OUR}, μ_{LL}, and μ_{CL} in equations (3.74)-(3.86), the first-order conditions in steady state become

(3.129) $$I^{MA}\text{:}\; -r(r+\delta^{MA})(\theta+n_f-n_r)+\delta_T^{MA}(r+\delta^{MA})n_u^{MA}+(r+\delta^{MA})n_{CMA}+rn_{CMA} -\delta^{MA}n_{ASD}+[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}][f_{MA}+h-n]r=0$$

(3.130) $$I^{BU}:\; -rn_r+[-r-(1-\tau^e)(1-\delta^{BU})+\tau^{\beta}\delta_T^{BU}](\theta+n_f)+n_{BU}-\eta^*[a(1+g)\delta^{BU} -\delta_T^{BU}]n_u^{PF}+[f_{MA}+h-n][(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}]=0$$

(3.131) $dofa:\; -r(\theta+n_f-n_r)+[h-n][(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}]+n_{OFA}=0$

(3.132) $dca:\; -r(\theta+n_f-n_r)+[f_{CA}+h-n][(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}]+n_{OFA}=0$

(3.133) $dsc:\; r(\theta+n_f-n_r)+n_{SC}=0$

(3.134) $drr:\; -rn_r+n_{RR}=0$

(3.135) $dour:\; r[\tau^e(\theta+n_f)-n_r]-r\eta^* a(1+g)n_u^{PF}+n_{OUR}=0$

(3.136) $p^{allo}:\; r[\tau^e(\theta+n_f)-n_r]-rn_u^{PF}+n_{PF}=0$

(3.137) $TDEP^{MA}:\; r[\tau^e(\theta+n_f)-n_r]-(r+\delta_T^{MA})n_u^{MA}-n_{CMA}+n_{ASD}-r\eta^* a(1+g)n_u^{PF}=0$

(3.138) $zpf:\; -r[\tau^e(\theta+n_f)-n_r]+n_{PF}+r\eta^* a(1+g)n_u^{PF}=0$

(3.139) $dll:\; r(\theta+n_f-n_r)-i_B[(1-\tau^e)(\theta+n_f)+\eta^* a(1+g)n_u^{PF}]+n_{LL}=0$

(3.140) $dcl:\ r(\theta + n_f - n_r) - (i_B + f_{CL})[(1-\tau^e)(\theta + n_f) + \eta^* a(1+g)n_u^{PF}] + n_{CL} = 0$

(3.141) $L:\ (f_L - W)\{(1-\tau^e)(\theta + n_f) + \eta^* a(1+g)n_u^{PF}\} = 0$

3.4 The Economic Relationships between Different Decision Variables

In this section we will subject the optimization to investigations of the comparative-static sort. The idea is to find out how a change in any parameter or exogenous variable will affect the equilibrium position of the model, which in the present context refers to the optimal values of the decision variables (and the optimal value of the objective function). If we let a disequilibria change occur in the model, the initial equilibrium will, of course, be upset. As a result, the various endogenous variables (decision variables) must undergo certain adjustments. If it is assumed that a new equilibrium state relevant to the new values of the data can be defined and attained, the question posed in the analysis is: how would the new equilibrium compare with the old? In this case, we merely compare the initial (pre-change) equilibrium state with the final (post-change) equilibrium state. Also, we preclude the possibility of instability of equilibrium, for we assume the new equilibrium to be attainable, just as we do for the old.

We are concerned with the magnitude of the change in decision variables' equilibrium values resulting from a given change in parameters or exogenous variables. The rate of change of the equilibrium values of the endogenous variable are found by total differentiating the first-order conditions in equilibrium (3.129)-(3.141). We are not interested in an analysis of qualitative nature. For the simulation model we need to do an analysis of quantitative nature. To be able to obtain a quantitative answer, however, we use econometric methods.

We know that the partial derivatives measure the rate of change in the first-order conditions with respect to an infinitesimal change in a parameter, an exogenous variable, or an endogenous variable. Taking advantage of the fact that, in the neighborhood *N*, (3.129)-(3.141) and (3.99)-(3.114) have the status of identities, we can take the total differential of each of these. The result is a set of equations involving the differentials.

The total differentiation of (3.129) yields

(3.142)
$$\begin{aligned}&\beta_1^{MA} dMA + \beta_2^{MA} dBU + \beta_3^{MA} dCA + \beta_4^{MA} dL + \beta_5^{MA} dn_f + \beta_6^{MA} dn_r \\ &+ \beta_7^{MA} dn_u^{MA} + \beta_8^{MA} dn_{CMA} + \beta_9^{MA} dn_{ASD} + \beta_{10}^{MA} dn_u^{PF} + \beta_{11}^{MA} n_{MA} + \beta_{21}^{MA} dCL = \beta_0^{MA}\end{aligned}$$

where

$$\begin{aligned}\beta_0^{MA} &= (\theta + n_f - n_r)d(r + \delta^{MA}) - n_u^{MA} d[\delta_T^{MA}(r + \delta^{MA})] - n_{CMA} d(r + \delta^{MA}) \\ &- n_{ASD} d\delta^{MA} - Ar[dh - dn] - n_{MA} dr - A[f_{MA} + h - n]dr \\ &- [f_{MA} + h - n]\{(\theta + n_f)d(1-\tau^e) + (1-\tau^e)d\theta + n_u^{PF} d[\eta^* a(1+g)]\}\end{aligned}$$

and $\beta_1^{MA} = rA\, f_{MAMA}$, $\beta_2^{MA} = rA\, f_{MABU}$, $\beta_3^{MA} = rA\, f_{MACA}$, $\beta_4^{MA} = rA\, f_{MAL}$, $\beta_5^{MA} = -(r + \delta^{MA}) + (1-\tau^e)[f_{MA} + h - n]r$, $\beta_6^{MA} = r[r + \delta^{MA}]$, $\beta_7^{MA} = \delta_T^{MA}(r + \delta^{MA})$, $\beta_8^{MA} = r + \delta^{MA}$, $\beta_9^{MA} = -\delta^{MA}$, $\beta_{10}^{MA} = [f_{MA} + h - n]r\eta^* a(1+g)$, $\beta_{11}^{MA} = r$, $\beta_{21}^{MA} = -rA\, f_{MACL}$, $A = (1-\tau^e)(\theta + n_f) + \eta^* a(1+g)n_u^{PF}$.

The total differentiation of (3.130) yields

$$\begin{aligned}(3.143)\quad & \beta_1^{BU} dMA + \beta_2^{BU} dBU + \beta_3^{BU} dCA + \beta_4^{BU} dL + \beta_5^{BU} dn_f + \beta_6^{BU} dn_r \\ & + \beta_{10}^{BU} dn_u^{PF} + \beta_{12}^{BU} dn_{BU} + \beta_{21}^{BU} dCL = \beta_0^{BU}\end{aligned}$$

where

$$\begin{aligned}\beta_0^{BU} &= -n_r dr - (\theta + n_f) d[-r - (1-\tau^e)(1-\delta^{BU}) + \tau^\beta \delta_T^{BU}] - [-r - (1-\tau^e)(1-\delta^{BU}) + \tau^\beta \delta_T^{BU}] d\theta \\ &\quad + n_u^{PF} d\{\eta * [a(1+g)\delta^{BU} - \delta_T^{BU}]\} - [(1-\tau^e)(\theta + n_f) + \eta * a(1+g) n_u^{PF}][dh - dn] \\ &\quad - [f_{BU} + h - n]\{(\theta + n_f) d(1-\tau^e) + (1-\tau^e) d\theta + n_u^{PF} d[\eta * a(1+g)]\}\end{aligned}$$

and $\beta_1^{BU} = A\, f_{BU\,MA}$, $\beta_2^{BU} = A\, f_{BU\,BU}$, $\beta_3^{BU} = A\, f_{BU\,CA}$, $\beta_4^{MA} = A\, f_{BU\,L}$, $\beta_5^{BU} = -r - (1-\tau^e)(1-\delta^{BU}) + \tau^\beta \delta_T^{BU} + (1-\tau^e)(f_{BU} + h - n)$, $\beta_6^{BU} = r$, $\beta_{10}^{BU} = -\eta * [a(1+g)\delta^{BU} - \delta_T^{BU}]$, $\beta_{12}^{BU} = 1$, $\beta_{21}^{BU} = -A\, f_{BU\,CL}$.

The total differentiation of (3.131) yields

$$(3.144)\quad \beta_5^{OFA} dn_f + \beta_6^{OFA} dn_r + \beta_{10}^{OFA} dn_u^{PF} + \beta_{13}^{OFA} dn_{OFA} = \beta_0^{OFA}$$

where

$$\begin{aligned}\beta_0^{OFA} &= d\theta - Ad\left[\frac{h-n}{r}\right] - \left[\frac{h-n}{r}\right]\{(\theta + n_f) d(1-\tau^e) + (1-\tau^e) d\theta + n_u^{PF} d[\eta * a(1+g)]\} \\ &\quad - n_{OFA} d\left[\frac{1}{r}\right] r\end{aligned}$$

and $\beta_5^{OFA} = -1 + (1-\tau^e)\left[\frac{h-n}{r}\right]$, $\beta_6^{OFA} = 1$, $\beta_{10}^{BU} = \left[\frac{h-n}{r}\right]\eta * a(1+g)$, $\beta_{13}^{OFA} = \left[\frac{h-n}{r^2}\right]$.

The total differentiation of (3.132) yields

$$\begin{aligned}(3.145)\quad & \beta_1^{CA} dMA + \beta_2^{CA} dBU + \beta_3^{CA} dCA + \beta_4^{CA} dL + \beta_5^{CA} dn_f + \beta_6^{CA} dn_r \\ & + \beta_{10}^{CA} dn_u^{PF} + \beta_{14}^{CA} dn_{CA} + \beta_{21}^{CA} dCL = \beta_0^{CA}\end{aligned}$$

where

$$\begin{aligned}\beta_0^{CA} &= (\theta + n_f - n_r) dr + r d\theta - A[dh - dn] \\ &\quad - [f_{CA} + h - n]\{(\theta + n_f) d(1-\tau^e) + (1-\tau^e) d\theta + n_u^{PF} d[\eta * a(1+g)]\}\end{aligned}$$

and $\beta_1^{CA} = A\, f_{CAMA}$, $\beta_2^{CA} = A\, f_{CABU}$, $\beta_3^{CA} = A\, f_{CACA}$, $\beta_4^{CA} = A\, f_{CAL}$, $\beta_5^{CA} = -r + (1-\tau^e)(f_{CA} + h - n)$, $\beta_6^{CA} = 1$, $\beta_{10}^{CA} = \eta * a(1+g)(f_{CA} + h - n)$, $\beta_{14}^{CA} = 1$, $\beta_{21}^{CA} = -A\, f_{CACL}$.

The total differentiation of (3.133) yields

(3.146) $\beta_5^{SC} dn_f + \beta_6^{SC} dn_r + \beta_{15}^{SC} dn_{SC} = \beta_0^{SC}$

where $\beta_0^{SC} = -rd\theta - (\theta + n_f - n_r)dr$, $\beta_5^{SC} = r$, $\beta_6^{CA} = -r$, and $\beta_{15}^{SC} = 1$.

The total differentiation of (3.134) yields

(3.147) $\beta_6^{RR} dn_r + \beta_{16}^{RR} dn_{RR} = \beta_0^{RR}$

where $\beta_0^{RR} = n_r dr$, $\beta_6^{RR} = -r$, and $\beta_{16}^{RR} = 1$.

The total differentiation of (3.135) yields

(3.148) $\beta_5^{OUR} dn_f + \beta_6^{OUR} dn_r + \beta_{10}^{OUR} dn_u^{PF} + \beta_{17}^{OUR} dn_{OUR} = \beta_0^{OUR}$

where $\beta_0^{OUR} = -r[d(\theta\tau^e) + n_f d\tau^e] - [\tau^e(\theta + n_f) - n_r]dr$, $\beta_5^{OUR} = r\tau^e$, $\beta_6^{OUR} = -r$, $\beta_{10}^{OUR} = -r\eta * a(1+g)$ and $\beta_{17}^{OUR} = 1$.

The total differentiation of (3.136) yields

(3.149) $\beta_5^{PPF} dn_f + \beta_6^{PPF} dn_r + \beta_{10}^{PPF} dn_u^{PF} + \beta_{18}^{PPF} dn_{PF} = \beta_0^{PPF}$

where $\beta_0^{PPF} = -r[d(\theta\tau^*) + n_f d\tau^*] - [\tau^*(\theta + n_f) - n_r]dr$, $\beta_5^{PPF} = r\tau^*$, $\beta_6^{PPF} = -r$, $\beta_{10}^{PPF} = -r$ and $\beta_{18}^{PPF} = 1$.

The total differentiation of (3.137) yields

(3.150) $\beta_5^{TMA} dn_f + \beta_6^{TMA} dn_r + \beta_7^{TMA} dn_u^{MA} + \beta_8^{TMA} dn_{CMA} + \beta_9^{TMA} dn_{ASD} + \beta_{10}^{TMA} dn_u^{PF} = \beta_0^{TMA}$

where $\beta_0^{TMA} = -r[d(\theta\tau^e) + n_f d\tau^e] - [\tau^e(\theta + n_f) - n_r]dr + n_u^{MA} d(r + \delta_T^{MA}) + n_u^{PF} d[r\eta * a(1+g)]$, $\beta_5^{TMA} = r\tau^e$, $\beta_6^{TMA} = -r$, $\beta_7^{TMA} = -(r + \delta_T^{MA})$, $\beta_8^{TMA} = -1$, $\beta_9^{TMA} = 1$, and $\beta_{10}^{TMA} = -r\eta * a(1+g)$.

The total differentiation of (3.138) yields

(3.151) $\beta_5^{ZPF} dn_f + \beta_6^{ZPF} dn_r + \beta_{10}^{ZPF} dn_u^{PF} + \beta_{18}^{ZPF} dn_{PF} = \beta_0^{ZPF}$

where $\beta_0^{ZPF} = r[d(\theta\tau^e) + n_f d\tau^e] - [\tau^e(\theta + n_f) - n_r]dr$, $\beta_5^{ZPF} = -r\tau^e$, $\beta_6^{ZPPF} = -r$, $\beta_{10}^{ZPF} = r\eta * a(1+g)$ and $\beta_{18}^{ZPF} = 1$.

The total differentiation of (3.139) yields

(3.152) $\beta_5^{LL} dn_f + \beta_6^{LL} dn_r + \beta_{10}^{LL} dn_u^{PF} + \beta_{19}^{LL} dn_{LL} = \beta_0^{LL}$

where $\beta_0^{LL} = -rd\theta - [\theta + n_f - n_r]dr + [\theta + n_f]d[(1-\tau^e)i_B] - (1-\tau^e)i_B d\theta + n_u^{PF} d[\eta * a(1+g)]$, $\beta_5^{LL} = r + (1-\tau^e)i_B$, $\beta_6^{LL} = r$, $\beta_{10}^{LL} = -\eta * a(1+g)i_B$ and $\beta_{19}^{LL} = 1$.

The total differentiation of (3.140) yields

$$
(3.153) \quad \begin{aligned} &\beta_1^{CL} dMA + \beta_2^{CL} dBU + \beta_3^{CL} dCA + \beta_4^{CL} dL + \beta_5^{CL} dn_f \\ &+ \beta_6^{CL} dn_r + \beta_{10}^{CL} dn_u^{PF} + \beta_{20}^{CL} dn_{CL} = \beta_0^{CL} \end{aligned}
$$

where
$\beta_0^{CL} = -rd\theta - [\theta + n_f - n_r]dr + (i_B + f_{CL})\{[\theta + n_f]d(1-\tau^e) + (1-\tau^e)d\theta + n_u^{PF} d[\eta * a(1+g)]\}$,
$\beta_1^{CL} = A\, f_{CL\,MA}$, $\beta_2^{CL} = A\, f_{CL\,BU}$, $\beta_3^{CL} = A\, f_{CL\,CA}$, $\beta_4^{CL} = A\, f_{CL\,L}$, $\beta_5^{CL} = [r - (1-\tau^e)(i_B + f_{CL})]$,
$\beta_6^{CL} = -r$, $\beta_{10}^{CL} = -\eta * a(1+g)(i_B + f_{CL})$, $\beta_{20}^{CL} = 1$, and $\beta_{21}^{CL} = -A\, f_{CL\,CL}$.

The total differentiation of (3.141) yields

(3.154) $\beta_1^L dMA + \beta_2^L dBU + \beta_3^L dCA + \beta_4^L dL + \beta_5^L dn_f + \beta_{10}^L dn_u^{PF} + \beta_{21}^L dCL = \beta_0^L$

where $\beta_0^L = AdW + (f_L - W)\{(\theta + n_f)d(1-\tau^e) + (1-\tau^e)d\theta + n_u^{PF} d[\eta * a(1+g)]\}$,
$\beta_1^L = A\, f_{L\,MA}$, $\beta_2^L = A\, f_{L\,BU}$, $\beta_3^L = A\, f_{L\,CA}$, $\beta_4^L = A\, f_{L\,L}$, $\beta_5^L = (f_L - W)(1-\tau^e)$,
$\beta_{10}^L = [f_L - W]\eta * a(1+g)$, $\beta_{11}^{MA} = r$, $\beta_{21}^L = -A\, f_{L\,CL}$.

It is important to mention that the second-order partial derivatives' β s are evaluated at their initial values. It is obvious that the second partial derivatives will assume specific values at the points we are examining as possible extremum points, and thus may be regarded as *constants*.

Before we proceed with the total differentiation of the complementary-slackness conditions in (3.99)-(3.114), the following definitions are necessary

(3.155) $dmcash = mcash - cashfl$

(3.156) $dmtdm = MTDM - TDEP^{MA}$

(3.157) $dmpa = MPA - p^{allo}$

where $mcash = NBI + URE - drr$. Using (3.155)-(3.157) we can rewrite the complementary-slackness conditions in (3.99), (3.101) and (3.102) as follows

(3.158) $n_r\, dmcash = 0$

(3.159) $n_u^{MA}\, dmtdm = 0$

(3.160) $n_u^{PF}\, dmpa = 0$

These conditions are easy to interpret. For example condition (3.160) stipulates that, if the periodical reserve option is not fully used in the optimal solution ($\frac{\delta F}{\delta n_u^{PF}} \geq 0$), the shadow price of that constraint, which is never allowed to be negative, must be equal to zero ($n_u^{PF} = 0$). On the other hand, if the constraint has a positive shadow price in the optimal solution ($n_u^{PF} > 0$), then it is perforce a fully utilized allocation to periodical reserves ($\frac{\delta F}{\delta n_u^{PF}} = 0$). The Lagrange-multiplier n_u^{PF} is a measure of how the optimal value of firms reacts to a slight relaxation of the constraint. In that light, complementary slackness means that, if the constraint is optimally not binding, then relaxing that particular constraint will not affect the optimal value of firms. If, on the other hand, a slight relaxation of the constraint does increase the value of firms, then that constraint must in fact be binding in the optimal solution. Let us now try to find out how slight relaxations of the constraints affect the complementary-slackness conditions in (3.155), (3.100), (3.156), (3.157), (3.103)-(3.114) by total differentiating these conditions.

(3.161) $dn_r\, dmcash + n_r\, ddmcash = 0$

(3.162) $dn_f\, cashfl + n_f\, dcashfl = 0$

(3.163) $dn_u^{MA}\, dmtdm + n_u^{MA}\, ddmtdm = 0$

(3.164) $dn_u^{PF}\, dmpa + n_u^{PF}\, ddmpa = 0$

(3.165) $dn_{CMA}\, CMA + n_{CMA}\, dCMA = 0$

(3.166) $dn_{MA}\, MA + n_{MA}\, dMA = 0$

(3.167) $dn_{BU}\, BU + n_{BU}\, dBU = 0$

(3.168) $dn_{OFA}\, OFA + n_{OFA}\, dOFA = 0$

(3.169) $dn_{CA}\, CA + n_{CA}\, dCA = 0$

(3.170) $dn_{SC}\, SC + n_{SC}\, dSC = 0$

(3.171) $dn_{RR}\, RR + n_{RR}\, dRR = 0$

(3.172) $dn_{PF}\, PF + n_{PF}\, dPF = 0$

(3.173) $dn_{OUR}\, OUR + n_{OUR}\, dOUR = 0$

(3.174) $dn_{ASD}\, ASD + n_{ASD}\, dASD = 0$

(3.175) $dn_{LL}\, LL + n_{LL}\, dLL = 0$

(3.176) $dn_{CL}\,CL + n_{CL}\,dCL = 0$

Solving (3.162) and (3.164) for dn_f and dn_u^{PF}, and inserting them into equation (3.154), we can derive the following expression for dL

(3.177) $dL = \gamma_0^L + \gamma_1^L dMA + \gamma_2^L dBU + \gamma_3^L dCA + \gamma_4^L dcashfl + \gamma_5^L ddmpa + \gamma_{32}^L dCL$

where $\gamma_0^L = (\beta_0^L / \beta_4^L)$, $\gamma_1^L = -(\beta_1^L / \beta_4^L)$, $\gamma_2^L = -(\beta_2^L / \beta_4^L)$, $\gamma_3^L = -(\beta_3^L / \beta_4^L)$, $\gamma_4^L = (\beta_5^L / \beta_4^L)(n_f / cashfl)$, $\gamma_5^L = (\beta_{10}^L / \beta_4^L)(n_u^{PF} / dmpa)$, and $\gamma_{32}^L = -(\beta_{21}^L / \beta_4^L)$. Now, solving (3.161)-(3.176) for dn_r, dn_f, dn_u^{MA}, dn_u^{PF}, dn_{CMA}, dn_{MA}, dn_{BU}, dn_{OFA}, dn_{CA}, dn_{SC}, dn_{RR}, dn_{PF}, dn_{OUR}, dn_{ASD}, dn_{LL}, and dn_{CL}, and inserting these expressions (together with the expression for dL) into (3.142)-(3.153), we obtain the following relationships between different flow variables

1) Investment in machinery and equipment: Equation (3.142) can be rewritten as follows

$$
\begin{aligned}
dMA = {} & \gamma_0^{MA} + \gamma_2^{MA} I^{BU} + \gamma_3^{MA} dca + \gamma_4^{MA} dcashfl + \gamma_5^{MA} ddmpa + \gamma_6^{MA} ddmcash \\
& + \gamma_7^{MA} ddmtdm + \gamma_8^{MA} TDEP^{MA} + \gamma_9^{MA} EDEP^{MA} + \gamma_{10}^{MA} EDEP^{BU} + \gamma_{32}^{MA} dCL
\end{aligned} \tag{3.178}
$$

where $B = \beta_1^{MA} + \beta_4^{MA}\gamma_1^L - \beta_8^{MA}(n_{CMA} / CMA)$, $\gamma_0^{MA} = (1/B)(\beta_0^{MA} - \beta_4^{MA}\gamma_0^L)$, $\gamma_2^{MA} = (1/B)(\beta_2^{MA} + \beta_4^{MA}\gamma_2^L)$, $\gamma_3^{MA} = (1/B)(\beta_3^{MA} + \beta_4^{MA}\gamma_3^L)$, $\gamma_6^{MA} = (1/B)\beta_6^{MA}(n_r / dmcash)$, $\gamma_4^{MA} = (1/B)[\beta_5^{MA}(n_f / cashfl) - \beta_4^{MA}\gamma_4^L]$, $\gamma_7^{MA} = (1/B)\beta_7^{MA}(n_u^{MA} / dmtdm)$, $\gamma_5^{MA} = (1/B)[\beta_{10}^{MA}(n_u^{PF} / dmpa) - \beta_4^{MA}\gamma_5^L]$, $\gamma_8^{MA} = (1/B)[\beta_8^{MA}(n_{CMA} / CMA) + \beta_9^{MA}(n_{ASD} / ASD)$, $\gamma_9^{MA} = -(1/B)[\beta_8^{MA}(n_{CMA} / CMA) + \beta_9^{MA}(n_{ASD} / ASD)$, $\gamma_{10}^{MA} = -(1/B)(\beta_2^{MA} + \beta_4^{MA}\gamma_2^L)$, and $\gamma_{32}^{MA} = -(1/B)(\beta_{21}^{MA} + \beta_4^{MA}\gamma_{32}^L)$.

From equation (3.32), we can derive the following expression for I^{MA}

(3.179) $I^{MA} = dMA + EDEP^{MA} + SMA$

Inserting equation (3.178) into (3.179) we get the following economic relationship for I^{MA}

$$
\begin{aligned}
I^{MA} = {} & \gamma_0^{MA} + \gamma_2^{MA} I^{BU} + \gamma_3^{MA} dCA + \gamma_4^{MA} dcashfl + \gamma_5^{MA} ddmpa + \gamma_6^{MA} ddmcash \\
& + \gamma_7^{MA} ddmtdm + \gamma_8^{MA} TDEP^{MA} + \gamma_9^{MA*} EDEP^{MA} + \gamma_{10}^{MA} EDEP^{BU} + \gamma_{11}^{MA} SMA + \gamma_{32}^{MA} dCL
\end{aligned} \tag{3.180}
$$

where $\gamma_9^{MA*} = 1 + \gamma_9^{MA}$.

2) Investment in buildings: Equation (3.143) can be rewritten as follows

(3.181) $$dBU = \gamma_0^{BU} + \gamma_1^{BU} I^{MA} + \gamma_3^{BU} dca + \gamma_4^{BU} dcashfl + \gamma_5^{BU} ddmpa + \gamma_6^{BU} ddmcash + \gamma_9^{BU} EDEP^{MA} + \gamma_{11}^{BU} SMA + \gamma_{32}^{BU} dCL$$

where $C = \beta_2^{BU} + \beta_4^{MA}\gamma_2^L - \beta_{12}^{BU}(n_{BU}/BU)$, $\gamma_0^{BU} = (1/C)(\beta_0^{BU} - \beta_4^{BU}\gamma_0^L)$, $\gamma_1^{BU} = -(1/C)(\beta_1^{BU} + \beta_4^{BU}\gamma_1^L)$, $\gamma_3^{BU} = -(1/C)(\beta_3^{BU} + \beta_4^{BU}\gamma_3^L)$, $\gamma_4^{BU} = (1/C)[\beta_5^{BU}(n_f/cashfl) - \beta_4^{BU}\gamma_4^L]$, $\gamma_5^{BU} = (1/C)[\beta_{10}^{BU}(n_u^{PF}/dmpa) + \beta_4^{BU}\gamma_5^L]$, $\gamma_6^{BU} = (1/C)\beta_6^{BU}(n_r/dmcash)$, $\gamma_9^{BU} = (1/C)(\beta_1^{BU} + \beta_4^{BU}\gamma_1^L)$, $\gamma_{11}^{BU} = (1/C)(\beta_1^{BU} + \beta_4^{BU}\gamma_1^L)$, and $\gamma_3^{BU} = -(1/C)(\beta_{21}^{BU} + \beta_4^{BU}\gamma_{32}^L)$.

From equation (3.33), we can derive the following expression for I^{BU}

(3.182) $$I^{BU} = dBU + EDEP^{BU}$$

Inserting equation (3.182) into (3.181) we get the following economic relationship for I^{BU}

(3.183) $$I^{BU} = \gamma_0^{BU} + \gamma_1^{BU} I^{MA} + \gamma_3^{BU} dca + \gamma_4^{BU} dcashfl + \gamma_5^{BU} ddmpa + \gamma_6^{BU} ddmacsh + \gamma_9^{BU} EDEP^{MA} + \gamma_{10}^{BU} EDEP^{BU} + \gamma_{11}^{BU} SMA + \gamma_{32}^{BU} dCL$$

where $\gamma_{10}^{BU} = 1$.

3) The change in other fixed assets: Equation (3.144) can be rewritten as follows

(3.184) $$dofa = \gamma_0^{OFA} + \gamma_4^{MA} dcashfl + \gamma_5^{MA} ddmpa + \gamma_6^{MA} ddmcash$$

where $D = \beta_{13}^{OFA}(n_{OFA}/OFA)$, $\gamma_0^{OFA} = -(1/D)\beta_0^{OFA}$, $\gamma_4^{OFA} = -(1/D)\beta_5^{OFA}(n_f/cashfl)$, $\gamma_5^{OFA} = -(1/D)\beta_{10}^{OFA}(n_u^{PF}/dmpa)$, and $\gamma_6^{OFA} = -(1/D)\beta_6^{OFA}(n_r/dmcash)$.

4) The change in current assets: Equation (3.145) can be rewritten as follows

(3.185) $$dca = \gamma_0^{CA} + \gamma_1^{CA} I^{MA} + \gamma_2^{CA} I^{BU} + \gamma_4^{CA} dcashfl + \gamma_5^{CA} ddmpa + \gamma_6^{CA} ddmcash + \gamma_9^{CA} EDEP^{MA} + \gamma_{10}^{CA} EDEP^{BU} + \gamma_{11}^{CA} SMA + \gamma_{32}^{CL} dCL$$

where $E = \beta_1^{CA} + \beta_4^{CA}\gamma_1^L - \beta_8^{CA}(n_{CA}/CA)$, $\gamma_0^{CA} = (1/E)(\beta_0^{CA} - \beta_4^{CA}\gamma_0^L)$, $\gamma_1^{CA} = -(1/E)(\beta_1^{CA} + \beta_4^{CA}\gamma_1^L)$, $\gamma_2^{CA} = -(1/E)(\beta_2^{CA} + \beta_4^{CA}\gamma_2^L)$, $\gamma_4^{CA} = (1/E)[\beta_5^{CA}(n_f/cashfl) - \beta_4^{CA}\gamma_4^L]$, $\gamma_5^{CA} = (1/E)[\beta_{10}^{CA}(n_u^{PF}/dmpa) - \beta_4^{CA}\gamma_5^L]$, $\gamma_6^{CA} = (1/E)\beta_6^{CA}(n_r/dmcash)$, $\gamma_9^{CA} = (1/E)(\beta_1^{CA} + \beta_4^{CA}\gamma_1^L)$, $\gamma_{10}^{CA} = (1/E)(\beta_2^{CA} + \beta_4^{CA}\gamma_2^L)$, $\gamma_{11}^{CA} = (1/E)(\beta_1^{CA} + \beta_4^{CA}\gamma_1^L)$, and $\gamma_{32}^{CA} = -(1/E)(\beta_{21}^{CA} + \beta_4^{CA}\gamma_{22}^L)$.

5) The change in share capital: Equation (3.146) can be rewritten as follows

(3.186) $dsc = \gamma_0^{sc} + \gamma_4^{SC} dcashfl + \gamma_6^{SC} ddmcash$

where $F = \beta_{15}^{SC}(n_{SC} / SC)$, $\gamma_0^{CA} = -(1/E)\beta_0^{SC}$, $\gamma_4^{SC} = -(1/E)\beta_5^{SC}(n_f / cashfl)$, and $\gamma_6^{SC} = -(1/E)\beta_6^{SC}(n_r / dmcash)$.

6) The change in restricted reserves: Equation (3.147) can be rewritten as follows

(3.187) $drr = \gamma_0^{RR} + \gamma_6^{RR} ddmcash$

where $G = \beta_{16}^{RR}(n_{RR} / RR)$, $\gamma_0^{RR} = -(1/G)\beta_0^{RR}$, and $\gamma_6^{RR} = -(1/G)\beta_6^{RR}(n_r / dmcash)$.

7) The change in other untaxed reserves: Equation (3.148) can be rewritten as follows

(3.188) $dour = \gamma_0^{OUR} + \gamma_4^{OUR} dcashfl + \gamma_5^{OUR} ddmpa + \gamma_6^{OUR} ddmcash$

where $I = \beta_{17}^{OUR}(n_{OUR} / OUR)$, $\gamma_0^{OUR} = -(1/I)\beta_0^{OUR}$, $\gamma_4^{OUR} = -(1/I)\beta_5^{OUR}(n_f / cashfl)$, $\gamma_5^{OUR} = -(1/I)\beta_{10}^{OUR}(n_u^{PF} / dmpa)$, and $\gamma_6^{OUR} = -(1/I)\beta_6^{OUR}(n_r / dmcash)$.

8) The allocations to periodical reserves: Equation (3.149) can be rewritten as follows

(3.189) $p^{allo} = \gamma_0^{PPF} + \gamma_4^{PPF} dcashfl + \gamma_5^{PPF} ddmpa + \gamma_6^{PPF} ddmcash + \gamma_{12}^{PPF} zpf$

where $J = \beta_{18}^{PPF}(n_{PF} / PF)$, $\gamma_0^{PPF} = -(1/J)\beta_0^{PPF}$, $\gamma_4^{PPF} = -(1/J)\beta_5^{PPF}(n_f / cashfl)$, $\gamma_5^{PPF} = -(1/J)\beta_{10}^{PPF}(n_u^{PF} / dmpa)$, $\gamma_6^{PPF} = -(1/J)\beta_6^{PPF}(n_r / dmcash)$, and $\gamma_{12}^{PPF} = 1$.

9) The tax depreciation of machinery and equipment: Equation (3.150) can be rewritten as follows

(3.190) $$\begin{aligned} TDEP^{MA} &= \gamma_0^{TMA} + \gamma_1^{TMA} I^{MA} + \gamma_4^{TMA} dcashfl + \gamma_5^{TMA} ddmpa + \gamma_6^{TMA} ddmcash \\ &\quad + \gamma_7^{TMAS} EDEP^{MA} + \gamma_{11}^{OFA} SMA \end{aligned}$$

where $K = \beta_8^{TMA}(n_{CMA} / CMA) - \beta_9^{TMA}(n_{ASD} / ASD)$, $\gamma_0^{TMA} = (1/K)\beta_0^{TMA}$,
$\gamma_1^{TMA} = (1/K)\beta_8^{TMA}(n_{CMA} / CMA)$, $\gamma_4^{TMA} = (1/K)\beta_5^{TMA}(n_f / cashfl)$,
$\gamma_5^{TMA} = (1/K)\beta_{10}^{TMA}(n_u^{PF} / dmpa)$, $\gamma_6^{TMA} = (1/K)\beta_6^{TMA}(n_r / dmcash)$,
$\gamma_7^{TMA} = -(1/K)\beta_9^{TMA}(n_{ASD} / ASD)$, and $\gamma_{11}^{TMA} = (1/K)\beta_6^{TMA}(n_{CMA} / CMA)$.

10) The removals from periodical reserves: Equation (3.151) can be rewritten as follows

$$(3.191)\quad zpf = \gamma_0^{ZPF} + \gamma_4^{ZPF} dcashfl + \gamma_5^{ZPF} ddmpa + \gamma_6^{ZPF} ddmcash + \gamma_{13}^{ZPF} p^{allo}$$

where $M = \beta_{18}^{ZPF}(n_{PF} / PF)$, $\gamma_0^{ZPF} = (1/M)\beta_0^{ZPF}$, $\gamma_4^{ZPF} = (1/M)\beta_5^{ZPF}(n_f / cashfl)$, $\gamma_5^{ZPF} = (1/M)\beta_{10}^{ZPF}(n_u^{PF} / dmpa)$, $\gamma_6^{ZPF} = (1/M)\beta_6^{ZPF}(n_r / dmcash)$, and $\gamma_{13}^{ZPF} = -1$.

11) The change in long-term liabilities: Equation (3.152) can be rewritten as follows

$$(3.192)\quad dll = \gamma_0^{LL} + \gamma_4^{LL} dcashfl + \gamma_5^{LL} ddmpa + \gamma_6^{LL} ddmcash$$

where $N = \beta_{19}^{LL}(n_{LL} / LL)$, $\gamma_0^{LL} = -(1/N)\beta_0^{LL}$, $\gamma_4^{LL} = -(1/N)\beta_5^{LL}(n_f / cashfl)$, $\gamma_5^{LL} = -(1/N)\beta_{10}^{LL}(n_u^{PF} / dmpa)$, and $\gamma_6^{LL} = -(1/N)\beta_6^{LL}(n_r / dmcash)$.

12) The change in current liabilities: Equation (3.153) can be rewritten as follows

$$(3.193)\quad \begin{aligned} dcl = {} & \gamma_0^{CL} + \gamma_1^{CL} I^{MA} + \gamma_2^{CL} I^{BU} + \gamma_3^{CL} dca + \gamma_4^{CL} dcashfl + \gamma_5^{CL} ddmpa + \gamma_6^{CL} ddmcash \\ & + \gamma_9^{CL} EDEP^{MA} + \gamma_{10}^{CL} EDEP^{BU} + \gamma_{11}^{CL} S^{MA} \end{aligned}$$

where $O = \beta_{21}^{CL} - \beta_{20}^{CL}(n_{CL} / CL) + \beta_4^{CL}\gamma_{32}^{L}$, $\gamma_0^{CL} = (1/O)[\beta_0^{CL} + \beta_4^{CL}\gamma_0^{L}]$, $\gamma_1^{CL} = -(1/O)[\beta_1^{CL} + \beta_4^{CL}\gamma_1^{L}]$, $\gamma_2^{CL} = -(1/O)[\beta_2^{CL} + \beta_4^{CL}\gamma_2^{L}]$, $\gamma_3^{CL} = -(1/O)[\beta_3^{CL} + \beta_4^{CL}\gamma_3^{L}]$, $\gamma_4^{CL} = -(1/O)[-\beta_4^{CL}\gamma_4^{L} + \beta_5^{CL}(n_f / cashfl)]$, $\gamma_5^{CL} = -(1/O)[-\beta_4^{CL}\gamma_5^{L} + \beta_{10}^{CL}(n_u^{PF} / dmpa)]$, $\gamma_6^{CL} = (1/O)\beta_6^{CL}(n_r / dmcash)$, $\gamma_9^{CL} = -\gamma_1^{L}$, $\gamma_{10}^{CL} = -\gamma_2^{L}$, and $\gamma_{11}^{CL} = -\gamma_1^{L}$.

The economic relationships (3.180) and (3.183)-(3.193) are the major relationships that should be estimated. The best way to estimate these relationships is to estimate them simultaneously. However, because of many different econometric problems (i.e. extreme values on dependent values- see further section 6.1), we have to rely on robust estimation techniques. These techniques are often very complicated. To use such techniques on a simultaneous equation system is not an easy task. Moreover, as we mentioned in the introduction to this book, for the simulation model it is important to use more or less "standardized estimation techniques". This makes the updating of the simulation model easier. Finally, as it is evident from the economic relationships (3.180) and (3.183)-(3.193), it is only a few dependent variables that are interrelated. We believe that estimating these relationships in a recursive manner would not jeopardize the findings of the simulation model. In the next section, we will outline the recursive system used to estimate each and every economic relationship in (3.180) and (3.183)-(3.193). There, we will also outline the economic relationships for those income statement variables that are related to the variables I^{MA}, I^{BU}, $dofa$, dca, dsc, drr, $dour$, p^{allo}, $TDEP^{MA}$, zpf, dll, and dcl.

3.5 The Econometric Models

In this section, we will present the econometric models for all the flow variables that are important for the simulation model. But, before we can do that, we must outline the economic relationships for those income statement variables that are related to the variables I^{MA}, I^{BU}, $dofa$, dca, dsc, drr, $dour$, p^{allo}, $TDEP^{MA}$, zpf, dll, and dcl. Moreover, the dynamic optimization model presented in sections 3.1-3.4 uses a continuous time approach. However, the simulation model has a discrete time approach. In this section, we will “translate” and show the findings in section 3.4 in discrete time.

3.5.1 The Economic Relationships for Different Income Statement Items

1) Economic depreciation and the sale of machinery and equipment: We know that $MA_t = MA_{t-1} + dMA$. By including equation (3.178) in this expression and using the definition of economic depreciation of machinery and equipment (from equation (3.12) we know that $EDEP_t^{MA} = \delta^{MA} MA_t$) and the definition for the sale of same assets (from equation (3.42) we know that $S_t^{MA} = sMA_t$), we can derive the following economic relationships for $EDEP_t^{MA}$ and S_t^{MA}.

(3.194)
$$\begin{aligned} EDEP^{MA} = {} & \gamma_0^{EMA} + \gamma_2^{EMA} I^{BU} + \gamma_3^{EMA} dca + \gamma_4^{EMA} dcashfl + \gamma_5^{EMA} ddmpa + \gamma_6^{EMA} ddmcash \\ & + \gamma_7^{EMA} ddmtdm + \gamma_8^{EMA} TDEP^{MA} + \gamma_{10}^{EMA} EDEP^{BU} + \gamma_{14}^{EMA} MA_{t-1} + \gamma_{32}^{EMA} dCL \end{aligned}$$

where $\gamma_0^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_0^{MA}$, $\gamma_2^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_2^{MA}$, $\gamma_3^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_3^{MA}$, $\gamma_4^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_4^{MA}$, $\gamma_5^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_5^{MA}$, $\gamma_6^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_6^{MA}$, $\gamma_7^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_7^{MA}$, $\gamma_8^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_8^{MA}$, $\gamma_{10}^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_{10}^{MA}$, $\gamma_{14}^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]$, and $\gamma_{32}^{EMA} = [\delta^{MA}/(1-\gamma_9^{MA})]\gamma_{32}^{MA}$.

(3.195)
$$\begin{aligned} S^{MA} = {} & \gamma_0^{SMA} + \gamma_2^{SMA} I^{BU} + \gamma_3^{SMA} dca + \gamma_4^{SMA} dcashfl + \gamma_5^{SMA} ddmpa + \gamma_6^{SMA} ddmcash \\ & + \gamma_7^{SMA} ddmtdm + \gamma_8^{SMA} TDEP^{MA} + \gamma_9^{SMA} EDEP^{MA} + \gamma_{10}^{SMA} EDEP^{BU} + \gamma_{14}^{SMA} MA_{t-1} \\ & + \gamma_{32}^{SMA} dCL \end{aligned}$$

where, $\gamma_0^{SMA} = s\gamma_0^{MA}$, $\gamma_2^{SMA} = s\gamma_2^{MA}$, $\gamma_3^{SMA} = s\gamma_3^{MA}$, $\gamma_4^{EMA} = s\gamma_4^{MA}$, $\gamma_5^{SMA} = s\gamma_5^{MA}$, $\gamma_6^{SMA} = s\gamma_6^{MA}$, $\gamma_7^{SMA} = s\gamma_7^{MA}$, $\gamma_8^{SMA} = s\gamma_8^{MA}$, $\gamma_9^{SMA} = s\gamma_9^{MA}$, $\gamma_{10}^{SMA} = s\gamma_{10}^{MA}$, $\gamma_{14}^{SMA} = s$, and $\gamma_{32}^{SMA} = s\gamma_{32}^{MA}$.

2) Economic depreciation and the tax depreciation of buildings: We know that $BU_t = BU_{t-1} + dBU$. By including equation (3.181) in this expression and using the definition of economic depreciation of buildings (from equation (3.13) we know that $EDEP_t^{BU} = \delta^{BU} BU_t$) and the definition of tax depreciation of same assets (from equation (3.24) we know that $TDEP_t^{BU} = \delta_T^{BU} BU_t$), we can derive the following economic relationships for $EDEP_t^{BU}$ and $TDEP_t^{BU}$.

$$EDEP_t^{MA} = \gamma_0^{EMA} + \gamma_2^{EMA} I^{MA} + \gamma_3^{EMA} dca + \gamma_4^{EMA} dcashfl_{t-1} + \gamma_5^{EMA} ddmpa_{t-1} + \gamma_6^{EMA} ddmcash_{t-1} + \gamma_7^{EMA} ddmtdm_{t-1} + \gamma_8^{EMA} TDEP^{MA} + \gamma_{10}^{EMA} EDEP^{BU} + \gamma_{14}^{EMA} MA_{t-1} + \gamma_{32}^{EMA} dCL \tag{3.196}$$

where, $\gamma_0^{EBU} = \delta^{BU} \gamma_0^{BU}$, $\gamma_1^{EBU} = \delta^{BU} \gamma_1^{BU}$, $\gamma_3^{EBU} = \delta^{BU} \gamma_3^{BU}$, $\gamma_4^{EBU} = \delta^{BU} \gamma_4^{BU}$, $\gamma_5^{EBU} = \delta^{BU} \gamma_5^{BU}$, $\gamma_6^{EBU} = \delta^{BU} \gamma_6^{BU}$, $\gamma_9^{EBU} = \delta^{BU} \gamma_9^{BU}$, $\gamma_{11}^{EBU} = \delta^{BU} \gamma_{11}^{BU}$, $\gamma_{15}^{EBU} = \delta^{BU}$, and $\gamma_{32}^{EBU} = \delta^{BU} \gamma_{32}^{BU}$.

$$TDEP^{BU} = \gamma_0^{TBU} + \gamma_1^{TBU} I^{MA} + \gamma_3^{TBU} dca + \gamma_4^{TBU} dcashfl + \gamma_5^{TBU} ddmpa + \gamma_6^{TBU} ddmcash + \gamma_9^{TBU} EDEP^{MA} + \gamma_{11}^{TBU} S^{MA} + \gamma_{15}^{TBU} BU_{t-1} + \gamma_{32}^{TBU} dCL \tag{3.197}$$

where $\gamma_0^{TBU} = \delta_T^{BU} \gamma_0^{BU}$, $\gamma_1^{TBU} = \delta_T^{BU} \gamma_1^{BU}$, $\gamma_3^{TBU} = \delta_T^{BU} \gamma_3^{BU}$, $\gamma_4^{TBU} = \delta_T^{BU} \gamma_4^{BU}$, $\gamma_5^{TBU} = \delta_T^{BU} \gamma_5^{BU}$, $\gamma_6^{TBU} = \delta_T^{BU} \gamma_6^{BU}$, $\gamma_9^{TBU} = \delta_T^{BU} \gamma_9^{BU}$, $\gamma_{11}^{TBU} = \delta_T^{BU} \gamma_{11}^{BU}$, $\gamma_{15}^{TBU} = \delta_T^{BU}$, and $\gamma_{32}^{TBU} = \delta_T^{BU} \gamma_{32}^{BU}$.

3) Operating income before depreciation: We know that $OIBD_t = f(MA_t, BU_t, WC_t, L) - WL$, $MA_t = MA_{t-1} + dMA_t$, $BU_t = BU_{t-1} + dBU_t$, $WC_t = WC_{t-1} + dwc_t$, and $L_t = L_{t-1} + dL_t$. By using (3.177), (3.178), (3.181), (3.185), and (3.193) we obtain the following expression for operating income before depreciation.

$$OIBD = \gamma_0^{OIBD} + \gamma_1^{OIBD} I^{MA} + \gamma_2^{OIBD} I^{BU} + \gamma_3^{OIBD} dca + \gamma_4^{OIBD} dcashfl + \gamma_5^{OIBD} ddmpa + \gamma_9^{OIBD} EDEP^{MA} + \gamma_{10}^{OIBD} EDEP^{BU} + \gamma_{11}^{OIBD} S^{MA} + \gamma_{14}^{OIBD} MA_{t-1} + \gamma_{15}^{OIBD} BU_{t-1} + \gamma_{16}^{OIBD} CA_{t-1} + \gamma_{32}^{OIBD} dCL + \gamma_{33}^{OIBD} dCL \tag{3.198}$$

where $\gamma_0^{OIBD} = \gamma_0^{OIBD}(\gamma_0^L, W)$, $\gamma_1^{OIBD} = \gamma_1^{OIBD}(\gamma_1^L, W)$, $\gamma_2^{OIBD} = \gamma_2^{OIBD}(\gamma_2^L, W)$, $\gamma_3^{OIBD} = \gamma_3^{OIBD}(\gamma_3^L, W)$, $\gamma_4^{OIBD} = \gamma_4^{OIBD}(\gamma_4^L, W)$, $\gamma_5^{OIBD} = \gamma_5^{OIBD}(\gamma_5^L, W)$, $\gamma_9^{OIBD} = \gamma_9^{OIBD}(\gamma_1^L, W)$, $\gamma_{10}^{OIBD} = \gamma_{10}^{OIBD}(\gamma_2^L, W)$, and $\gamma_{11}^{OIBD} = \gamma_{11}^{OIBD}(\gamma_1^L, W)$.

3) Financial income: We know that $FI_t = h(OFA_t + MA_t + BU_t + CA_t)$, and $OFA_t = OFA_{t-1} + dofa_t$. By using (3.178), (3.181), (3.184), and (3.185), we obtain the following expression for financial income.

$$FI = \gamma_0^{FI} + \gamma_1^{FI} I^{MA} + \gamma_2^{FI} I^{BU} + \gamma_3^{FI} dca + \gamma_9^{FI} EDEP^{MA} + \gamma_{10}^{FI} EDEP^{BU} + \gamma_{11}^{FI} S^{MA} + \gamma_{14}^{FI} MA_{t-1} + \gamma_{15}^{FI} BU_{t-1} + \gamma_{16}^{FI} CA_{t-1} + \gamma_{17}^{FI} OFA_{t-1} + \gamma_{31}^{FI} dofa_t \tag{3.199}$$

where $\gamma_1^{FI} = ... = \gamma_{31}^{fi} = h$. It is important to mention that this is true where financial income is modeled as a fraction of firms' capital stock.

4) Financial expenses: We know that $FE_t = i_B(CL_t + LL_t) + n(OFA_t + MA_t + BU_t + CA_t)$, $CL_t = CL_{t-1} + dcl_t$ and $LL_t = LL_{t-1} + dll_t$. By using (3.178), (3.181), (3.184), (3.185), (3.192), and (3.193), we obtain the following expression for financial expenses.

$$
\begin{aligned}
FE = {} & \gamma_0^{FE} + \gamma_1^{FE} I^{MA} + \gamma_2^{FE} I^{BU} + \gamma_3^{FE} dca + \gamma_9^{FE} EDEP^{MA} + \gamma_{10}^{FE} EDEP^{BU} \\
& + \gamma_{11}^{FE} S^{MA} + \gamma_{14}^{FE} MA_{t-1} + \gamma_{15}^{FE} BU_{t-1} + \gamma_{16}^{FE} CA_{t-1} + \gamma_{17}^{FE} OFA_{t-1} \\
& + \gamma_{18}^{FE} CL_{t-1} + \gamma_{19}^{FE} dcl_t + \gamma_{20}^{FE} LL_{t-1} + \gamma_{21}^{FE} dll_t + \gamma_{31}^{FE} dofa_t
\end{aligned} \tag{3.200}
$$

where $\gamma_1^{FE} = \gamma_{17}^{FE} = n$ and $\gamma_{18}^{FE} = ... = \gamma_{31}^{FE} = i_B$. It is important to mention that this is true where financial expenses are modeled as a fraction of both firms' capital stock and their liabilities.

5) Group contributions: We know that $GC_t = g(OIBD_t - EDEP_t^{BU} + FI_t - FE_t - TDEP_t^{MA} + zpf_t - dour_t)$ so that

$$
\begin{aligned}
GC = {} & \gamma_0^{GC} + \gamma_{10}^{GC} EDEP^{BU} + \gamma_{22}^{GC} OIBD + \gamma_{23}^{GC} FI + \gamma_{24}^{GC} FE + \gamma_{25}^{GC} TDEP^{MA} \\
& + \gamma_{26}^{GC} zpf + \gamma_{27}^{GC} dour
\end{aligned} \tag{3.201}
$$

where $\gamma_{10}^{GC} = \gamma_{27}^{GC} = g$. It is important to mention that this is true where group contributions are modeled as a fraction of net income before group contributions ($NIBGC_t$).

5) Other allocations: We know that $OA_t = -dour_t + GC_t$ so that

$$
OA = \gamma_0^{OA} + \gamma_{28}^{OA} dour + \gamma_{29}^{OA} GC \tag{3.202}
$$

where $\gamma_{28}^{OA} = -1$ and $\gamma_{29}^{OA} = 1$. It is important to mention that this is true where other allocations are modeled as a sum of group contributions and the change in other untaxed reserves.

6) Tax liabilities: We know that $TL_t = \alpha EBT_t = \alpha(1+g)(OIBD_t - EDEP_t^{BU} + FI_t - FE_t - TDEP_t^{MA} + zpf_t - dour_t) - \alpha p_t^{allo}$ so that

$$
\begin{aligned}
TL = {} & \gamma_0^{TL} + \gamma_{10}^{TL} EDEP^{BU} + \gamma_{13}^{TL} p^{allo} + \gamma_{22}^{TL} OIBD + \gamma_{23}^{TL} FI + \gamma_{24}^{TL} FE + \gamma_{25}^{TL} TDEP^{MA} \\
& + \gamma_{26}^{TL} zpf + \gamma_{27}^{TL} dour
\end{aligned} \tag{3.203}
$$

where $\gamma_{13}^{TL} = -\alpha$ and $\gamma_{10}^{TL} = \gamma_{27}^{TL} = \alpha(1+g)$.

7) Other tax adjustments: We know that $OTA_t = \gamma^{OT} NI_t = \gamma^{OT}(EBT_t - TL_t)$ so that

(3.204)
$$OTA = \gamma_0^{OTA} + \gamma_{10}^{OTA} EDEP^{BU} + \gamma_{13}^{OTA} p^{allo} + \gamma_{22}^{OTA} OIBD + \gamma_{23}^{OTA} FI + \gamma_{24}^{OTA} FE + \gamma_{25}^{OTA} TDEP^{MA} + \gamma_{26}^{OTA} zpf + \gamma_{27}^{OTA} dour + \gamma_{28}^{OTA} TL$$

where $\gamma_{13}^{OTA} = -\gamma^{OT}$, $\gamma_{10}^{OTA} = \gamma_{27}^{OTA} = \gamma^{OT}(1+g)$, and $\gamma_{28}^{OTA} = -1$. This, of course, is true where other tax adjustments are modeled as a fraction of net income. However, we could have modeled other tax adjustments as a function of NI_t. In that case, $\gamma_{10}^{OTA} - \gamma_{28}^{OTA}$ could take other values.

8) Reduction of taxes: We know that $ROT_t = \beta^{RO} TAX_t = \beta^{RO} \tau (EBT_t - TL_t + OTA_t - TDEP_t^{BU})$ so that

(3.205)
$$ROT = \gamma_0^{ROT} + \gamma_{10}^{ROT} EDEP^{BU} + \gamma_{13}^{ROT} p^{allo} + \gamma_{22}^{ROT} OIBD + \gamma_{23}^{ROT} FI + \gamma_{24}^{ROT} FE + \gamma_{28}^{ROT} TDEP^{MA} + \gamma_{12}^{ROT} zpf + \gamma_{25}^{ROT} dour + \gamma_{26}^{ROT} TL + \gamma_{27}^{ROT} OTA + \gamma_{28}^{ROT} TDEP^{BU}$$

where $\gamma_{13}^{ROT} = -\beta^{RO}\tau$, $\gamma_{10}^{ROT} = ... = \gamma_{25}^{ROT} = \gamma_{29}^{ROT} = \beta^{RO}\tau(1+g)$, and $\gamma_{26}^{ROT} = \gamma_{28}^{ROT} = -\beta^{RO}\tau$.

3.5.2 The Recursive System Used to Estimate Firms' Decision Variables

In this section, we will summarize the recursive system that we use in the next chapter to estimate each and every decision variable within a firm. Here, we also introduce the time index to make the recursive system more straightforward.

(3.206)
$$EDEP_t^{MA} = \gamma_0^{EMA} + \gamma_2^{EMA} I_{t-1}^{MA} + \gamma_3^{EMA} dca_{t-1} + \gamma_4^{EMA} dcashfl_{t-1} + \gamma_5^{EMA} ddmpa_{t-1} + \gamma_6^{EMA} ddmcash_{t-1} + \gamma_7^{EMA} ddmtdm_{t-1} + \gamma_8^{EMA} TDEP_{t-1}^{MA} + \gamma_{10}^{EMA} EDEP_{t-1}^{BU} + \gamma_{14}^{EMA} MA_{t-1} + \gamma_{32}^{EMA} dcl_{t-1}$$

(3.207)
$$S_t^{MA} = \gamma_0^{SMA} + \gamma_2^{SMA} I_{t-1}^{BU} + \gamma_3^{SMA} dca_{t-1} + \gamma_4^{SMA} dcashfl_{t-1} + \gamma_5^{SMA} ddmpa_{t-1} + \gamma_6^{SMA} ddmcash_{t-1} + \gamma_7^{SMA} ddmtdm_{t-1} + \gamma_8^{SMA} TDEP_{t-1}^{MA} + \gamma_9^{SMA} EDEP_t^{MA} + \gamma_{10}^{SMA} EDEP_{t-1}^{BU} + \gamma_{14}^{SMA} MA_{t-1} + \gamma_{32}^{SMA} dcl_{t-1}$$

(3.208)
$$I_t^{MA} = \gamma_0^{MA} + \gamma_2^{MA} I_{t-1}^{BU} + \gamma_3^{MA} dca_{t-1} + \gamma_4^{MA} dcashfl_{t-1} + \gamma_5^{MA} ddmpa_{t-1} + \gamma_6^{MA} ddmcash_{t-1} + \gamma_7^{MA} ddmtdm_{t-1} + \gamma_8^{MA} TDEP_{t-1}^{MA} + \gamma_9^{MA*} EDEP_t^{MA} + \gamma_{10}^{MA} EDEP_{t-1}^{BU} + \gamma_{11}^{MA} S_t^{MA} + \gamma_{32}^{MA} dcl_{t-1}$$

(3.209)
$$EDEP_t^{BU} = \gamma_0^{EBU} + \gamma_1^{EBU} I_t^{MA} + \gamma_3^{EBU} dca_{t-1} + \gamma_4^{EBU} dcashfl_{t-1} + \gamma_5^{EBU} ddmpa_{t-1} + \gamma_6^{EBU} ddmcash_{t-1} + \gamma_9^{EBU} EDEP_t^{MA} + \gamma_{11}^{EBU} S_t^{MA} + \gamma_{15}^{EBU} BU_{t-1} + \gamma_{32}^{EBU} dcl_{t-1}$$

(3.210)
$$I_t^{BU} = \gamma_0^{BU} + \gamma_1^{BU} I_t^{MA} + \gamma_3^{BU} dca_{t-1} + \gamma_4^{BU} dcashfl_{t-1} + \gamma_5^{BU} ddmpa_{t-1} + \gamma_6^{BU} ddmacsh_{t-1} + \gamma_9^{BU} EDEP_t^{MA} + \gamma_{10}^{BU} EDEP_t^{BU} + \gamma_{11}^{BU} S_t^{MA} + \gamma_{32}^{BU} dcl_{t-1}$$

(3.211) $$dofa_t = \gamma_0^{OFA} + \gamma_4^{MA} dcashfl_{t-1} + \gamma_5^{MA} ddmpa_{t-1} + \gamma_6^{MA} ddmcash_{t-1}$$

(3.212) $$\begin{aligned} dca_t = {} & \gamma_0^{CA} + \gamma_1^{CA} I_t^{MA} + \gamma_2^{CA} I_t^{BU} + \gamma_4^{CA} dcashfl_{t-1} + \gamma_5^{CA} ddmpa_{t-1} + \gamma_6^{CA} ddmcash_{t-1} \\ & + \gamma_9^{CA} EDEP_t^{MA} + \gamma_{10}^{CA} EDEP_t^{BU} + \gamma_{11}^{CA} S_t^{MA} + \gamma_{32}^{CA} dcl_{t-1} \end{aligned}$$

(3.213) $$dll_t = \gamma_0^{LL} + \gamma_4^{LL} dcashfl_{t-1} + \gamma_5^{LL} ddmpa_{t-1} + \gamma_6^{LL} ddmcash_{t-1}$$

(3.214) $$\begin{aligned} dcl_t = {} & \gamma_0^{CL} + \gamma_1^{CL} I_t^{MA} + \gamma_2^{CL} I_t^{BU} + \gamma_3^{CL} dca_t + \gamma_4^{CL} dcashfl_{t-1} + \gamma_5^{CL} ddmpa_{t-1} + \gamma_6^{CL} ddmcash_{t-1} \\ & + \gamma_9^{CL} EDEP_t^{MA} + \gamma_{10}^{CL} EDEP_t^{BU} + \gamma_{11}^{CL} S_t^{MA} \end{aligned}$$

(3.215) $$dsc_t = \gamma_0^{sc} + \gamma_4^{SC} dcashfl_{t-1} + \gamma_6^{SC} ddmcash_{t-1}$$

(3.216) $$drr_t = \gamma_0^{RR} + \gamma_6^{RR} ddmcash_{t-1}$$

(3.217) $$\begin{aligned} OIBD_t = {} & \gamma_0^{OIBD} + \gamma_1^{OIBD} I_t^{MA} + \gamma_2^{OIBD} I_t^{BU} + \gamma_3^{OIBD} dca_t + \gamma_4^{OIBD} dcashfl_{t-1} + \gamma_5^{OIBD} ddmpa_{t-1} \\ & + \gamma_9^{OIBD} EDEP_t^{MA} + \gamma_{10}^{OIBD} EDEP_t^{BU} + \gamma_{11}^{OIBD} S_t^{MA} + \gamma_{14}^{OIBD} MA_{t-1} + \gamma_{15}^{OIBD} BU_{t-1} \\ & + \gamma_{16}^{OIBD} CA_{t-1} + \gamma_{32}^{OIBD} CL_{t-1} + \gamma_{33}^{OIBD} dcl_t \end{aligned}$$

(3.218) $$\begin{aligned} FI_t = {} & \gamma_0^{FI} + \gamma_1^{FI} I_t^{MA} + \gamma_2^{FI} I_t^{BU} + \gamma_3^{FI} dca_t + \gamma_9^{FI} EDEP_t^{MA} + \gamma_{10}^{FI} EDEP_t^{BU} \\ & + \gamma_{11}^{FI} S_t^{MA} + \gamma_{14}^{FI} MA_{t-1} + \gamma_{15}^{FI} BU_{t-1} + \gamma_{16}^{FI} CA_{t-1} + \gamma_{17}^{FI} OFA_{t-1} + \gamma_{31}^{FI} dofa_t \end{aligned}$$

(3.219) $$\begin{aligned} FE_t = {} & \gamma_0^{FE} + \gamma_1^{FE} I_t^{MA} + \gamma_2^{FE} I_t^{BU} + \gamma_3^{FE} dca_t + \gamma_9^{FE} EDEP_t^{MA} + \gamma_{10}^{FE} EDEP_t^{BU} \\ & + \gamma_{11}^{FE} S_t^{MA} + \gamma_{14}^{FE} MA_{t-1} + \gamma_{15}^{FE} BU_{t-1} + \gamma_{16}^{FE} CA_{t-1} + \gamma_{17}^{FE} OFA_{t-1} \\ & + \gamma_{18}^{FE} CL_{t-1} + \gamma_{19}^{FE} dcl_t + \gamma_{20}^{FE} LL_{t-1} + \gamma_{21}^{FE} dll_t + \gamma_{32}^{FE} dofa_t \end{aligned}$$

(3.220) $$\begin{aligned} TDEP_t^{MA} = {} & \gamma_0^{TMA} + \gamma_1^{TMA} I_t^{MA} + \gamma_4^{TMA} dcashfl_{t-1} + \gamma_5^{TMA} ddmpa_{t-1} + \gamma_6^{TMA} ddmcash_{t-1} \\ & + \gamma_7^{TMAS} EDEP_t^{MA} + \gamma_{11}^{OFA} S_t^{MA} \end{aligned}$$

(3.221) $$zpf_t = \gamma_0^{ZPF} + \gamma_4^{ZPF} dcashfl_{t-1} + \gamma_5^{ZPF} ddmpa_{t-1} + \gamma_6^{ZPF} ddmcash_{t-1} + \gamma_{13}^{ZPF} p_{t-1}^{allo}$$

(3.222) $$dour_t = \gamma_0^{OUR} + \gamma_4^{OUR} dcashfl_{t-1} + \gamma_5^{OUR} ddmpa_{t-1} + \gamma_6^{OUR} ddmcash_{t-1}$$

(3.223) $$\begin{aligned} GC_t = {} & \gamma_0^{GC} + \gamma_{10}^{GC} EDEP_t^{BU} + \gamma_{22}^{GC} OIBD_t + \gamma_{23}^{GC} FI_t + \gamma_{24}^{GC} FE_t + \gamma_{25}^{GC} TDEP_t^{MA} \\ & + \gamma_{26}^{GC} zpf_t + \gamma_{27}^{GC} dour_t \end{aligned}$$

(3.224) $$OA_t = \gamma_0^{OA} + \gamma_{28}^{OA} dour_t + \gamma_{29}^{OA} GC_t$$

(3.225) $$\begin{aligned} TL_t = {} & \gamma_0^{TL} + \gamma_{10}^{TL} EDEP_t^{BU} + \gamma_{13}^{TL} p_{t-1}^{allo} + \gamma_{22}^{TL} OIBD_t + \gamma_{23}^{TL} FI_t + \gamma_{24}^{TL} FE_t + \gamma_{25}^{TL} TDEP_t^{MA} \\ & + \gamma_{26}^{TL} zpf_t + \gamma_{27}^{TL} dour_t \end{aligned}$$

(3.226) $$\begin{aligned} OTA_t = {} & \gamma_0^{OTA} + \gamma_{10}^{OTA} EDEP_t^{BU} + \gamma_{13}^{OTA} p_t^{allo} + \gamma_{22}^{OTA} OIBD_t + \gamma_{23}^{OTA} FI_t + \gamma_{24}^{OTA} FE_t \\ & + \gamma_{25}^{OTA} TDEP_t^{MA} + \gamma_{26}^{OTA} zpf + \gamma_{27}^{OTA} dour_t + \gamma_{28}^{OTA} TL_t \end{aligned}$$

$$\text{(3.227)}\quad \begin{aligned} TDEP_t^{BU} &= \gamma_0^{TBU} + \gamma_1^{TBU} I_t^{MA} + \gamma_3^{TBU} dca_t + \gamma_4^{TBU} dcashfl_{t-1} + \gamma_5^{TBU} ddmpa_{t-1} \\ &+ \gamma_6^{TBU} ddmcash_{t-1} + \gamma_9^{TBU} EDEP_t^{MA} + \gamma_{11}^{TBU} S_t^{MA} + \gamma_{15}^{TBU} BU_{t-1} + \gamma_{33}^{TBU} dcl_t \end{aligned}$$

The next variable in the recursive system is firms' allocations to periodical reserves: The equation of interest has the following structure

$$\text{(3.228)}\quad p_t^{allo} = \gamma_0^{PPF} + \gamma_4^{PPF} dcashfl_{t-1} + \gamma_5^{PPF} ddmpa_t + \gamma_6^{PPF} ddmcash_{t-1} + \gamma_{12}^{PPF} zpf_t$$

where $ddmap_t = (MPA_t - p_t^{allo}) - (MPA_{t-1} - p_{t-1}^{allo})$. Solving equation (3.228) for p_t^{allo} gives us the following economic relationship for p_t^{allo}

$$\text{(3.229)}\quad p_t^{allo} = \hat{\gamma}_0^{PPF} + \hat{\gamma}_4^{PPF} dcashfl_{t-1} + \gamma_{33}^{PPF} MPA_t + \gamma_{34}^{PPF} dmpa_{t-1} + \hat{\gamma}_6^{PPF} ddmcash_{t-1} + \hat{\gamma}_{12}^{PPF} zpf_t$$

where $\hat{\gamma}_0^{PPF} = \gamma_0^{PPF}/(1+\gamma_5^{PPF})$, $\hat{\gamma}_4^{PPF} = \gamma_4^{PPF}/(1+\gamma_5^{PPF})$, $\hat{\gamma}_{33}^{PPF} = \gamma_5^{PPF}/(1+\gamma_5^{PPF})$, $\hat{\gamma}_{34}^{PPF} = -\gamma_5^{PPF}/(1+\gamma_5^{PPF})$, $\hat{\gamma}_6^{PPF} = \gamma_6^{PPF}/(1+\gamma_5^{PPF})$, and $\hat{\gamma}_{12}^{PPF} = \gamma_{12}^{PPF}/(1+\gamma_5^{PPF})$.

Finally, we have the following economic relationship for ROT_t

$$\text{(3.230)}\quad \begin{aligned} ROT_t &= \gamma_0^{ROT} + \gamma_{10}^{ROT} EDEP_t^{BU} + \gamma_{13}^{ROT} p_t^{allo} + \gamma_{22}^{ROT} OIBD_t + \gamma_{23}^{ROT} FI_t + \gamma_{24}^{ROT} FE_t \\ &+ \gamma_8^{ROT} TDEP_t^{MA} + \gamma_{12}^{ROT} zpf_t + \gamma_{25}^{ROT} dour_t + \gamma_{26}^{ROT} TL_t + \gamma_{27}^{ROT} OTA_t + \gamma_{28}^{ROT} TDEP_t^{BU} \end{aligned}$$

To be able to learn something about the unknown parameters in our economic models (3.206)-(3.230), we need to specify corresponding *statistical models* that are very specific about the sampling process by which we think the data (the sample observations) were generated. In statistical models we must realize that economic relationships are not exact, containing both a predictable systematic component and an unobserved and unpredictable random error component, u. Therefore we add a disturbance term to each and every equation above. Giving the errors a random interpretation converts our economic models into statistical probability models and gives us a basis for statistical inference, that is, a basis for estimating unknown parameters and testing hypotheses about them.

Because of the interdependence between the stochastic disturbance and the endogenous explanatory variables, it is inappropriate to use OLS to estimate any equation in a system of simultaneous equations. The estimators in such a case are not only biased, but also inconsistent.

On the other hand, we estimate our model in a recursive way. As can be seen from the set of the equations, they will all be estimated one-way, with no feedback looping. A recursive model is a situation where OLS can be applied appropriately even in the context of simultaneous equation systems – OLS can be applied to each equation separately. Moreover, recursive models are never under-identified. Crucial for this model, however, is the fact that it is assumed that disturbance terms for the endogenous variables are uncorrelated, that is, the assumption of zero contemporaneous correlation must be fulfilled. By applying OLS in a recursive model we obtain consistent estimates.

Meanwhile, we may improve our estimation technique and construct estimators – maximum likelihood estimators – that are asymptotically efficient under additional assumptions. In this case, even if the assumption of no correlation among error terms is not satisfied, we still acquire consistent estimates, and there is thus no need for recursivity assumptions. Though preferable, we find it a rather complicated task to apply this kind of estimators and will leave it as a suggestion for future improvement of the model estimation[43].

Klevmarken (2001) argues that, in practice, even the maximum likelihood technique will meet several barriers and becomes unfeasible. In order to overcome certain obstacles, the Simulated Method of Moments (SMM) is suggested and a discussion of an alternative approach, namely the “moment calibrated” estimator, is in place. However, some indications of unfulfilled assumptions in our data make us rather cautious in using these methods and we will leave it as a future exercise, reinforced by the fact that these approaches are computationally demanding.

[43] When disturbances are correlated across equations, it may be advisable to consider the Seemingly Unrelated Regression (SURE) approach to cope with this assumption. Still, this method will not be applied to our model.

4 Description of Data

The data used in this paper are accounting data. The reporting of performance is the principal use commonly made of accounting information. It measures past success of the organization. It performs a policing action to ensure honesty, and also acts as a basis for rewards - both to managers and to stockholders. To the extent that accounting information actually is used in making management decisions, it should be valuable to use these data in a simulation model. The characteristics of accounting information can impart important dynamic characteristics to a business. We can anticipate that the psychological attitudes derived from the financial information have the potential for influencing decisions in such a way as to reinforce other difficulties to which the organization may be subject. It can be argued that the accounting system is an essential part of internal information loops affecting attitudes and decisions. Usually it measures symptoms, not causes.

It is important to pay close attention to the data being used when modeling firms' behavior regarding the flow (decision) variables. The model can only be as good as the data used. Statistical theory tells us that the data must contain relevant information, in order to capture the true behavior. Thus, if the data contain errors or show signs of selective information, the results will become incorrect. In this section we will describe our data, and briefly mention the methods used to improve the contained information.

4.1 The Source of the Database

The data at our disposal come mainly from the National Tax Board (RSV) and include the Standardized Accounting Statements (SRU) and the Tax Assessment (TA). The SRU contains information on accounts and tax adjustments and the TA contains information on the income tax paid by the firms.

RSV collects information, from individuals and companies, which is necessary for the assessment of tax. The TA files contain pure information on assessed income, preliminary tax, final tax, and some administrative data. However, the TA data do not contain any background information on how the final tax is calculated. Thus, before 1992, the local tax offices (LOK) gathered the necessary background information, in order to calculate the final tax for the companies. This information stayed within the local tax office. There were no obligations to pass on this vital information to the RSV. Therefore, up to 1992 the TA files where the only available information from the RSV.

In 1992, RSV introduced the so-called Standardized Account Statements (SRU). The SRU are the result of a long process in which the aim of the tax authorities has been to minimize the amount of information that firms are obliged to gather for their tax statements. The SRU information has helped the different LOK to perform tax assessments. It is also the underlying information for the tax authorities in their system for tax audits. All firms are obliged to report their SRU information to the LOK. They, in turn, are also obliged to send the information to the RSV. Thus, since 1992 there exists data on assessed income, preliminary tax, and final tax as well as the background information (accounts, balance sheet and tax adjustments) on how to calculate final tax. Through the progress of computerized accounting systems, approximately 44 percent of the companies reported their 1994 SRU file electronically. The remaining part was manually converted into data files at the RSV. The share of electronically reported SRU files has been increased since then. Thus, since 1992 two data files are available from the tax authorities.

Both the TA and the SRU files are designed to cover the total population of firms. This is true in the case of the TA files but not for the SRU. When starting a new database there are always problems associated with gathering the information. The local tax offices have had difficulties in passing on the SRU information (and any changes during tax assessment) to RSV. This has particularly been the case for the larger metropolitan areas (see Andersson 1996). Since the larger firms are concentrated in these areas, the implication of underreporting is a selective sample in a geographical sense. However, there were further selectivity problems with regard to company size. This is the main reason why we have chosen to have 1997 as a base year for our database.

In 1997, the Ministry of Finance and Statistics Sweden (SCB) started developing FRIDA, which stands for Firm Register and Individual DAtabases. FRIDA is composed of several databases for firms with different organization forms. This includes databases for joint stock companies (which also includes closed companies), cooperatives, partnerships (which also includes limited partnerships), associations, foundations, and proprietorships (or sole traders). Apart from partnerships and proprietorships, these enterprises are subject to corporation tax.

In this paper, we will present the database for stock companies only. The gathered information for these firms includes accounts, balance sheets, wages and other compensation, depreciation, untaxed reserves and dividends, etc. Moreover, it also includes information on tax adjustments.

4.2 The Sampling Construction

The sampling frame for the database is based on register data in TA and SRU. For the joint stock companies, we select those firms that provide the income tax return form S2. This register is then completed with further information about the organization form from SCB's Central Register of Enterprises and Establishments. The sampling frame is then adjusted by removing the income tax return form for those firms that have provided two identical forms.

The stratification is made according to company size and whether they are a closed company. For this purpose, we select those firms that complete their income tax return forms with another form (K10) that is used by shareholders in closed companies. The firms' size is based on total assets (K), net income (NI), and net business income (NBI). The sampling frame is stratified in three different strata.

The first stratum contains each and every financial firm that has the industry classification 65, 66, and 67 according to SNI92.[44] Further, this stratum also contains all those firms that fulfill the following conditions: $K > 100$ MSEK, $NI > 5$ MSEK (which is the case for both positive and negative net income), and $NBI > 5$ MSEK (which is the case for both positive and negative net business income). Remaining firms are classified between two different strata depending on whether there is a K10 form assigned to the company.[45]

In the first stratum, all units are selected. In the other two strata, the numbers of units drawn are a function of NBI. For this purpose a simple random sampling (SRS) was used.[46] The idea behind the database is to get a fairly good approximation of the total net business income and final tax payments.

---

[44] The industries are classified according to the Swedish Standard Industrial Classification 1992 (SNI92). This classification standard is based on the classification used by Eurostat, NACE Rev. 1.

[45] Firms with few employees but with an unusually large balance sheet or accounts are, by means of this construction, big in size, and do not influence the results through an otherwise large sampling weight in the other two strata.

[46] Thus, from the sampling construction we see that for the larger firms we have in principle a panel. For small firms we have a random sample. We use a sampling construction that gives each firm a random number. In each stratum, we sort the firms according to the random number in ascending order. We then select a stratum sample of n units, and choose the first n firms. Thus, if no new firms have been established, so that the units in the strata are the same, there is a small chance of getting a panel for small firms as well.

Today we have databases for 1995, 1996, 1997, 1998, and 1999. In our estimations, we use the information from 1997-1999. For these years, we have two time series observations on different variables for each and every firm. For 1997-1999, we have 27370, 27440, and 35457 cross-sectional observations respectively. This gives us a total of 90267 pooled observations.

From the sampling frame of 256171 companies in 1997, a sample of 27370 was drawn. The sample size was originally 33887. However, this sample was checked for inconsistencies and errors, etc.[47] Observations that do not fulfill the constructed criteria were excluded from the sample. Consequently, we end up with a sample size of 29363. Moreover, for the estimation purposes, we also need two time series observations on different variables. Therefore, we exclude those firms that did not provide information the previous year. By this we end up with a sample size of 27370. The sample was then re-weighted.[48]

A description of the sample for 1997

Strata	*Sampling frame*	*Sample size*	*Sample size after correction***	*Sample size with the observations for1995*
1. Large companies*	13291	13291	11258	10216
2. Small closed companies	134769	11129	10527	10103
3. Small joint stock companies	108111	9467	7578	7051
Total	256171	33887	29363	27370

* Companies (joint stock companies or closed companies) are regarded as large if their total assets are greater than 100 MSEK, their net income is greater than 5 MSEK, or their net business income is greater than 5 MSEK.
** After the sampling, we check for inconsistencies and errors, etc. If the firm does not fulfill the constructed criteria, it is regarded as an outlier and excluded from the sample. The sample is then re-weighted.

From the sampling frame of 250058 companies in 1998, a sample of 27400 was drawn. After correction, we end up with a sample size of 31400. By excluding those firms that do not have two time series observations on different variables, we end up with a sample size of 27400, which is then re-weighted.

A description of the sample for 1998

Strata	*Sampling frame*	*Sample size*	*Sample size after correction*	*Sample size with the observations for1996*
1. Large companies	15406	15406	13900	11700
2. Small closed companies	134726	11310	10700	9800
3. Small joint stock companies	99936	8391	6800	5900
Total	250058	35107	31400	27400

From the sampling frame of 243131 companies in 1999, a sample of 35457 was drawn.

[47] For a closer description, see section 4.3.
[48] It is important to note that the original sample, as well as the sample after correction, has unique weights that can be used to analyze the behavior of the population.

A description of the sample for 1999

Strata	*Sampling frame*	*Sample size*	*Sample size after correction*	*Sample size with the observations for1997*
1. Small financial stock companies	4575	4575	4575	4518
2. Large financial stock companies	705	705	705	698
3. Small companies that are not stock companies	482	482	482	482
4. Large companies that are not stock companies	118	118	118	118
5 Small closed companies	135476	11960	11868	11661
6. Small joint stock companies	91089	8040	7955	7775
7. Large stock companies*	10686	10686	10535	10205
Total	243131	36566	36238	35457

4.3 Check for Inconsistencies and Errors

As we mentioned in the previous section, the original sample for each and every year is checked for inconsistencies and errors, etc. The data program for audition and correction contains 30-60 modules. The structure of audition and correction is as follows:

First, the program starts by controlling whether the observations lack balance sheet data, income statement data, or tax adjustments. For 1996, approximately 50 percent of the observations that were regarded as outliers lacked such information. Although the firms are under statutory obligation to supply the data, non-response occurs. Some observations lack all the information. The reason for this could be that some of the firms do not provide the income tax return form S2. We also suspect that some of the firms are so-called dormant corporations. We do not use imputation methods to handle the non-response. Instead, we regard these firms as outliers and exclude them from our sample. There are no major selectivity problems in the data. Thus, we can conclude that the non-response can be considered random.

Second, before being used, the information collected for the observations that have information about the balance sheet, income statement, or tax adjustments, undergoes a detailed examination by the program developed. Routines of testing and improving the data quality have been developed to make the SRU and TA files reliable. Usually the errors originate from the following: clerical or typing errors, summation errors, or changes in the assessment of tax that is not registered in SRU files. In each module, we check whether firms have made a correct addition of the information requested by the tax authorities. If the deviation is lower than 100 SEK, we accept the addition made by firms. If not, we correct the information. If such correction is made, the firm is given a code that shows the way in which the information was corrected. All corrections are made automatically to avoid costly revision. In the data program for 1996, there were 800 different controls, which made it possible to correct the information provided by the firms. These controls corrected 85 percent of all inconsistencies and errors found in our database for 1996. The information for those observations that could not be corrected was assigned a special error code. Different parts of the income tax return form S2 are dependent on each other. This necessitates an extensive check of different variables against each other. This method of checking is very restrictive and requires that the information be error-free. To make this method less restrictive, we created some help variables. If the check indicates that there is a small deviation between some variables, we let the help variable include this difference.

From the above we see that the data source given to us contains some problems. The SRU and TA information is still in its infancy, and entails problems as regards reporting the information to RSV. At the RSV, some routines for checking and improving the quality of the SRU and TA files have been developed. These have enhanced the SRU and we believe that in the future we can expect the SRU and TA data to improve further.

We regard those firms that do not pass through the error and correction program as outliers and exclude them from our sample. The sample is then re-weighted.

5 Different Variables Used in the Statistical Module

In Chapters 1 and 2, we examined the information content of the firms' basic financial statements. In section 5.1, we introduce different variables that capture the legal and accounting constraints. To capture the firms' expectations regarding the business cycle and the development of the market, we also introduce, in section 5.2, different macroeconomic variables. In section 5.3, we define two different variables as proxies for the maturity of firms. In section 5.4, we define a variable that captures the location of the firms. In section 5.5, we define a variable that identifies the market in which firms may have their business. Moreover, we also introduce a variable that captures the market share of these firms. Finally, in section 5.6 we introduce different dummy variables that control the decision made by firms.

5.1 Legal and Accounting Constraints on the Firms' Behavior

In what follows, we will explain the way we have chosen to handle various legal and institutional constraints that are imposed on the firms.

1) The constraints that the tax code in Sweden puts on the amount of tax depreciation, which firms may deduct from their taxable income: In section 2.6, we defined the variable $MTDM_t$, which captures the constraint that the tax code in Sweden puts on the amount of tax depreciation that firms may deduct from their taxable income. Using equation (2.53)-(2.55) we define a variable that is the difference between the maximum amount of tax depreciation and the depreciation for income tax purposes made by these firms ($TDEP_t^{MA}$)

(5.1) $dmtdm_t = MTDM_t - TDEP_t^{MA}$

This variable gives us the opportunity to analyze the impact of underutilization (or overutilization) of depreciation allowances on firms' investment and financial behavior. We use equation (5.1) to construct the variable of interest in the economic relationships, $ddmtdm_t$, as follows

(5.2) $ddmtdm_t = dmtdm_t - dmtdm_{t-1} = (MTDM_t - TDEP_t^{MA}) - (MTDM_{t-1} - TDEP_{t-1}^{MA})$

This variable captures whether firms increase or decrease their utilization of depreciation allowances.

2) The constraint that the tax code in Sweden puts on the amount of allocations firms can make to periodical reserves each year: In section 2.7, we defined the variable MPA_t, which captures the maximum amount of allocations firms can make to periodical reserves. Using equation (2.56)-(2.58) we define a variable that is the difference between the maximum amount of allocations to periodical reserves and the allocation made by the firms during the current period (p_t^{allo})

(5.3) $dmpa_t = MPA_t - p_t^{allo}$

This variable gives us the opportunity to analyze the impact of underutilization (or overutilization) of allocations to periodical reserves on firms' investment and financial behavior. We use equation (5.3) to construct the variable of interest in the economic relationships, $ddmpa_t$, as follows

(5.4) $ddmpa_t = dmpa_t - dmpa_{t-1} = (MPA_t - p_t^{allo}) - (MPA_{t-1} - p_{t-1}^{allo})$

This variable captures whether firms change their utilization of their allocations to periodical reserves.

3) Dividends should be positive and they cannot exceed unrestricted equity in period t-1 plus the current period net business income minus allocations to restricted reserves. Using equation (2.41)-(2.42) we define a variable that is the difference between the maximum dividend firms can pay to their shareholders ($mcash_t$) and the amount of dividends they actually pay to their shareholders ($cashfl_t$)

(5.5) $dmcash_t = mcash_t - cashfl_t$

This variable gives us the opportunity to analyze the way in which firms' investment and financial behavior would be influenced by the fact that firms' dividend policy does not coincide with the legal constraint on dividends. We use equation (5.5) to construct the variable of interest in the economic relationships, $ddmcash_t$, as follows

(5.6) $ddmcash_t = dmcash_t - dmcash_{t-1} = (mcash_t - cashfl_t) - (mcash_{t-1} - cashfl_{t-1})$

This variable captures whether firms change their dividend policy so that it comes closer to the legal constraint on dividends.

5.2 Different Macroeconomic Variables

To capture firms' expectations regarding the business cycle and the development of the market, we also include one of the following two macroeconomic variables as explanatory variables in our estimations of the firms' decision variables: the change in gross national product ($dgnp_t$) and the real interest rate on a government bond with a maturity of 10 years ($ranta10_t$). These variables are defined as follows

(5.7) $dgnp_t = GNP_t - GNP_{t-1}$

(5.8) $realr_t = \dfrac{1 + ranta10_t}{1 + \inf_t} - 1 = \dfrac{ranta10_t - \inf_t}{1 + \inf_t}$

where $\inf_t$ is the inflation rate. The reason for using macroeconomic variables in our estimations is twofold. First, the simulation model describes only the behavior of the firms. All things outside this system are regarded as exogenous. However, these exogenous variables influence the flow variables in our closed system. The influence on behavior is captured by including exogenous factors as explanatory variables in our estimations of the flow variables. Second, the simulation model will be used as a practical tool at the Ministry of Finance. The forecasts of the model described in this book should be consistent with the macroeconomic forecasts at the Ministry of Finance.

5.3 The Maturity of the Firms

We believe that the maturity of the firms has an impact on their decision regarding investment and financing structure. As a proxy for the maturity of the firms we use the legal definition of “public” and “private” firms. “Private” firms are firms that have share capital of 100000 SEK or more in their possession. On the other hand, firms that have share capital of 500000 SEK or more in their possession are called “public” firms. We introduce a dummy variable that takes on values of 1 or 0, 1 indicating that the firm is “public” and 0 if the firm is “private”.

(5.9) $$\begin{aligned} Public_t &= 1 \quad \text{if} \quad bpptyp = 1 \\ &= 0 \quad \text{if} \quad bpptyp = 0 \end{aligned}$$

Further, a complementary proxy for the maturity of firms is whether or not these firms are closed companies. As we explained in the previous chapter, firms that complete their income tax return forms using another form (K10- which is used by shareholders in closed companies) are defined as closed firms. We introduce a dummy variable that takes on values of 1 or 0, 1 indicating that the firm is a closed company and 0 if the firm is not a closed company.

(5.10) $$\begin{aligned} FAAB_t &= 1 \quad \text{if} \quad K10 = 1 \\ &= 0 \quad \text{if} \quad K10 = 0 \end{aligned}$$

5.4 The Location of the Firms

Another factor that we believe has a major impact on firms’ behavior is their location. The competition that firms face is often dependent on where the firms are located. Another important factor is the firms’ financing ability. This is often dependent on the capital market that is close to the location of the firms. This is why we are interested to know whether these firms are located in larger localities, smaller localities or more sparsely populated areas. For this purpose, we introduce two dummy variables that take on values of 1 or 0 according to equation (5.11) and (5.12) below

(5.11) $$\begin{aligned} largcity_t &= 1 \quad if \quad the\ firm\ is\ located\ in\ larger\ city \\ &= 0 \quad if \quad otherwise \end{aligned}$$

(5.12) $$\begin{aligned} ruralare_t &= 1 \quad \text{if} \quad \text{the firm is located in a more sparsely populated area} \\ &= 0 \quad \text{if} \quad \text{otherwise} \end{aligned}$$

5.5 The Concentration of Incorporated Firms with the Industry

We believe that firms' behavior concerning their investment and financial structure is highly dependent of the concentration of incorporated firms within the industry in which they operate. This means that the behavior of firms that have a monopoly[49] in their market should differ from the behavior of firms operating in a duopoly, oligopoly, monopolistic competition, or competitive markets. The way we have chosen to identify the kind of market in which firms operate can be described as follows. First, we sort firms according to their industry classification in SNI92 at the three-digit level (there are 216 different industry classifications at the three-digit level). Second, we count the number of firms within each "market", which we call "*count*". Finally, we introduce the following variable

$$\textbf{(5.13)} \qquad market_{t-1} = \frac{1}{count}$$

A $market_{t-1} = 1$ indicates that the firm has a monopoly in its market. A $market_{t-1} = 0.5$ indicates that the firm operates in a duopoly market. As $market_{t-1} \to 0$, the competition within the market increases.

So far, we have identified markets in which the firms operate. However, we don't know the market share of the firms. The way we define the market share for each and every firm can be described as follows. First, we sum up the book value of assets within each "market", which we call "*msum*". Second, we define the market share as follows

$$\textbf{(5.14)} \qquad marketw_{t-1} = \frac{K_{t-1}}{msum}$$

where $msum_{t-1} = \sum_{j=1}^{i} K_{jt-1}$, $i = 1,...,470$. A $marketw_{t-1} = 1$ indicates that the firm has a 100 percent share of its market (this, of course, is the case for a monopoly). A $marketw_{t-1} = 0.5$ indicates that the firm has 50 percent of its market. As $marketw_{t-1} \to 0$, the firm's share of the market decreases.

5.6 Controlling for the Decisions Made by the Firms

As we showed in Chapter 3, we use a recursive system to estimate the behavior of the firms. The recursive system requires that we be aware of the firms' decisions during our estimation procedure. For example, the economic relationship for the change in other fixed assets does not include investment in machinery and equipment or investment in buildings. However, those firms that do not undertake such investment have better opportunities to invest in other fixed assets. Therefore, the behavior of these firms may be different compared to the behavior of firms who actually invest in machinery and buildings. Therefore, we introduce two dummy variables that take on values of 1 or 0 according to following equations

---

[49] This is not a monopoly in a traditional sense since we only look at the incorporated firms within an industry. For a traditional definition of a monopoly, we must also take into consideration that there exist other firms in the market of interest that have other organization forms. However, the industries here are classified according to SNI92 at the three-digit level. At such a level, we believe that firms organized in other forms should not have such an impact on the concentration within an industry.

$$(5.15)\quad \begin{aligned} DI_t^{MA} &= 1 &&\text{if} && I_t^{MA} > 0 \\ &= 0 &&\text{if} && I_t^{MA} = 0 \end{aligned}$$

$$(5.16)\quad \begin{aligned} DI_t^{BU} &= 1 &&\text{if} && I_t^{BU} \neq 0 \\ &= 0 &&\text{if} && I_t^{BU} = 0 \end{aligned}$$

We use this technique in our estimations of other economic relationships as well. To be able to do so, we introduce the following dummy variables

$$(5.17)\quad \begin{aligned} Ddofa_t &= 1 &&\text{if} && dofa_t \neq 0 \\ &= 0 &&\text{if} && dofa_t = 0 \end{aligned}$$

$$(5.18)\quad \begin{aligned} Ddll_t &= 1 &&\text{if} && dll_t \neq 0 \\ &= 0 &&\text{if} && dll_t = 0 \end{aligned}$$

$$(5.19)\quad \begin{aligned} Ddsc_t &= 1 &&\text{if} && dsc_t \neq 0 \\ &= 0 &&\text{if} && dsc_t = 0 \end{aligned}$$

$$(5.20)\quad \begin{aligned} DTDEP_t^{MA} &= 1 &&\text{if} && TDEP_t^{MA} > 0 \\ &= 0 &&\text{if} && TDEP_t^{MA} = 0 \end{aligned}$$

$$(5.21)\quad \begin{aligned} Dzpf_t &= 1 &&\text{if} && zpf_t > 0 \\ &= 0 &&\text{if} && zpf_t = 0 \end{aligned}$$

$$(5.22)\quad \begin{aligned} Ddour_t &= 1 &&\text{if} && dour_t \neq 0 \\ &= 0 &&\text{if} && dour_t = 0 \end{aligned}$$

6 The Statistical Module

In this chapter, we discuss different strategies and methods that we use in order to estimate certain relationships and concentrate on how to deal with different assumptions that underlie these relationships. In Appendix C, we outline the general structure and assumptions behind the classical regression models. However, our work shows that these assumptions are not fulfilled. The focus of this chapter will be on resistant estimation and methods for evaluating the impact of particular data elements on classical regression estimates.

6.1 Influence Diagnostics

Ordinary regression models are well known for their sensitivity to influential observations. Influential observations are anomalous values in the data and may have a strong influence over the fitted coefficients, giving a poor fit to the bulk of the data observations. In this step of the analysis we are merely interested in testing whether a number of data points is influential in the regression results and examine observations that strongly affect coefficient values (attempts to reduce their effect is discussed in section 6.1.3). We will test for influential observations associated with both the dependent variable and the regressors of the model (high-leverage points).

Influence diagnostics measure how each individual observation contributes to determining the parameter estimates and/or the fitted values. The influence of an observation depends on two factors: leverage, which is the "horizontal" distance of the *x*-value from the mean and *y*-discrepancy, being the vertical distance between y_i and a regression line that ignores y_i[51].

In our analysis we will consider three influence measures. Starting with high-leverage data points, we define the diagonal entries, h_i, of the Hat matrix, *H*, as measures of the (squared) distance of observation x_i from the vector of means, $\bar{x}$, or more formally $h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_k (x_k - \bar{x})^2}$. In practice, observations are typically considered high-leverage if $h_i > 2p/n$, where *n* is the number of observations used to fit the model and *p* is the number of parameters in the model.

In the presence of influential observations, the ordinary residuals, $e_i = y_i - \hat{y}_i$ are not very useful because the variance of the residuals, $var(e_i) = \sigma^2(1 - h_i)$, varies with leverage and aberrant on y tend to pull the regression line towards themselves. To correct these difficulties, a commonly used measure is the *studentized* (case-*deleted*) *residual,* RSTUDENT, which is the residual estimated for y_i without using that observation. Let $s^2_{(-i)}$ be the error mean square estimate of σ^2 omitting the *i:*th observation. Then, we define $RSTUDENT_i = \frac{e_i}{s_{(-i)}\sqrt{(1 - h_i)}}$.

Two influence measures based on the above ideas are commonly used. Both are functions of the product of leverage and *y*-discrepancy. The first is DFFITS and the

51 A conceptual formula is: Influence = Leverage × y-Discrepancy. Thus, an observation must have *both* high leverage and a large residual to be influential.

other, which is very similar to the latter and will thus not be used in the analysis, is COOKS'D. DFFITS is a scaled measure of the change in the predicted value, $\hat{y}_i$, when the ith observation is omitted, that is $DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(-i)}}{s_{(-i)}\sqrt{h_i}} = \frac{e_i}{s_{(-i)}}\sqrt{h_i/(1-h_i)}$. Large values of DFFITS indicate influential observations and a general cutoff point is 2. However, in large data sets, it is recommended to use $2\sqrt{p/n}$ as a cutoff with p and n defined as previously.

6.2 Robust Estimation

Least squares estimation enjoys a number of attractive properties when we are dealing with ideal data. If the error term is normally, independently and identically distributed, then least squares regression is the most efficient method in the class of unbiased estimators. However, this popular approach loses its advantages as complications, often associated with data, arise. OLS regression models are quite sensitive to influential observations, which may be a consequence of heavy-tailed distributions or simply are data points caused by coding errors. Careful examination of regression diagnostics, both in the direction of the dependent and explanatory variables, should thus be a part of any regression analysis.

When influential cases are present and have an unacceptable effect on the least squares method, the best remedy is to apply a robust estimator rather than deleting these observations either directly or indirectly by entering binary variables for anomalous data. Hampel (1985) argues that robust estimators are superior in practice to classical non-robust estimators and shows that they are even better than classical methods combined with rejection. Moreover, rejecting influential observations affects the sampling distribution of classical methods in ways that have not been fully studied, while large-sample distribution of most robust estimators can be easily calculated (Carroll and Ruppert (1987)).

If the influential observations are suspected to be gross errors, a robust estimator should be used (Krasker et al. (1983)). Meanwhile, Carroll and Ruppert (1987) suggest that a robust estimator is also useful when the influential cases indicate a model deficiency, but either the data do not allow a better model to be developed or the analyst does not want to add complexity to the existing model merely in order to accommodate a small fraction of observations.

After previously dealing with issues such as model specification, homoscedasticity and influence statistics, we concentrate here on robust methods as a modification of least squares regression.

The basic purpose of robust statistics is to develop procedures that perform well when the model to be estimated is correctly specified, while being relatively insensitive to small departures from the distributional assumptions of the model.

Three basic types of estimators – *M*-estimators (maximum-likelihood type estimators), *L*-estimators (linear combinations of order statistics) and *R*-estimators (derived from rank tests) – have played an important role in the topic of robust estimation. However, our emphasis will be on the first type, namely the *M*-estimators. This class of estimators is the most flexible one, involves relatively easy computations and generalizes straightforwardly to multi-parameter problems. Particularly, we are interested in the estimation of a location parameter. Although the family of *M*-estimators is large, in our analysis we will concentrate on one of them, namely the Huber's estimator. Usually, the squared-residual objective function makes the estimator too sensitive to outliers. At the same time, using absolute residuals makes the estimate too sensitive to the middle region. Huber (1964) proposed as a robust estimator the maximum-likelihood estimator of the location parameter associated with a density function that is like a normal one in the middle part, but like a double exponential in the tails. This study formed the first basis for a theory of robust estimation. In what follows we look upon some parts of the material.

The underlying density of the normal distribution is given as

$$(6.1)\quad f(z|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\, exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)$$

while the density of the double exponential distribution is

$$(6.2)\quad g(z|\mu,\sigma) = \frac{exp\left(-\left|\frac{z-\mu}{\delta}\right|\right)}{2\delta}$$

In fact, the Huber estimator is just a slight generalization of maximum-likelihood estimators for the *least favorable ε-contaminated* Gaussian distribution. Roughly speaking, ε is the amount of contamination that affects our distribution of interest. Assume that $\Re$ is the set of all ε-contaminated normal distribution, that is, the set of all distributions of the form $F = (1-\varepsilon)\Phi + \varepsilon H$, where $0 \le \varepsilon < 1$ is a fixed number, Φ is the standard normal (Gaussian) cumulative distribution and H is a variable distribution function over the set of all symmetric probability distributions (an unknown contaminating distribution). This model arises for instance if the observations are assumed to be normal with variance 1, but a fraction ε of them is affected by gross errors. A convenient measure of robustness for asymptotically normal estimators seems to be the supremum of the asymptotic variance ($n \to \infty$) when F ranges over some suitable set of underlying distributions, in particular over the set of all $F = (1-\varepsilon)\Phi + \varepsilon H$ for fixed ε and symmetric H. As mentioned, this *M*-estimator is the maximum-likelihood estimator corresponding to a unique least favorable distribution F_0 with density given as

$$(6.3)\quad f_0(z) = \frac{1-\varepsilon}{\sqrt{2\pi}}\, exp(-\rho_0(z))$$

This density behaves like a normal density for small z and like an exponential density for large z. The function $\rho(z) = -\ln f(z)$[52] is a robust non-constant loss function given by

[52] Including the class of all maximum-likelihood estimators where f is the density of the untranslated distribution. Two simple examples contain in particular the sample *mean* ($\rho(z) = z^2$) and the sample median ($\rho(z) = |z|$).

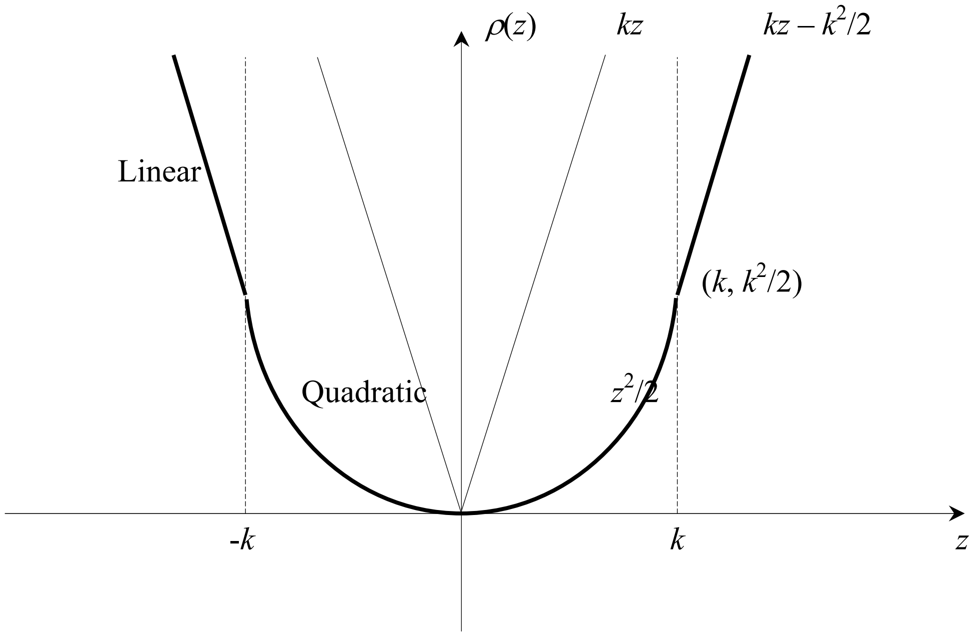
$$(6.4)\quad \rho(z)=\begin{cases}\frac{1}{2}z^2 & \text{if } |z|\le k\\ k|z|-\frac{1}{2}k^2 & \text{if } |z|>k\end{cases}$$

This objective function is drawn in the figure below for illustrative purposes
The parameter k is related to ε by

$$(6.5)\quad \frac{2}{k}\varphi(k)-2\Phi(-k)=\frac{\varepsilon}{1-\varepsilon}$$

where φ is the density function of the Φ distribution. This parameter can be seen as a constant that can be tuned in order to adjust the efficiency of the resulting estimator for specific distributions. Hence, the density of the least favorable distribution is given as

$$(6.6)\quad f(z)=\begin{cases}\frac{1-\varepsilon}{\sqrt{2\pi}}\,exp\left(-\frac{z^2}{2}\right) & \text{if } |z|\le k\\ \frac{1-\varepsilon}{\sqrt{2\pi}}\,exp\left(\frac{k^2}{2}-k|z|\right) & \text{if } |z|>k\end{cases}$$



An auxiliary scale estimator is essential for the data to fix the point of transition from quadratic to linear. Let e_i represent the ith-case absolute residual and define the ith scaled residual as

(6.7) $r_i = \frac{e_i}{\sigma} = \frac{(y_i - \mathbf{x_i\beta})}{\sigma}$

where $\boldsymbol{\beta}$ is a vector of parameter estimates and σ is a residual scale estimate. Assuming this residual to enjoy the above-mentioned Huber properties, i.e. the known parametric form of the error distribution is assumed to be contaminated by a distribution with probability epsilon, we may rewrite equation (6.6) as follows

(6.8) $$f(r_i) = \begin{cases} \frac{1-\varepsilon}{\sqrt{2\pi}}\, exp\left(-\frac{e_i^2}{2\sigma^2}\right) & \text{if } |r_i| \le k \\ \frac{1-\varepsilon}{\sqrt{2\pi}}\, exp\left(\frac{k^2}{2} - k\left|\frac{e_i}{\sigma}\right|\right) & \text{if } |r_i| > k \end{cases}$$

The log-likelihood function for a sample of n disturbances from the above distribution is given as $\ln L(\beta) = \sum_{i=1}^{n} \ln f(r_i)$ or as

(6.9) $$\ln L(\beta) = \begin{cases} \sum_{i=1}^{n} -\left[\text{ - ln}(1\text{-}\varepsilon) + \ln(\sqrt{2\pi}) + \frac{1}{2}\left(\frac{e_i}{\sigma}\right)^2\right] & \text{if } |r_i| \le k \\ \sum_{i=1}^{n} -\left[\text{ - ln}(1\text{-}\varepsilon) + \ln(\sqrt{2\pi}) + \left(\text{k}\left|\frac{e_i}{\sigma}\right| - \frac{k^2}{2}\right)\right] & \text{if } |r_i| > k \end{cases}$$

The maximization of the log-likelihood function thus reduces to the minimization of the summation over n disturbances, $Q = \sum \rho\left[\frac{(y_i - \mathbf{x_i\beta})}{\sigma}\right]$, or equivalently of the following expression over the two regions

(6.10) $$\ln L(\beta) = \begin{cases} \frac{1}{2}\sum_{i=1}^{n}\left(\frac{e_i}{\sigma}\right)^2 = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \beta_1 x_{1i} - \dots - \beta_p x_{pi}}{\sigma}\right)^2 & \text{if } |r_i| \le k \\ \sum_{i=1}^{n}\left(k\left|\frac{e_i}{\sigma}\right| - \frac{k^2}{2}\right) = \sum_{i=1}^{n}\left(k\left|\frac{y_i - \beta_1 x_{1i} - \dots - \beta_p x_{pi}}{\sigma}\right| - \frac{k^2}{2}\right) & \text{if } |r_i| > k \end{cases}$$

Next we are interested in finding the maximum-likelihood estimator as the solution of equation (6.10). Since ρ is convex, it suffices to look at the first derivative of

$$Q = \sum \rho\left[\frac{(y_i - \mathbf{x}_i\boldsymbol{\beta})}{\sigma}\right] = \sum_{r_i \le k} \frac{1}{2}\left(\frac{e_i}{\sigma}\right)^2 + \sum_{|r_i|>k}\left(k\left|\frac{e_i}{\sigma}\right| - \frac{1}{2}k^2\right)$$

(6.11)

$$= \sum_{r_i \le k} \frac{1}{2} r_i^2 + \sum_{r_i > k}\left(kr_i - \frac{1}{2}k^2\right) + \sum_{r_i > k}\left(-kr_i - \frac{1}{2}k^2\right)$$

Setting the first partial derivatives with respect to the elements of $\boldsymbol{\beta}$, say β_j, equal to zero gives

(6.12) $$Q' = \sum_{r_i \le k} r_i x_i' + \sum_{|r_i|>k}\left(ksign(r_i)x_i'\right) = \sum_{r_i \le k} r_i x_i' + \sum_{r_i > k}\left(kx_i'\right) - \sum_{r_i < -k}\left(kx_i'\right) = 0$$

This is actually equivalent to finding the maximizing solution associated with the p equations

(6.13) $$\sum_{i=1}^{n} x_{ij}\psi\left[\frac{(y_i - \mathbf{x}_i\boldsymbol{\beta})}{\sigma}\right] = \mathbf{0}\,, \qquad j = 1, 2, \ldots, p$$

where x_{ij} is the element in the ith row and jth column of $\mathbf{X}$. In general, (6.13) is a set of non-linear equations, and iterative methods are required. The ψ-function is defined as $\psi(z) = \rho'(z)$. We may express the ψ-function for the two regions as follows

(6.14) $$\psi(r_i) = \begin{cases} \dfrac{e_i}{\sigma} & \text{if } |r_i| \le k \\[2ex] k\,\mathrm{sgn}\left(\dfrac{e_i}{\sigma}\right) & \text{if } |r_i| > k \end{cases}$$

A simple reformulation of M-estimators, namely by defining a weight function as $w(z) = \frac{\psi(z)}{z}$, gives a weighted mean where the weights are dependent on the data, providing both intuitive and computational support for this approach. The class of W-estimators provides a modification of the OLS estimation and form the basis of the technique of iteratively re-weighted least squares (IRLS).

Generally, we can rewrite the p equations of (6.13) as

(6.15) $$\sum_{i=1}^{n} x_{ij} \cdot \left[\frac{(y_i - \mathbf{x}_i\boldsymbol{\beta})}{\sigma}\right] \cdot w_i = \mathbf{0}\,, \qquad j = 1, 2, \ldots, p$$

A solution (not necessarily unique) to this system of equations will be the W-estimate. Rearrangement of the above equation in matrix terms yields the solution vector

(6.16) $$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{W}\mathbf{y})(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$$

In fact, these are the normal equations of a weighted least squares regression. However, the weights in this case are neither equal nor determined by **X**, while they depend on the residuals r_i.

Specifically, we may write the weight function for the two regions in the Huber's case as

$$(6.17) \qquad w(r_i) = \begin{cases} 1 & \text{if } |r_i| \le k \\ \dfrac{k}{|r_i|} & \text{if } |r_i| > k \end{cases}$$

In this way, the W-estimator is a computational procedure for solving the set of equations in (6.13) and will be the starting point in our programming in order to obtain the Huber's estimate.

Observations with small residuals will receive their own value, while those with large residuals will be gradually down-weighted. The scaled residual, r_i, is expressed as the absolute residual, e_i, divided by the scale estimator, σ. When defining the solution of (6.13), we assume that this error scale parameter is known and fixed. A class of research supports the idea of estimating $\boldsymbol{\beta}$ and σ simultaneously. However, we will assume this parameter to be estimated beforehand, that is, before each iteration step we choose a scale estimator and calculate its value, s. Considering this estimate as a known and fixed constant, we proceed with W-estimation for $\boldsymbol{\beta}$. The most commonly used resistant scale estimator is the median absolute deviation from the median residual (MAD)

$$(6.18) \qquad s = \frac{1}{0.6745}\,\text{median}\big[\,|\mathrm{e_i} - \text{median}(e_i)|\,\big]$$

where 0.6745 is the average value of the MAD for samples from the standard Gaussian distribution.

The tuning constant, k, in Huber's estimation is chosen to be equal to 1.345 in our application. Such a constant is chosen in order to refine the estimator so that it has a specific efficiency for the given distribution. Similarly, when the type of the ψ-function is decided, the tuning constant determines the properties of the underlying estimator – efficiency, gross error sensitivity, and so on. Setting the tuning constant to the above value produces approximately 95% asymptotic efficiency in the Gaussian distribution, i.e. in the efficiency of OLS given normally distributed errors. Lower tuning constants down-weight outliers more rapidly but may lead to unstable estimates. Higher tuning constants yield milder down-weighting, and the estimator's behavior is more similar to that in the OLS approach. From equation (6.17) we see that down-weighting begins with cases whose absolute residual, e_i, exceed $k{\cdot}\sigma$ = $(1.345/0.6745)\cdot\text{MAD} \approx 2{\cdot}\text{MAD}$.

Robust regression standard errors are calculated as discussed in Street, Carroll and Ruppert (1988). Given that the robust regression model is correct, the final estimates are asymptotically normally distributed and hypothesis tests are performed in the usual manner.

Within the frame of robust estimation, we discuss and apply two further issues. The *first* is related to the robustness of the Huber's estimator. Although the ψ–function is monotone, thus guaranteeing a single solution, the Huber's estimator may not be resistant enough to influential observations. An *M*-estimator can become more resistant by choosing a ψ–function that returns to zero. This is one of the properties of the redescending *M*-estimators (Biweight estimator, Andrew's sine estimator, Hampel' three-part estimator – which in many cases are reasonable substitutes for each other). One desirable feature of a redescending ψ–function is the way they treat severe influential observations by assigning them *zero weights* and thus increasing the efficiency at heavy-tailed distributions compared to the Huber case[53]. However, effectively dropping observations from the estimation may lead to undesirably inefficient prediction, making Huber a more preferable estimation procedure. Moreover, the ρ–function of, for example, the three above-mentioned redescending estimators is not convex, which leads to convergence failure or multiple solutions.[54]

The *second* issue is associated with the bounded-influence estimation. The influence function measures the effect (on the asymptotic bias of an estimator) of an arbitrarily small contamination of the assumed statistical model. In other words, the influence function is a measure of the sensitivity of an estimator to small changes in the data, and is thus a measure of robustness. OLS and other commonly used maximum-likelihood techniques have an unbounded influence function. Any small subset of contaminated data (data errors and/or imperfectly specified models) can have an arbitrarily large effect on any coefficient estimate. *M*-estimators, discussed above, provide certain protection against influential observations associated with the dependent variable, or *y*- influential observations. Since the *X*-data could also be aberrant, these estimators still have an unbounded influence function. Bounded-influence methods are developed to obtain further resistance against extreme (high-leverage) *x*-values.

In our analysis, we apply a bounded influence technique proposed by Schweppe (Handschin et. al. (1975)). This technique reduces the influence of influential observations in *x*-space. The Schweppe-type estimator satisfies the equations

$$(6.19) \qquad \sum_{i=1}^{n} x_{ij}\psi\left[\frac{(y_i - \mathbf{x}_i\boldsymbol{\beta})}{\sigma\sqrt{1-h_i}}\right]\sigma\sqrt{1-h_i} = \mathbf{0}\,, \qquad j = 1, 2, \ldots, p$$

where h_i is defined as previously. Huber (1983) arrives at this same form of estimator in a minimax framework. Again, this estimator may suffer from poor efficiency and may not be optimal, but it is reasonably simple to apply within the framework of our

---

[53] Generally, we may want to apply Huber's estimator when we are prepared to sacrifice efficiency in heavier-tailed distribution to achieve higher efficiency near the Gaussian. For overall efficiency and resistance we may want to choose a redescending estimator.

[54] In our study we applied, besides the Huber's estimator, an estimator suggested by Li (1985) which combines Huber's estimation and the biweight estimation (proposed by Beaton and Tukey (1974)). As mentioned, while the biweight estimator is more efficient, it is sensitive to starting values. The biweights are given as $w(r_i) = w(r_i) = \left(1-(r_i/b)^2\right)^2$ if $|r_i| \le b$ and $w(r_i) = 0$ if $w(r_i) = |r_i| > b$. Since the Huber estimator is much more stable and has a unique solution, the Huber weights are used first until convergence. Then, based on those results, the estimates are used as initial parameter values and biweights are used until convergence. This way, the initial Huber weighting should improve the behavior of the biweight estimator.

analysis. Moreover, this procedure may be less tolerant with respect to pairs or groups of outlying observations and should be used with some caution. From the equation above we derive the weight function for the Schweppe estimation given as[55]

$$(6.20) \qquad w(r_i) = \begin{cases} 1 & \text{if } \left| \dfrac{e_i}{\sigma\sqrt{1-h_i}} \right| \le k \\ \dfrac{k\sqrt{1-h_i}}{|r_i|} & \text{if } \left| \dfrac{e_i}{\sigma\sqrt{1-h_i}} \right| > k \end{cases}$$

where h_i are the diagonal elements of the hat matrix H. This estimator can be regarded as an M-estimator with Huber's ψ-function differing from observation to observation (the parameter k is replaced by $k\sqrt{1-h_i}$).

As in the case of M-estimators, there exist several bounded-influence estimators. Compared to M-estimators, bounded-influence estimators are generally less efficient based on a correctly specified model, but come up with a higher degree of protection.

A Schweppe-type estimator with suitable weight functions (which may differ from the original weights) is the most intuitive, if not optimal, way of bounding the influence, either on the fit or on the parameter (Hampel (1978)). Krasker and Welsch (1982) derive a Schweppe-type weight function that satisfies both the sensitivity bound and efficiency conditions. It can, however, be argued theoretically that the underlying assumptions of their model are too strong in order to satisfy this weighting. Moreover, some criticism is related to the protection against influential observations, which corresponds to situations where the gross errors in y occur selectively at the highest leverage points and may be seen as an unrealistically pessimistic assumption. However, we do not exclude the possibility of such a situation in our data and the relevance of the Krasker and Welsch estimator for certain variable estimations.

Since different estimators use different approaches to protect against influential observations, it is not an easy task to choose arbitrarily among them. The bounded-influence method that will be applied in our study is a simple modification of the Huber-type estimator. It relies on the available hat matrix diagonal elements, but may lack the efficiency offered by other techniques, which are more difficult to apply with standard program packages.

---

[55] In our analysis, we applied another bounded-influence technique suggested by Hamilton (1991)[55]. To account for x-outliers, we down-weight based on leverage. The predicted value obtained by least squares is expressed as $\hat{y} = Xb = Hy$, where $H = X(X'X)^{-1}X'$ is the hat matrix. Thus, strong leverage points (those with large diagonal elements of the hat matrix, h_i) dominate predicted or fitted values. By performing an initial OLS estimation, we define leverage weights as $w_i^L = 1$ if $|h_i| \le k^L$ and $w_i^L = (k^L / h_i)^2$ if $|h_i| > k^L$, where k^L is the 90th sample percentile of the diagonal elements of the hat matrix. From here we proceed with IRLS estimation using Huber's weights or biweights, $w(r_i)$, multiplied by the new weights provided above, that is, $w(r_i) \cdot w_i^L$ at each iteration step. While $w(r_i)$ changes at each step, w_i^L remains constant. The question of standard errors and different tests for bounded-influence estimation is more complicated than in the case of M-estimators and will not be discussed here. This method is easy to apply, but may contain certain problems with respect to efficiency and/or robustness.

6.3 Logistic Models with Complementary Log-Log Function

A look at different decision variables within firms indicates that many firms do not take any action at all regarding these variables, while other firms do. Decision variables that have such a characteristic are estimated in two steps. First, we investigate whether firms will take action or not. This is done with a logit model with which we find the probability of taking action. Second, we estimate the level of the decision variable for those firms that take action. We do this using the Huber-Schweppes robust estimation method.

In Appendix C, we outline some theoretical aspects of logistic models with cumulative logistic distribution function. Since we deal with contaminated data, we use the following representation of the firms' action regarding a decision variable

$$(6.21) \qquad P_i = F(X_i^{'}\beta) = 1 - e^{-e^{X_i^{'}\beta}}$$

where $F(\cdot)$ is the cumulative distribution of the complementary log-log function

$$(6.22) \qquad g(p) = \log(-\log(1-p)$$

The complementary log-log function is the inverse of the cumulative extreme-value function of (6.21), which is also called the Gombertz distribution or Weibull distribution. The characteristic of this distribution is that it is an asymmetric distribution, approaching 0 on the left more slowly than it approaches 1 on the right. This distribution seems to capture the underlying distribution of some firms' decision variables better then the cumulative logit distribution.

The model of interest in this case is

$$(6.23) \qquad \eta_i = \log\left[-\log\left(1-P_i\right)\right] = \beta_1 + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}$$

The predicted value for $P(Y \leq 1|x)$, when the cumulative extreme-value function is obtained by back-transforming the corresponding measures for the linear predictor, is as follows

$$(6.24) \qquad \hat{P}_i = 1 - e^{-e^{\hat{\eta}_i}}$$

Finally, we generate some artificial sample observations (or hypothetical critical values) $P_i^{critical}$, by using a random number generator. We draw randomly different random numbers from a uniform distribution between 0 and 1. Also, since the interest centers on the prediction ability of our model we design our simulated experiment so that we can compare the different outcome values for $\hat{P}_i$ with different values of $P_i^{critical}$. If the predicted probability $\hat{P}_i$ is higher than the critical probability $P_i^{critical}$, we draw the conclusion that the firm in this case takes action. If the predicted probability $\hat{P}_i$ is lower than $P_i^{critical}$, we draw the conclusion that the firm does not take action. These outcomes are then compared with the actual number of firms that take action and the actual number of firms that do not take action in our database. We compare the average predictive ability and decide whether it can be regarded as acceptable. However, we also examine whether the logit model predicts some firms, which in the database are observed to be non-acting companies, to take action.

6.4 Multinomial Models with Complementary Log-Log Function

This type of model applies to cases where an observation can fall into one of j categories. Many of the decision variables within firms can take positive, zero, and negative values. In Appendix C, we outline some theoretical aspects of multinomial models. Since we are dealing with contaminated data, we use the following representation of the firms' actions regarding a decision variable.

The model of interest in this case is

(6.25) $\quad \eta_{ic} = \log[-\log(1 - P_{ic})] = \mu_c + X_i{'}\beta$

where $c = 1.2$. The predicted value for $\hat{P}_{i1} = \hat{p}_{i1}$ is obtained by back-transforming the corresponding measures for the linear predictor as follows

(6.26) $\quad \hat{p}_{i1} = 1 - e^{-e^{\hat{\eta}_{i1}}}$

which is the probability that the firm takes negative action. The predicted value for $\hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2}$ is obtained by back-transforming the corresponding measures for the linear predictor as follows

(6.27) $\quad \hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2} = 1 - e^{-e^{\hat{\eta}_{i2}}}$

Using (C.17) we can solve an expression for $\hat{p}_{i2}$ from equation (C.18) as follows

(6.28) $\quad \hat{p}_{i2} = e^{\hat{\eta}_{i1}} - e^{\hat{\eta}_{i2}}$

which is the probability that the firm does not take any action. We know that $\hat{p}_{i3} = 1 - \hat{p}_{i1} - \hat{p}_{i2}$ so that

(6.29) $\quad \hat{p}_{i3} = e^{\hat{\eta}_{i2}}$

Let us now generate some artificial sample observations (or hypothetical critical values) $P_i^{critical}$, by using a random number generator. We draw randomly different random numbers from a uniform distribution between 0 and 1. Also, since the interest centers on the predictive ability of our model we design our simulated experiment so that we could compare the different outcome values for the cumulative probability $\hat{P}_{ic}$ with different values of $P_i^{critical}$. If the predicted probability $\hat{P}_{i1}$ is higher than the critical probability $P_i^{critical}$, we draw the conclusion that the firm in this case takes negative action. If $P_i^{critical}$ lies between the predicted probability $\hat{P}_{i1}$ and $\hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2}$ we draw the conclusion that the firm does not take any action. Finally, if $P_i^{critical}$ lies between the predicted probability $\hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2}$ and $P_{i3} = 1$ we draw the conclusion that the firm takes positive action.

These outcomes are then compared with the actual number of firms in our database that take negative, positive, or no action. We compare the average predictive ability and decide whether it can be regarded as acceptable. However, we also examine whether the multinomial logit model may predict some firms to take action, which in the database are observed not to do so.

6.5 The Tobit Model with a Logistic Distribution

The Tobit model, described by Tobin (1958), is a regression model for left-censored data assuming a normally distributed error term. The model parameters are estimated by maximum likelihood. Greene (2000) has a complete discussion of censored normal data and related distributions. In this section, we will derive left-censored data assuming a logistic distributed error term. The reason for this is that the parameter estimates for the normal distribution are sensitive to values that are extremely large. We know that this is the case in our database. We have both observations with large residuals and extreme values of covariates that affect the model parameters. The logistic distribution gives robust parameter estimates in the sense that the estimates have a bounded influence function. This is actually one of the good properties of the logistic distribution. The MLE has a bounded influence function with respect to the response variable. As regards the robustness of the MLE, it is worth mentioning here that the influence function is not bounded with respect to the leverage points. Quite a few papers talk about this. A good paper is that by Kunsch, Stefanski,and Carroll (1989). More papers can be found in their reference list.

In what follows, we will show how the predicted value for the TOBIT model is computed when the censored variable follows a logistic distribution. Consider a continuous random variable Y, and a constant a. If you were to sample from the distribution of Y but discard values less than (greater than) a, the distribution of the remaining observations would be **truncated** on the left (right). If you were to sample from the distribution of Y and report values less than (greater than) a, the distribution of the sample would be left- (right-) **censored**.

The probability density function of the truncated random variable Y is given by

$$(6.30) \qquad f_Y(y|y > a) = \frac{f_Y(y)}{P(y > a)}$$

for $y > a$ where $f_Y(y)$ is the probability density function of Y. If y has a logistic distribution with mean μ and standard deviation σ

$$(6.31) \qquad P(y > a) = 1 - \Phi^L(\alpha) = 1 - \Phi^L\left(\frac{a-\mu}{\sigma}\right)$$

The truncated logistic distribution is then

$$(6.32) \qquad f_Y(y|y > a) = \frac{\phi^L(\alpha)}{1 - \Phi^L(\alpha)} = \frac{e^{\alpha}}{\sigma[1 - \Phi^L(\alpha)][1 + e^{\alpha}]^2}$$

ϕ and Φ represent the logistic probability density and cumulative distribution functions. Let us take a closer look at the mean of the truncated logistic distribution.

(6.33) $$E(y|y>a)=\int_a^\infty y\, f_Y(y|y>a)dy=\frac{1}{[1-\Phi^L(\alpha)]}\int_a^\infty \frac{ye^\alpha}{\sigma[1+e^\alpha]^2}dy$$

where $\alpha=\dfrac{a-\mu}{\sigma}$. By developing the integral in (6.33) we get the following expression for the truncated logistic distribution.

(6.34) $$E(y|y>a)=\frac{1}{[1-\Phi^L(\alpha)]}\left[\mu-a\Phi^L(\alpha)+\sigma\ln\left(1+e^\alpha\right)\right]$$

However, we know that $CDF=\Phi^L(\alpha)=1/[1+e^{-\alpha}]$. This implies that $1+e^\alpha=1/[1-\Phi^L(\alpha)]$ so that

(6.35) $$E(y|y>a)=\frac{1}{[1-\Phi^L(\alpha)]}\left[\mu-a\Phi^L(\alpha)+\sigma\ln\left(\frac{1}{1-\Phi^L(\alpha)}\right)\right]$$

Suppose the model being fit is specified as follows:

(6.36) $$y_i^*=X_i'\beta+\varepsilon_i$$

Define the censored random variable y_i as

(6.37) $$\begin{aligned} y_i &= 1 \quad if \qquad y_i^*>0 \\ y_i &= 0 \quad if \qquad y_i^*=0 \end{aligned}$$

This is the Tobit model for left-censored data. y_i^* is sometimes called the *latent variable*. Let us now derive the mean of the censored logistic variable y_i.

(6.38) $$E(y_i)=a\Phi^L(\alpha)+[1-\Phi^L(\alpha)]\frac{1}{[1-\Phi^L(\alpha)]}\left[\mu-a\Phi^L(\alpha)+\sigma\ln\left(1+e^\alpha\right)\right]$$

After few simplifications, we can rewrite equation (6.38) as follows

(6.39) $$E(y_i)=\mu+\sigma\ln\left(1+e^{\frac{q-\mu}{\sigma}}\right)$$

In our case $a=0$ so that

(6.40) $$E(y_i)=\sigma\ln\left(1+e^{\mu/\sigma}\right)$$

The computer program SAS computes predicted values based on the mean functions of the latent and observed variables. The mean of the latent variable y_i^* is $X_i'\beta$ and we can compute values of the mean for different settings of X_i. Predicted values of the variable y_i can be computed based on the mean specified in (6.40).

We use the computer program SAS to estimate parameters of the distribution of y_i^* by maximum likelihood. Suppose there are n observations from the model $y_i^* = X_i'\beta + \sigma\varepsilon_i$, where X is an $n \times k$ matrix of covariate values (including the intercept), y is a vector of responses, and ε is a vector of errors with the cumulative distribution function F and probability density function f. That is, $F(t) = P(\varepsilon_i \le t)$, and $f(t) = dF(t)/dt$, where ε_i is a component of the error vector. Then, if all the responses are observed, the log likelihood, *L*, can be written as

$$(6.41) \qquad L = \sum \ln\left(\frac{f(w_i)}{\sigma}\right)$$

where $w_i = \frac{1}{\sigma}(y_i^* - X_i'\beta)$. If some of the responses are left-censored, the log likelihood can be written as

$$(6.42) \qquad L = \sum \ln\left(\frac{f(w_i)}{\sigma}\right) + \sum \ln\left(F(w_i)\right)$$

with the first sum regarding uncensored observations and the second sum regarding left-censored observations. Additional information on censored and limited dependent variable models can be found in Maddala (1983). The estimated covariance matrix of the parameter estimates is computed as the negative inverse of I, which is the information matrix of second derivatives of L with respect to the parameters evaluated at the final parameter estimates. If I is not positive definite, a positive definite submatrix of I is inverted, and the remaining rows and columns of the inverse are set to zero. If some of the parameters, such as the scale and intercept, are restricted, the corresponding elements of the estimated covariance matrix are set to zero. The standard error estimates for the parameter estimates are taken as the square roots of the corresponding diagonal elements.

For restrictions placed on the intercept, scale, and shape parameters, one-degree-of-freedom Lagrange multiplier test statistics are computed. These statistics are computed as

$$(6.43) \qquad \chi^2 = \frac{g^2}{V}$$

where g is the derivative of the log likelihood with respect to the restricted parameter at the restricted maximum and $V = I_{11} - I_{12}I_{22}^{-1}I_{21}$ where the 1 subscripts refer to the restricted parameter and the 2 subscripts refer to the unrestricted parameters. The information matrix is evaluated at the restricted maximum. These statistics are asymptotically distributed as chi-squares with one degree of freedom under the null hypothesis that the restrictions are valid, provided that some regularity conditions are satisfied.

A Lagrange multiplier test statistic is also computed to test this constraint. Notice that this test statistic is comparable to the Wald test statistic for testing that the scale is 1. The Wald statistic is the result of squaring the difference of the estimate of the scale parameter from 1 and dividing this by the square of its estimated standard error.

Predicted values: The estimated parameters of the TOBIT model can be used to predict the behavior of firms who must choose whether or not to take action. This is done for the actual values of the explanatory variables, $X_i^{*'}$. In (6.31), we showed that $P(y > a) = 1 - \Phi^L(\alpha) = 1 - \Phi^L\left(\frac{a-\mu}{\sigma}\right)$ where $CDF = \Phi^L(\alpha) = 1/[1+e^{-\alpha}]$. Using the CDF, we find out that $P(y > a) = 1 - \frac{1}{1+e^{-\alpha}} = \frac{e^{-\alpha}}{1+e^{-\alpha}}$. For $a = 0$, it is easy to show that

$$(6.44) \qquad \hat{P}_i(y > 0) = \frac{e^{\mu/\sigma}}{1+e^{\mu/\sigma}} = \frac{e^{(X_i'\beta)/\sigma}}{1+e^{(X_i'\beta)/\sigma}}$$

we generate some artificial sample observations (or hypothetical critical values) $P_i^{critical}$, by using a random number generator. We draw randomly different random numbers from a uniform distribution between 0 and 1. Also, since the interest centers on the predictive ability of our model we design our simulated experiment so that we could compare the different outcome values for $\hat{P}_i$ with different values of $P_i^{critical}$. If the predicted probability $\hat{P}_i$ is higher than the critical probability $P_i^{critical}$, we draw the conclusion that the firm in this case takes action. If the predicted probability $\hat{P}_i$ is lower than $P_i^{critical}$, we draw the conclusion that the firm does not take action. These outcomes are then compared with the actual number of firms that take action and the actual number of firms that do not take action, in our database. We compare the average predictive ability and decide whether it can be regarded as acceptable. However, we also examine whether the Tobit model may predict some firms, which in the database are observed to be non-acting firms, to take action.

Having investigated whether the firm will take action or not, we go on to estimate the level of the decision variable for those firms as follows. If the predicted probability $\hat{P}_i$ is higher than the critical probability $P_i^{critical}$, the predicted value for those firms that take action is computed by $E(y_i) = \sigma \ln\left(1 + e^{\mu/\sigma}\right)$, which is the mean of the censored logistic variable y_i.

If the predicted probability $\hat{P}_i$ is lower than $P_i^{critical}$, the predicted value for those firms that do not take any action is $E(y_i) = 0$.

Marginal effects: In addition we also present the marginal effects for the TOBIT model given the censoring and the underlying logistic distribution as follows

(6.45) $$\frac{\delta E(y_i \mid x_i)}{\delta x_i} = \beta \frac{e^{\mu/\sigma}}{1+e^{\mu/\sigma}} = \beta \frac{e^{(X_i'\beta)/\sigma}}{1+e^{(X_i'\beta)/\sigma}}$$

In interpreting the marginal effects, it will be useful to evaluate these effects in the means of the regressors.

7 The Estimation Results

In this chapter, we present the estimation results for I_t^{MA}, S_t^{MA}, $EDEP_t^{MA}$, I_t^{BU}, $EDEP_t^{BU}$, $TDEP_t^{BU}$, $dofa_t$, dca_t, dsc_t, drr_t, $dour_t$, p_t^{allo}, zpf_t, $TDEP_t^{MA}$, dll_t, dcl_t, $OIBD_t$, FI_t, FE_t, OA_t, GC_t, TL_t, OTA_t, and ROT_t. As we mentioned in the previous chapter, we use different estimation methods for estimating these flow variables. In our presentation, we begin by showing the predication made by each estimation method. This only gives us information about the predictive accuracy of different estimation methods. It should be notified that the estimations are made on pooled data from 1997-1999. However, the starting point for the simulation model is 1999, which of course has fewer observations. This means that a method that happens to give the best prediction may fail to give a good prediction in the simulation model. Moreover, in the simulation model, the prediction made by a decision variable is dependent on predictions made by other variables. This means that we may choose an estimation method that does not give the best sample prediction.

The estimation methods that we use are the following:

1) The Huber-Schweppes robust estimation method (HS).

2) For those variables that only take non-negative values, we use a logistic model with the *cumulative logistic distribution function* to find the probability that the variable is positive. Then, we use the Huber-Schweppes robust estimation method to estimate the positive level of the variable. (LLN)

3) For those variables that only take non-negative values, we use a logistic model with the *complementary log-log distribution function* to find the probability that the variable is positive. Then, we use the Huber-Schweppes robust estimation method to estimate the positive level of the variable. (LLG)

4) For those variables that can be negative, zero or positive, we use a logistic model with the *cumulative logistic distribution function* to find the probability that the variable is positive. Second, we use another logistic model with the *cumulative logistic distribution function* to find the probability that the variable is negative. Third, we use the Huber-Schweppes robust estimation method to estimate the positive level of the variable. Finally, we use the Huber-Schweppes robust estimation method to estimate the negative level of the variable. (LSN)

5) For those variables that can be negative, zero or positive, we use a logistic model with the *complementary log-log distribution function* to find the probability that the variable is positive. Second, we use another logistic model with the *complementary log-log distribution function* to find the probability that the variable is negative. Third, we use the Huber-Schweppes robust estimation method to estimate the positive level of the variable. Finally, we use the Huber-Schweppes robust estimation method to estimate the negative level of the variable. (LSN)

6) For those variables that can be negative, zero or positive, we use a multinomial model with the *complementary log-log distribution function* to find the probabilities that the variable is positive, equal to zero, or negative. Second, we use the Huber-Schweppes robust estimation method to estimate the positive level of the variable. Finally, we use Huber-Schweppes robust estimation method to estimate the negative level of the variable. (LSG)

7) For those variables that only take non-negative values, we use a *Tobit model with a logistic distribution function*, which is a combination of a truncated regression model and a probit. (Tobit 1)

8) For those variables that can be negative, zero or positive, we use two *Tobit models with a logistic distribution function*. (Tobit 2)

7.1 Estimating Economic Depreciation of Machinery and Equipment

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate economic depreciation of machinery and equipment.

Variable	Method	Mean	Sum	Std. Dev.	MSE
$EDEP^{MA}$		2164552.96	171555973575	34911990.83	
$pEDEP^{MA}$	LLN	1845239.45	146248142951	29319815.50	20868631.27
	LLG	1719706.51	136298779006	28391034.14	21789638.35
	Tobit 1	2041317.35	161788689115	20039680.60	25790214.53

As is evident by looking at the descriptive statistics, the Tobit 1 method better predicts firms' economic depreciation of machinery and equipment. At the same time, we calculate the mean square error as a measure of model performance. The best prediction should in this case be the one that minimizes MSE. However, our simulations indicate that the LLG method gives better predictions than any other method. The performance of each prediction in the simulation context will be the most important criterion in choosing the proper method. In what follows, we will outline the estimation results for the LLG method.

Economic depreciation of machinery and equipment is estimated in two steps. First (in section 7.1.1), we investigate whether firms make economic depreciation of their machinery and equipment. This will be done using a logistic model with which we find the probability of economic depreciation of machinery and equipment being positive. Second (in section 7.1.2), we estimate the level of economic depreciation of machinery and equipment investment for those firms that make economic depreciation of their machinery and equipment. This is done with the Huber-Schweppes robust estimation method.

7.1.1 The Probability that Economic Depreciation of Machinery and Equipment is Positive

The data regarding economic depreciation of machinery and equipment indicate that 15857 firms (of a total of 79257 firms) do not account for economic depreciation of their machinery and equipment at all. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.1)\quad \begin{array}{lll} DEDEP_i^{MA} = 1 & \quad if \quad & EDEP_i^{MA} > 0 \\ DEDEP_i^{MA} = 0 & \quad if \quad & EDEP_i^{MA} = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms account for economic depreciation of its machinery and equipment and 0 if firms do not account for economic depreciation. Table 1 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 1a. As the value of explanatory variables changes, the value of the index $DEDEP_i^{MA}$ varies. The larger the value of $DEDEP_i^{MA}$ the greater the incentive firms *i* receive from choosing the option $EDEP_i^{MA} > 0$. Thus, the greater the value of $DEDP_i^{MA}$, the greater will be P_i, the probability that firms *i* choose the option $EDEP_i^{MA} > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the probability that firms choose the option $EDEP_i^{MA} > 0$ increases with firms' accounted tax depreciation of their machinery and equipment. Second, the probability also increases with the level of machinery and equipment. Third, the probability increases with firms' investment in machinery and equipment. However, the increase in probability is decreasing. Fourth, the probability increases with the accounted economic depreciation of buildings. This increase in probability is also decreasing. Fifth, the probability increases also with the change in current assets.

Sixth, the change in the difference between the maximum amount of allocations to periodical reserves and the actual amount of allocations to periodical reserves has a positive impact on the probability that firms choose the option $EDEP_i^{MA} > 0$ (however, the increase in the probability is increasing.). This needs some explanation. We know that $n_u^{PF}\, dmpa = 0$. This condition stipulates that if the periodical reserve option is not fully used in the optimal solution ($\frac{\delta F}{\delta n_u^{PF}} > 0$), the shadow price of that constraint, which is never allowed to be negative, must be equal to zero ($n_u^{PF} = 0$). On the other hand, if the constraint has a positive shadow price in the optimal solution ($n_u^{PF} > 0$), then it is perforce a fully utilized allocation to periodical reserves ($\frac{\delta F}{\delta n_u^{PF}} = 0$). The Lagrange multiplier n_u^{PF} is a measure of how the optimal value of firms reacts to a slight relaxation of the constraint. In that light, complementary slackness means that if the constraint is optimally not binding, relaxing that particular constraint will not affect the optimal value of firms. If, on the other hand, a slight relaxation of the constraint does increase the value of the firms, the constraint must in fact be binding in the optimal solution. As we showed earlier, a slight relaxation of the constraint affects the complementary-slackness conditions in the following way: $dn_u^{PF} = -\frac{n_u^{PF}}{dmpa} ddmpa$. If we assume that the constraint is optimally binding, changing that particular constraint will affect the optimal value of firms. Here, this has a positive impact on the probability that firms choose the option $EDEP_i^{MA} > 0$. Another way of putting this is the following: Assume that firms do not maximally utilize the tax rules regarding allocations to periodical reserves (this means that $dmpa > 0$). Assume further that firms increase their underutilization in the next period so that $ddmpa > 0$. What the positive coefficient in front of $ddmpa$ says is that the probability that firms undertake an economic depreciation of their machinery and equipment increases with their increase of under-utilization.

Seventh, the probability increases if the change in gross national product increases. A higher value for this macro variable indicates an upward economic trend and reflects the state of the market.

(b) First, the probability that firms choose the option $EDEP_i^{MA} > 0$ decreases with the change in the difference between the maximum allowed tax depreciation of machinery and equipment, and the actual tax depreciation of these assets. We know that $n_u^{MA}\, dmtdm = 0$. This condition stipulates that if the tax rules for depreciation of machinery and equipment are not fully used in the optimal solution ($\frac{\delta F}{\delta n_u^{MA}} \geq 0$), the shadow price of that constraint must be equal to zero ($n_u^{MA} = 0$). On the other hand, if the constraint has a positive shadow price in the optimal solution ($n_u^{MA} > 0$), it is perforce a fully utilized tax depreciation of machinery and equipment ($\frac{\delta F}{\delta n_u^{MA}} = 0$). The Lagrange multiplier n_u^{MA} is a measure of how the optimal values of firms react to a slight relaxation of the constraint. As we showed earlier, a slight relaxation of the constraint affects the complementary-slackness conditions in the following way: $dn_u^{MA} = -\frac{n_u^{MA}}{dmtdm} ddmtdm$. If we assume that the constraint is optimally binding, changing that particular constraint will affect the optimal values of firms. Here, this has a negative impact on the probability that firms choose the option $EDEP_i^{MA} > 0$.

Second, the probability that firms choose the option $EDEP_i^{MA} > 0$ decreases with the change in cash flow. We know that $n_f \, dcashfl = 0$. The Lagrange multiplier n_f is a measure of how the optimal values of firms react to a slight relaxation of the constraint. As we showed earlier, a slight relaxation of the constraint affects the complementary-slackness conditions in the following way: $dn_f = -\frac{n_f}{cashfl} dcashfl$. If we assume that the dividend floor constraint is optimally non-binding, changing that particular constraint will not affect the optimal value of firms. A $dcashfl < 0$ implies that firms come closer to the dividend floor. Firms that reach the dividend floor faster have lower incentives to choose the option $EDEP_i^{MA} > 0$.

Third, the probability that firms choose the option $EDEP_i^{MA} > 0$ decreases with the change in the difference between the maximum allowed dividend payments according to legislation and the actual maximum amount available for dividends. We know that $n_r \, dmcash = 0$. The Lagrange multiplier n_r is a measure of how the optimal value of the firms reacts to a slight relaxation of the dividend constraint. As we showed earlier, a slight relaxation of the constraint affects the complementary-slackness conditions in the following way: $dn_r = -\frac{n_r}{dmcash} ddmcash$. Assume that the dividend constraint is not binding $dmcash > 0$. A $ddmcash < 0$ implies that firms come closer to the dividend constraint. The probability that firms choose the option $EDEP_i^{MA} > 0$ decreases if firms reach the dividend constraint faster.

(c) Moreover, the magnitude of the probability that firms choose the option $EDEP_i^{MA} > 0$ can be different for firms that are closed firms, for firms that are public firms, and for firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $EDEP_i^{MA} > 0$ is higher for firms with a market concentration index near 1. The probability is also higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicted zeros (19.95 percent) and positive economic depreciation of machinery and equipment (80.05 percent) almost coincide with the observed responses (20.01 percent and 79.99 percent, respectively). However, it is also evident that there is 33 percent discordance between the predicted probabilities and the observed responses.

7.1.2 The Level of Economic Depreciation of Machinery and Equipment

A look at economic depreciation of machinery and equipment undertaken by firms indicates that 63400 firms have accounted for economic depreciation of their machinery and equipment. Influence diagnostics show that 442 observations have severe influence on the response variable. Further, we find that we also have problems with leverage points (for 2007 observations) - points that are outliers in the design matrix. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, Cramer von Mises test, and the Anderson-Darling test). The distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of economic depreciation of machinery and equipment. Table 1b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Economic depreciation of machinery and equipment increases with firms' tax depreciation of their machinery and equipment, the level of machinery and equipment, firms' investment in machinery and equipment (the increase is decreasing), the change in the utilization of allocations to periodical reserves (this increase is also decreasing), and the change in cash flow.

(b) Economic depreciation of machinery and equipment decreases if firms reach the dividend constraint faster. Second, economic depreciation of machinery and equipment decreases with the accounted economic depreciation of buildings (this decrease is increasing), firms' change in their utilization of the tax rules for the maximum allowed tax depreciation of machinery and equipment (this decrease is decreasing), the change in current assets, and the change in gross national product.

(c) Moreover, the level of economic depreciation of machinery and equipment is lower for those firms that are closed firms, and for firms that are located in large cities and rural districts. On the other hand, the level of economic depreciation of machinery and equipment is higher for firms that are public firms.

(d) Finally, economic depreciation of machinery and equipment is higher for firms with a market concentration index near 1. This means that economic depreciation of machinery and equipment is lower for firms that operate in more competitive markets. Economic depreciation is also higher for firms that have high market shares.

7.2 Estimating Net Sales of Machinery and Equipment

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate net sales of machinery and equipment. As is evident, none of the estimation methods make a good prediction of S^{MA}. As will be evident in the next section, our simulations indicate that the MUNO method gives a satisfactory prediction of S^{MA}. In what follows, we will outline the estimation results for the MUNO method.

Variable	Method	Mean	Sum	Std. Dev.	MSE
S^{MA}		-1645164.33	-1.303908E11	387737778	
pS^{MA}	LLN	566874.34	44928759400	89576150.49	372090901.79
	LLG	606405.17	48061854446	82913608.98	379423802.33
	LSN	1736282.13	137612512393	146030660	367271734.22
	LSG	746744.93	59184763154	101444787	378577136.37
	MUNO	330153.28	26166958492	107218176	379952703.11

Net sales of machinery and equipment are estimated in two steps. First (in section 7.2.1), we use a multinomial logit model to investigate whether the sale of machinery and equipment is positive, equal to zero, or negative. Second (in section 7.2.2), we estimate the level of the positive sales with the Huber-Schweppes robust estimation method. Third (in section 7.2.3), we estimate the level of the negative sales with the Huber-Schweppes robust estimation method.

7.2.1 The Probabilities that Firms Make Net Sales of Their Machinery and Equipment

A look at net sales of machinery and equipment undertaken by firms indicates that 21452 firms account for positive sales (also includes revaluation), 19308 firms account for negative sales (devaluation), while 38497 firms do not account for sales of their machinery and equipment. To capture this fact we generate a dependent variable, which is dichotomous in nature, as

$$\begin{aligned} &DS^{MA} = 1 &\quad if \quad & S_i^{MA} < 0 \\ (7.2)\quad &DS^{MA} = 2 &\quad if \quad & S_i^{MA} = 0 \\ &DS^{MA} = 3 &\quad if \quad & S_i^{MA} > 0 \end{aligned}$$

Hence, the dependent variable can take three values: 1 if the sales are negative, 2 if firms do not sell any of their machinery and equipment, and 3 if the sales are positive. Table 2a concludes the estimated multinomial logit model with a complementary log-log function. The coefficients in this model are difficult to interpret. As is evident from section 6.1.5, it is rather difficult to find the marginal effects of the regressors on the probabilities. We will not do that here. However, these marginal effects can be computed within the statistical program package SAS, which is the program package that we use for all our estimations.[56] As we explained in section 6.1.5, the obtained cumulative probabilities from the multinomial logit model are then compared with a generated critical probability value to be able to find out whether firms will undertake a positive sale, a negative sale or no sale at all of their machinery and equipment. Having established the choice of the firms, we need to decide the level of positive and negative sales.

[56] See the SAS/STAT User Guide, Version 8, Volume 2, Chapter 29 “PROC GENMOD”, Example 29.4.

7.2.2 The Level of Positive Sales of Machinery and Equipment

Influence diagnostics show that there are 275 observations that have severe influence on the response variable. Further, it is also evident that there are 877 observations with high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the (net) sales of machinery and equipment. Table 2b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the positive sales of machinery and equipment increase with the level of machinery and equipment, the change in current assets, the change in current liabilities (this increase is increasing), the change in gross national product, the economic depreciation of buildings (this increase is increasing), the change in firms' utilization of their allocations to periodical reserves (this increase is decreasing), and the change in their utilization of the tax rules for tax depreciation of machinery and equipment.

(b) Second, the positive sales of machinery and equipment decrease with firms' economic depreciation of their machinery and equipment, the change in cash flow, and firms' net investment in buildings (this decrease is decreasing).

(c) Moreover, the positive sales of machinery and equipment are lower for firms that are closed firms and for firms that are located in rural districts. On the other hand, the positive sales are higher for firms that are public firms and firms that are located in large cities.

(d) Finally, the positive sales of machinery and equipment are lower for firms with a market concentration index close to 1. This means that the positive sales are higher for firms that operate in more competitive markets. The positive sales are also lower for firms that have high market shares.

7.2.3 The Level of Negative Sales of Machinery and Equipment

Influence diagnostics show that there are 17 observations that have severe influence on the response variable. Further, it is also evident that there are 490 observations with high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the (net) sales of machinery and equipment. Table 2c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the negative sales of machinery and equipment increase with the change in current liabilities, the level of machinery and equipment, the economic depreciation of buildings (this increase is decreasing), the change in firms' utilization of their allocations to periodical reserves (this increase is decreasing), the change in their utilization of the tax rules for tax depreciation of machinery and equipment, and the change in gross national product.

(b) Second, the negative sales of machinery and equipment decrease with the change in current assets, firms' investment in buildings and the economic depreciation of machinery and equipment (this decrease is increasing), and the change in cash flow. Finally, the negative sales of machinery and equipment decrease for firms that change their dividend policy so that they come closer to the legal constraint on dividend payments.

(c) Moreover, the negative sales of machinery and equipment are lower for firms that are public firms and firms that are located in rural districts. On the other hand, the negative sales are higher for firms that are closed firms and firms that are located in large cities.

(d) Finally, the negative sales of machinery and equipment are lower for firms with a market concentration index close to 1. This means that the negative sales are higher for firms that operate in more competitive markets. The negative sales are also lower for firms that have high market shares.

7.3 Estimating Investment in Machinery and Equipment

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate investment in machinery and equipment.

Variable	Method	Mean	Sum	Std. Dev.	MSE
I^{MA}		2985313.01	236606953116	67211295.62	
pI^{MA}	LLN	2177917.44	172615202707	36782034.47	50502739.64
	LLG	1888348.88	149664866914	36406739.84	50640693.23
	Tobit 1	2439182.18	193322261876	28282487.98	53529925.76

As is evident, the Tobit 1 method gives the best prediction of firms' investment in machinery and equipment. Our simulations indicate that this method is also best suited to the simulation model. In section 7.3.1, we will outline the estimation results for the Tobit 1 method.

7.3.1 The Tobit Model for Investment in Machinery and Equipment

Let us generate a dependent variable, which has the following characteristics:

$$(7.3)\quad \begin{array}{lll} LOWER = I_i^{MA} & if & I_i^{MA} > 0 \\ LOWER = . & if & I_i^{MA} \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: I_i^{MA} if firms invest in machinery and equipment and missing values otherwise. Table 3 concludes the estimated Tobit 1 model with a logistic distribution function. Let us now interpret the results in Table 3a. Direct interpretation of various regression coefficients given in Table 3a is not easy. But from the estimated coefficients one can assess the marginal effects. As Table 3a shows:

(a) An increase in investment in buildings, economic depreciation of machinery and equipment, economic depreciation of machinery and equipment, tax depreciation of machinery and equipment, the change in firms' utilization of their allocations to periodical reserves, the change in current liabilities, and the change in gross national product will increase firms' investment in machinery and equipment.

(b) An increase in net sales of machinery and equipment, the change in current assets, the change in cash flow, and the change in firms' dividend policy so that they come closer to the legal constraint on dividend payments, will decrease firms' investment in machinery and equipment.

(c) Moreover, investment in machinery and equipment is lower for those firms that are located in large cities. On the other hand, investment is higher for those firms that are public firms, closed firms, and firms that are located in rural districts.

(d) Finally, investment in machinery and equipment is higher for firms with a market concentration index close to 1. This means that investment in machinery and equipment is lower for firms that operate in more competitive markets. Investment in machinery and equipment is higher for firms that have high market shares.

7.4 Estimating Economic Depreciation of Buildings

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate economic depreciation of buildings.

Variable	Method	Mean	Sum	Std. Dev.	MSE
$EDEP^{BU}$		373828.87	29628554820	4455424.75	
$pEDEP^{BU}$	LLN	316462.19	25081843872	4002355.18	3421189.42
	LLG	217608.30	17246981021	6786438.77	6735867.49
	Tobit 1	305685.26	24227696647	3305222.31	3066486.87

As is evident, the LLN method gives the best prediction of firms' economic depreciation of buildings. However, our doubts regarding the normality of the variable distribution and simulation results indicate that this method is not best suited to the simulation model. Instead, the LLG method gives better prediction than the LLN method. In what follows, we will outline the estimation results for the LLG method.

Economic depreciation of buildings is estimated in two steps. First (in section 7.4.1), we investigate whether firms make economic depreciation of their buildings. This is done using a logistic model with which we find the probability that economic depreciation of buildings is positive. Second (in section 7.4.2), we estimate the level of economic depreciation of buildings for those firms that undertake an investment. This is done with the Huber-Schweppes robust estimation method.

7.4.1 The Probability that Economic Depreciation of Buildings is Positive

A look at economic depreciation of buildings undertaken by firms indicates that 61410 firms (of 79257 firms) do not account for economic depreciation of their buildings. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.4) \quad \begin{array}{lll} DEDEP_i^{BU} = 1 & if & EDEP_i^{BU} > 0 \\ DEDEP_i^{BU} = 0 & if & EDEP_i^{BU} \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms account for economic depreciation of their buildings and 0 if firms do not account for economic depreciation. Table 4 concludes the estimated logistic model with a complementary log-log distribution function. Let us now interpret the results in Table 4a. As the value of explanatory variables changes, the value of the index $DEDEP_i^{BU}$ varies. The larger the value of $DEDEP_i^{BU}$ the greater the incentive firms *i* receive from choosing the option $EDEP_i^{BU} > 0$. Thus, the greater the value of $DEDEP_i^{BU}$, the greater will be P_i, the probability that firms *i* choose the option $EDEP_i^{BU} > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $EDEP_i^{BU} > 0$ increases with the economic depreciation of machinery and equipment (this increase is decreasing), the level of buildings (this increase is decreasing), the change in current assets, the change in current liabilities, the change in gross national product, the change in cash flow, and the change in firms' dividend policy so that they come closer to the legal constraint on dividend payments.

(b) The probability that firms choose the option $EDEP_i^{BU} > 0$ decreases with the sales of machinery and equipment, the investment in machinery and equipment, and the change in firms' utilization of their allocations to periodical reserves.

(c) Moreover, the magnitude of the probability that firms choose the option $EDEP_i^{BU} > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $EDEP_i^{BU} > 0$ is higher for firms with a market concentration index close to 1. The probability is also higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not account for economic depreciation is 77.71 percent, and the frequency for predicting positive economic depreciation of buildings is 22.29 percent. This coincides approximately with the observed responses (77.48 percent and 22.52 percent, respectively). However, it is also evident that there is 31.3 percent discordance between the predicted probabilities and the observed responses.

7.4.2 The Level of Economic Depreciation of Buildings

A look at economic depreciation of buildings undertaken by firms indicates that 17847 firms have accounted for economic depreciation of their buildings. Influence diagnostics show that there are 278 observations that have severe influence on the response variable. Further, we find out that we also have problems with leverage points (for 624 observations) - points that are outliers in the design matrix. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of economic depreciation of buildings. Table 4d concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Economic depreciation of buildings increases with firms' economic depreciation of their machinery and equipment (the increase is decreasing), investment in machinery and equipment, the level of buildings, and the change in current liabilities.

(b) Economic depreciation of buildings decreases with the change in current assets, the change in cash flow, net sales of machinery and equipment (this decrease is decreasing), the change in utilization of the tax rules for maximum allowed allocations to periodical reserves, the change in gross national product, and the change in firms' dividend policy so that they come closer to the legal constraint on dividend payments.

(c) Moreover, the level of economic depreciation of buildings is lower for firms that are closed firms and firms that are located in large cities and rural districts. On the other hand, the level of economic depreciation of buildings is higher for public firms.

(d) Finally, economic depreciation of buildings is higher for firms with a market concentration index close to 1. This means that economic depreciation of buildings is lower for firms that operate in more competitive markets. Economic depreciation is higher for firms that have high market shares.

7.5 Estimating Investment in Buildings

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate investment in buildings.

Variable	Method	Mean	Sum	Std. Div.	MSE
I^{BU}		1111386.90	88085191902	154251931	
pI^{BU}	LLN	759764.06	60216620473	26478001.05	142217009.64
	LLG	615033.70	48745726307	25966895.91	142147310.36
	LSN	790922.01	62686105418	56493877.54	144652797.51
	LSG	578639.13	45861201900	51570328.16	152262776.63

As is evident, the LSN method gives the best prediction of firms' investment in buildings. However, our doubts regarding the normality of the variable distribution and simulation results indicate that this method is not best suited to the simulation model. Instead, the LLG method gives a better prediction than the LSN method. In what follows, we will outline the estimation results for the LLG method.

Investment in buildings is estimated in two steps. First (in section 7.5.1), we investigate whether firms make positive or negative investment in buildings. This is done using a logistic model with which we find the probability that firms undertake investment in buildings. Second (in section 7.5.2), we estimate the level of net investment in buildings.

7.5.1 The Probability that Investment in Buildings is Either Positive or Negative

A look at the database indicates that 11090 firms (of a total of 79257 firms) have made positive investment in buildings. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.5)\quad \begin{array}{lll} DI_i^{BU} = 1 & if & I_i^{BU} \neq 0 \\ DI_i^{BU} = 0 & if & I_i^{BU} = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms invest (either positive or negative) in buildings and 0 otherwise. Table 5 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 5a. As the value of the explanatory variables change, the value of the index DI_i^{BU} varies. The larger the value of DI_i^{BU} the greater the incentive for firms *i* to choose the option $I_i^{BU} \neq 0$. Thus, the greater the value of DI_i^{BU}, the greater will be P_i, the probability that firms *i* choose the option $I_i^{BU} > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $I_i^{BU} \neq 0$ increases with the economic depreciation of machinery and equipment (this increase is decreasing), investment in machinery and equipment (this increase is decreasing), net sales of machinery and equipment, economic depreciation of buildings (this increase is decreasing), the change in current assets, the change in the utilization of tax rules regarding allocations to periodical reserves (this increase is decreasing), the change in cash flow, the change in firms' dividend policy so that they come closer to the legal constraint on dividend payments, and the change in current liabilities.

(b) The probability that firms choose the option $I_i^{BU} \neq 0$ decreases with investment in machinery and equipment and the change in gross national product.

(c) Moreover, the magnitude of the probability that firms choose the option $I_i^{BU} \neq 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $I_i^{BU} \neq 0$ is higher for firms with a market concentration index close to 1. The probability is also higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not invest in buildings is 78.35 percent, and the frequency for predicting that firms will invest in buildings is 21.65 percent. These probabilities coincide with the observed responses (78.35 percent and 21.65 percent, respectively). However, it is also evident that there is 27.2 percent discordance between the predicted probabilities and the observed responses.

7.5.2 The Level of Positive and Negative Investment in Buildings

Influence diagnostics show that 291 observations have severe leverage points - points that are outliers in the design matrix. Further, 740 observations have severe influence on the response variable. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of investment in buildings. Table 5b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Investment in buildings increases with investment in machinery and equipment, economic depreciation of buildings (this increase is decreasing), the change in current liabilities, and the change in the utilization of tax rules regarding allocations to periodical reserves.

(b) Investment in buildings decreases with economic depreciation of machinery and equipment (this decrease is increasing), sales of machinery and equipment, the change in gross national product, the change in current assets, the change in cash flow, and the change in firms' dividend policy so that they come closer to the legal constraint on dividend payments.

(c) Moreover, investment in buildings is lower for those firms that are closed firms and for firms that are located in rural districts or large cities. On the other hand, investment in buildings is higher for public firms.

(d) Finally, investment in buildings is higher for firms with a market concentration index close to 1. This means that investment in buildings is lower for firms that operate in more competitive markets. On the other hand, investment in buildings is lower for firms that have high market shares.

7.6 Estimating the Net Change in Other Fixed Assets

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate the net change in other fixed assets. As is evident, the Tobit 2 method gives the best prediction of the change in other fixed assets. However, our simulations indicate that this method is not best suited to the simulation model. Instead, the LSG method gives a better prediction than the Tobit 2 method. In what follows, we will outline the estimation results for the LSG method.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dofa		23846605.85	1.8900104E12	4035157262	
pdofa	LLN	1473176.40	116759541548	13389577.87	4034867537.3
	LLG	1506503.81	119400972759	13327541.38	4034875860.6
	LSN	2077685.24	164671098728	166896919	4034801288.8
	LSG	2594347.53	205620201844	165284005	4034892245.3
	MUNO	1775894.26	140752051428	134714877	4033907686.7
	Tobit 2	14732304.96	1.1676383E12	27125335.76	4035161853.7

The net change in other fixed assets is estimated in four steps. First (in section 7.6.1), we investigate whether firms make a positive change in other fixed assets. This is done using a logistic model with which we find the probability that the net change is positive. Second (in section 7.6.2), we investigate those firms that have made negative net changes in other fixed assets or have decided not to make such negative changes. This is done using a logistic model with which we find the probability that the net change is negative. Third (in section 7.6.3), we estimate the level of the positive net change in other fixed assets. Finally (in section 7.6.4), we estimate the level of the negative net change in other fixed assets. In both latter steps, we use the Huber-Schweppes robust estimation method.

7.6.1 The Probability that the Net Change in Other Fixed Assets is Positive

A look at the database indicates that 25951 firms (of a total of 79257 firms) have made a positive net change in other fixed assets. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.6)\quad \begin{aligned} &DPdofa_i = 1 \qquad &&if \quad dofa_i > 0 \\ &DPdofa_i = 0 \qquad &&if \quad dofa_i \le 0 \end{aligned}$$

Hence, the dependent variable can take only two values: 1 if firms make a positive net change and 0 otherwise. Table 6 concludes the estimated logistic model with normal function. Let us now interpret the results in Table 6a. As the value of explanatory variables changes, the value of the index $DPdofa_i$ varies over the real number line. The larger the value of $DPdofa_i$ the greater the incentive for firms *i* to choose the option $dofa_i > 0$. Thus, the greater the value of $DPdofa_i$, the greater will be P_i, that is the probability that firms *i* choose the option $dofa_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $dofa_i > 0$ increases with the change in cash flow, the change in the utilization of tax rules regarding allocations to periodical reserves (the increase is decreasing), and the possible investment in machinery and equipment, or in buildings. The real interest rate affects the probability for positive net changes in other fixed assets positively.

(b) The probability that firms choose the option $dofa_i > 0$ decreases by the indicator of whether or not firms reach their dividend constraint faster.

(c) Moreover, the magnitude of the probability that firms choose the option $dofa_i > 0$ can be different for firms that are closed firms, public firms, or firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $dofa_i > 0$ is higher for firms with a market concentration index close to 1. The probability is lower for firms with high market shares.

7.6.2 The Probability that Investment in Buildings is Negative

A look at the database indicates that 18146 firms (of a total of 53306 firms) have made a negative net change in other fixed assets. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$
(7.7) \quad \begin{array}{lll} DNdofa_i = 1 & if & dofa_i < 0 \\ DNdofa_i = 0 & if & dofa_i = 0 \end{array}
$$

Hence, the dependent variable can take only two values: 1 if firms make a negative net change and 0 otherwise. Table 6b concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results. As the value of explanatory variables change, the value of the index $DNdofa_i$ varies over the real number line. The larger the value of $DNdofa_i$ the greater the incentive for firms *i* to choose the option $dofa_i < 0$. Thus, the greater the value of $DNdofa_i$, the greater will be P_i, the probability that firms *i* choose the option $dofa_i < 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $dofa_i < 0$ increases with the change in cash flow and the indicator of whether or not firms reach their dividend constraint faster. The real interest rate and the possible investment in machinery and equipment, or in buildings, also increase this probability.

(b) The probability that firms choose the option $dofa_i < 0$ decreases with the change in firms' utilization of their allocations to periodical reserves (the decrease is increasing).

(c) Moreover, the magnitude of the probability that firms choose the option $dofa_i < 0$ can be different for firms that are closed firms, public firms, or firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $dofa_i < 0$ is higher for firms with a market concentration index close to 1. The probability is also higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. The frequency table indicates that the models in section 7.6.1 and 7.6.2 predict that 23.03 percent of firms will make a negative net change in other fixed assets, 32.76 percent of firms will make a positive net change in other fixed assets, and 44.21 percent will not make any change at all. The corresponding frequencies for the observed responses are 22.90

percent, 32.74 percent, and 44.36 percent, respectively. However, it is also evident that there is discordance between the predicted probabilities and the observed responses.

7.6.3 The Level of the Positive Net Change in Other Fixed Assets

Influence diagnostics show that 4.03 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.05 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in other fixed assets. Table 6c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in other fixed assets increases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the change in firms' utilization of their allocations to periodical reserves, and the possible investment in buildings (at an increasing rate).

(b) The net change in other fixed assets decreases with the possible investment in machinery and equipment as well as with the real interest rate.

(c) Moreover, the net change in other fixed assets is lower for those firms that are closed firms and for firms that are located in rural districts. On the other hand, the net change in other fixed assets is higher for public firms and for firms that are located in large cities.

(d) Finally, the net change in other fixed assets is lower for firms with a market concentration index close to 1. This means that the net change in other fixed assets is higher for firms that operate in more competitive markets. Meanwhile, the net change in other fixed assets increases when the market share of firms increases.

7.6.4 The Level of the Negative Net Change in Other Fixed Assets

Influence diagnostics show that 2.68 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.71 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality, namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in other fixed assets. Table 6d concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The negative net change in other fixed assets increases with the change in firms' utilization of their allocations to periodical reserves (at an increasing rate), the possible investment in machinery and equipment, and the real interest rate.

(b) The negative net change in other fixed assets decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, and whether firms invest in buildings.

(c) Moreover, the negative net change in other fixed assets is lower for firms that are closed firms, public firms, and firms that are located in large districts. On the other hand, the negative net change in other fixed assets is higher for firms that are located in rural areas.

(d) Finally, the net change in other fixed assets is higher for firms with a market concentration index close to 1. This means that the net change in other fixed assets is lower for firms that operate in more competitive markets. Further, the negative net change in other fixed assets decreases when the market share of firms increases.

7.7 Estimating the Net Change in Current Assets

Let us begin by looking at the predictive ability when estimating the net change in current assets.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dca		8093669.95	641479998924	1090458896	
pdca	HS	2862281.33	226855831152	229821137	931529870.7

Motivated by the small number of "zeros" for this variable (about 1.4 percent of all observations), we use only the Huber-Schweppes robust method in our estimation and outline the results for the HS method.

7.7.1 Estimation Results for the Net Change in Current Assets

Influence diagnostics show that 2.93 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.76 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in current assets. Table 7a concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in current assets increases with the economic depreciation of machinery and equipment (the increase is increasing), net investment in machinery and equipment, economic depreciation of buildings (a decreasing increase), and net investment in buildings (the increase is increasing).

(b) The net change in current assets decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, net sales of machinery and equipment, the net change in current liabilities, the change in firms' utilization of their allocations to periodical reserves, the possible investment in buildings (at a decreasing rate), and the change in gross national product.

(c) Moreover, the net change in current assets is lower for those firms that are closed firms, and for firms that are located in rural districts. On the other hand, the net change in current assets is higher for public firms and for firms that are located in large cities.

(d) Finally, the net change in current assets is higher for firms with a market concentration index close to 1. This means that the net change in current assets is lower for firms that operate in more competitive markets. Meanwhile, the net change in current assets decreases when the market share of firms increases.

7.8 Estimating the Net Change in Long-Term Liabilities

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate the net change in long-term liabilities.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dll		16822279.61	1.3332834E12	3523901883	
pdll	LLN	1449731.78	114901391510	6807263.00	3523838278.6
	LLG	1460229.75	115733429164	6807826.55	3523837651.4
	LSN	2910431.93	230672103498	68459394.77	3522233845.6
	LSG	3099542.44	245660435103	67597981.17	3522105457.3
	MUNO	2790251.10	221146931091	68120452.41	3522356369.7
	Tobit 2	12654858.50	1.0029861E12	16611810.68	3523926957.7

As is evident, the Tobit method gives the best prediction of firms' net change in long-term liabilities. However, the simulation results indicate that this method is not best suited to the simulation model. Instead, the LLN method gives a better prediction. In what follows, we will outline the estimation results for the LLN method.

The net change in long-term liabilities is estimated in two steps. First (in section 7.8.1), we investigate whether firms make a positive or negative net change in long-term liabilities. This is done using a logistic model with which we find the probability that firms make any net change in long-term liabilities. Second (in section 7.8.2), we estimate the level of the net change in long-term liabilities.

7.8.1 The Probability that the Net Change in Long-Term Liabilities is Either Positive or Negative

A look at the database indicates that 56852 firms (of a total of 79257 firms) have made a positive net change in long-term liabilities. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.8)\quad \begin{aligned} Ddll &= 1 \qquad\qquad if \qquad dll_i \neq 0 \\ Ddll &= 0 \qquad\qquad if \qquad dll_i = 0 \end{aligned}$$

Hence, the dependent variable can take only two values: 1 if firms invest (either positive or negative) in buildings and 0 otherwise. Table 8 concludes the estimated logistic model with a cumulative logistic distribution function. Let us now interpret the results in Table 8a. As the value of explanatory variables change, the value of the index $Ddll$ varies. The larger the value of $Ddll$ the greater the incentive for firms *i* to choose the option $dll_i \neq 0$. Thus, the greater the value of $Ddll$, the greater will be P_i, the probability that firms *i* choose the option $dll_i \neq 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $dll_i \neq 0$ increases with the change in cash flow, the indicator whether or not firms reach their dividend constraint faster, the change in the utilization of tax rules regarding allocations to periodical reserves (at an increasing rate), the possible net investment in machinery and equipment, the possible net investment in buildings, and the possible net change in other fixed assets.

(b) The probability that firms choose the option $dll_i \neq 0$ decreases with the real interest rate.

(c) Moreover, the magnitude of the probability that firms choose the option $dll_i \neq 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $dll_i \neq 0$ is higher for firms with a market concentration index close to 1. The probability is also higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have made any net change in long-term liabilities is 71.83 percent, while the frequency for predicting that firms will have made net changes in long-term liabilities is 28.17 percent. These probabilities coincide with the observed responses (71.73 percent and 28.27 percent, respectively). However, it is also evident that there is 27.2 percent discordance between the predicted probabilities and the observed responses.

7.8.2 The Level of the Positive and Negative Net Change in Long-Term Liabilities

Influence diagnostics show that 1396 observations have severe leverage points - points that are outliers in the design matrix. Further, 28 observations have severe influence on the response variable. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in long-term liabilities. Table 8b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in long-term liabilities increases with the indicator of whether or not firms reach their dividend constraint faster, the change in the utilization of tax rules regarding allocations to periodical reserves (the increase is increasing), the possible net investment in machinery and equipment, the possible net change in other fixed assets, and the real interest rate.

(b) The net change in long-term liabilities decreases with the change in cash flow and the possible net investment in buildings.

(c) Moreover, the net change in long-term liabilities is lower for those firms that are closed firms and for firms that are located in rural districts. On the other hand, the net change in long-term liabilities is higher for firms that are public firms and firms that are located in large cities.

(d) Finally, the net change in long-term liabilities is higher for firms with a market concentration index close to 1. This means that the net change in long-term liabilities is lower for firms that operate in more competitive markets. On the other hand, the net change in long-term liabilities is lower for firms that have high market shares.

7.9 Estimating the Net Change in Current Liabilities

Let us begin by looking at the predictive ability when estimating the net change in current liabilities.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dcl		7839932.89	621369560928	871769194	
dcl	HS	7389177.18	585644015925	785209320	370119301.48
	LLG	12025873.51	953134657119	841631319	413578505.03
	LSG	-257283.41	-20391511440	2020578511	1810705578.1

Motivated by the small number of “zeros” for this variable (about 2.2 percent of all observations), we primarily use the Huber-Schweppes robust method in our estimation and, since it gives the best prediction, outline the results for the HS method.

7.9.1 Estimation Results for the Net Change in Current Liabilities

Influence diagnostics show that 2.93 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.76 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in current liabilities. Table 9a concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in current liabilities increases with the net investment in machinery and equipment, economic depreciation of buildings (a decreasing increase), net investment in buildings (the increase is increasing), and the net change in current assets.

(b) The net change in current liabilities decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, economic depreciation of machinery and equipment (the decrease is decreasing), net sales of machinery and equipment, the change in firms' utilization of their allocations to periodical reserves, the possible investment in buildings (at an increasing rate), and the change in gross national product.

(c) Moreover, the net change in current liabilities is lower for those firms that are closed firms, public firms, and firms that are located in rural districts. On the other hand, the net change in current liabilities is higher for firms that are located in large cities.

(d) Finally, the net change in current liabilities is lower for firms with a market concentration index close to 1. This means that the net change in current liabilities is higher for firms that operate in more competitive markets. Meanwhile, the net change in current liabilities decreases when the market share of firms increases.

7.10 Estimating the Net Change in Share Capital

Let us begin by looking closer at the predictive ability of different methods that we have used to estimate the net change in share capital.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dsc		1836259.98	145536457269	218866630	
pdsc	LLN	284636.85	22559462788	2869634.76	218838233.32
	LLG	285105.82	22596631848	2838789.79	218839574.83
	LSN	90375.06	7162856486	32201264.52	210705195.01
	LSG	110033.85	8720953132	31972705.77	210701400.39
	MUNO	-63165.27	-5006289667	33442752.48	211024629.20
	Tobit 2	1010988.30	80127899693	2877860.80	218848929.65

As is evident, the Tobit 2 method gives the best prediction of the change in share capital. However, our simulations indicate that this method is not best suited to the simulation model. Instead, the LSG method gives a better prediction than the Tobit 2 method. In what follows, we will outline the estimation results for the LSG method.

The net change in share capital is estimated in four steps. First (in section 7.10.1), we investigate whether firms make a positive change to share capital. This is done using a logistic model with which we find the probability that the net change is positive. Second (in section 7.10.2), we investigate those firms that have made negative net changes to share capital or have decided not to make such negative changes. This is done using a logistic model with which we find the probability that the net change is negative. Third (in section 7.10.3), we estimate the level of the positive net change in share capital. Finally (in section 7.10.4), we estimate the level of the negative net change in share capital. In both latter steps, we use the Huber-Schweppes robust estimation method.

7.10.1 The Probability that the Net Change in Share Capital is Positive

A look at the database indicates that 15502 firms (of a total of 79257 firms) have made a positive net change in share capital. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.9)\quad \begin{array}{lll} DPdsc_i = 1 & if & dsc_i > 0 \\ DPdsc_i = 0 & if & dsc_i \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms make a positive net change and 0 otherwise. Table 10 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 10a. As the value of explanatory variables change, the value of the index $DPdsc_i$ varies over the real number line. The larger the value of $DPdsc_i$ the greater the incentive for firms *i* to choose the option $dsc_i > 0$. Thus, the greater the value of $DPdsc_i$, the greater will be P_i, that is the probability that firms *i* choose the option $dsc_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $dsc_i > 0$ increases with the change in the utilization of tax rules regarding allocations to periodical reserves (the increase is decreasing), the possible investment in machinery and equipment or in buildings, as well as the possible net change in other fixed assets. The real interest rate affects the probability of positive net changes in share capital positively.

(b) The probability that firms choose the option $dsc_i > 0$ decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, and the possible net change in long-term liabilities.

(c) Moreover, the magnitude of the probability that firms choose the option $dsc_i > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $dsc_i > 0$ is higher for firms with a market concentration index close to 1. The probability is lower for firms with high market shares.

7.10.2 The Probability that Investment in Buildings is Negative

A look at the database indicates that 3717 firms (of a total of 63755 firms) have made a negative net change in share capital. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.10)\quad \begin{array}{lll} DNdsc_i = 1 & if & dsc_i < 0 \\ DNdsc_i = 0 & if & dsc_i = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms make a negative net change and 0 otherwise. Table 10b concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results. As the value of explanatory variables change, the value of the index $DNdsc_i$ varies over the real number line. The larger the value of $DNdsc_i$ the greater the incentive for firms *i* to choose the option $dsc_i < 0$. Thus, the greater the value of $DNdsc_i$, the greater will be P_i, the probability that firms *i* choose the option $dsc_i < 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $dsc_i < 0$ increases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the possible investment in buildings, the possible net change in other fixed assets, and the possible net change in long-term liabilities.

(b) The probability that firms choose the option $dsc_i < 0$ decreases with the change in firms' utilization of their allocations to periodical reserves (the decrease is increasing). The real interest rate and the possible investment in machinery and equipment also decrease this probability.

(c) Moreover, the magnitude of the probability that firms choose the option $dsc_i < 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $dsc_i < 0$ is lower for firms with a market concentration index close to 1. The probability is higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. The frequency table indicates that the models in section 7.10.1 and 7.10.2 predict that 4.69 percent of firms will make a negative net change to share capital, 19.56 percent of firms will make a positive net change to share capital, and 75.75 percent will not make any change at all. The corresponding frequencies for the observed responses are 4.65 percent, 19.71 percent, and 75.64 percent, respectively. However, it is also evident that there is discordance between the predicted probabilities and the observed responses.

7.10.3 The Level of the Positive Net Change in Share Capital

Influence diagnostics show that 2.15 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.26 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in share capital. Table 10c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in share capital increases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the change in firms' utilization of their allocations to periodical reserves and the possible investment in buildings (at a decreasing rate), the possible net change in other fixed assets, the possible net change in long-term liabilities, and the real interest rate.

(b) The net change in share capital decreases with the possible investment in machinery and equipment.

(c) Moreover, the net change in share capital is lower for those firms that are closed firms and for firms that are located in rural districts. On the other hand, the net change in share capital is higher for public firms and for firms that are located in large cities.

(d) Finally, the net change in share capital is higher for firms with a market concentration index close to 1. This means that the net change in share capital is lower for firms that operate in more competitive markets. Meanwhile, the net change in share capital increases when the market share of firms increase.

7.10.4 The Level of the Negative Net Change in Share Capital

Influence diagnostics show that 5.11 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 1.99 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net change in share capital. Table 10d concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The negative net change in share capital increases with the change in firms' utilization of their allocations to periodical reserves and the real interest rate.

(b) The negative net change in share capital decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the possible investment in machinery and equipment or in buildings, and the possible net change in other fixed assets or in long-term liabilities.

(c) Moreover, the negative net change in share capital is higher for firms that are closed firms. On the other hand, the negative net change in share capital is lower for public firms and firms that are located in large or rural districts.

(d) Finally, the net change in share capital is higher for firms with a market concentration index close to 1. This means that the net change in share capital is lower for firms that operate in more competitive markets. Further, the negative net change in share capital decreases when the market share of firms increases.

7.11 Estimating the Net Changes in Restricted Reserves

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate the net changes in restricted reserves.

Variable	Method	Mean	Sum	Std. Dev.	MSE
drr		4937664.41	391344468325	900834626	
pdrr	LLN	238730.20	18921039110	14626104.65	900623493.27
	LLG	255213.98	20227494096	14634401.95	900619068.05
	LSN	383433.74	30389808190	15326474.81	900614547.39
	LSG	403282.40	31962953513	15364359.58	900617538.51
	MUNO	214995.26	17039879552	15403761.54	900646184.12

As is evident, none of the estimation methods provide a good prediction. As will be evident in the next section, our simulations indicate that LLG method gives a satisfactory prediction of drr **.** In what follows, we will outline the estimation results for the LLG method.

The net change in restricted reserves is estimated in two steps. First (in section 7.11.1), we investigate whether firms make positive or negative net changes to restricted reserves. This is done using a logistic model with which we find the probability that firms make any net changes to restricted reserves. Second (in section 7.11.2), we estimate the level of the net changes in restricted reserves.

7.11.1 The Probability that the Net Changes in Restricted Reserves are Either Positive or Negative

A look at the database indicates that 35649 firms (of a total of 79257 firms) have made positive net changes to restricted reserves. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.11)}\qquad \begin{array}{lll} Ddrr = 1 & \quad if \quad & drr_i \neq 0 \\ Ddrr = 0 & \quad if \quad & drr_i = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms invest (either positive or negative) in buildings and 0 otherwise. Table 11 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 11a. As the value of explanatory variables change, the value of the index $Ddrr$ varies. The larger the value of $Ddrr$ the greater the incentive for firms *i* to choose the option $drr_i \neq 0$. Thus, the greater the value of $Ddrr$, the greater will be P_i, the probability that firms *i* choose the option $drr_i \neq 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $drr_i \neq 0$ increases with the possible net investment in machinery and equipment, the possible net investment in buildings, the possible net change in other fixed assets, the possible net change in long-term liabilities, the possible net change in share capital, and the real interest rate.

(b) The probability that firms choose the option $drr_i \neq 0$ decreases with the change of firms' dividend policy so that they come closer to the legal constraint on dividend payments (the decrease is increasing).

(c) Moreover, the magnitude of the probability that firms choose the option $drr_i \neq 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $drr_i \neq 0$ is lower for firms with a market concentration index close to 1. The probability is also lower for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have made any net changes to restricted reserves is 45.3 percent, while the frequency for predicting that firms will have made net changes to long-term liabilities is 54.7 percent. These probabilities coincide with the observed responses (44.98 percent and 55.02 percent, respectively). However, it is also evident that there is 27.2 percent discordance between the predicted probabilities and the observed responses.

7.11.2 The Level of the Positive and Negative Net Changes in Restricted Reserves

Influence diagnostics show that 1062 observations have severe leverage points - points that are outliers in the design matrix. Further, 28 observations have severe influence on the response variable. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net changes in restricted reserves. Table 11b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net change in restricted reserves increases with the possible net investment in machinery and equipment, the possible net change in other fixed assets, the possible net change in share capital, and the real interest rate.

(b) The net change in restricted reserves decreases with the change of firms' dividend policy so that they come closer to the legal constraint on dividend payments (the decrease is decreasing), the possible net investment in buildings, and the possible net change in long-term liabilities.

(c) Moreover, the net change in restricted reserves is lower for those firms that are closed firms and for firms that are located in rural districts. On the other hand, the net change in restricted reserves is higher for public firms and firms that are located in large cities.

(d) Finally, the net change in restricted reserves is higher for firms with a market concentration index close to 1. This means that the net change in restricted reserves is lower for firms that operate in more competitive markets. On the other hand, the net change in restricted reserves is lower for firms that have high market shares.

7.12 Estimating Operating Income Before Depreciation

Let us begin by looking at the predictive ability when estimating operating income before depreciation.

Variable	Method	Mean	Sum	Std. Dev.	MSE
OIBD		6550522.15	519174733773	75916333.37	
pOIBD	HS	5312982.55	421091058308	51678146.59	53322832.947

Motivated by the small number of “zeros” for this variable (about 2.7 percent of all observations), we use the Huber-Schweppes robust method in our estimation and outline the results for the HS method.

7.12.1 Estimation Results for Operating Income Before Depreciation

Influence diagnostics show that 3.01 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.97 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of operating income before depreciation. Table 12a concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Operating income before depreciation increases with the level of current assets, the level of machinery and equipment, the net investment in machinery and equipment, economic depreciation of machinery and equipment (at a decreasing rate), the level of buildings, the net investment in buildings, economic depreciation of buildings (at an increasing rate), the net change in current assets (at a decreasing rate), the change in the utilization of tax rules regarding allocations to periodical reserves (the increase is decreasing), and the change in cash flow (at an increasing rate).

(b) Operating income before depreciation decreases with the level of current liabilities, the net sales of machinery and equipment, the net change in current liabilities, and the change in gross national product.

(c) Moreover, operating income before depreciation is lower for closed firms and public firms, as well as for firms that are located in rural districts. On the other hand, operating income before depreciation is higher for firms that are located in large cities.

(d) Finally, operating income before depreciation is higher for firms with a market concentration index close to 1. This means that operating income before depreciation is lower for firms that operate in more competitive markets. Meanwhile, operating income before depreciation decreases when the market share of firms increases.

7.13 Estimating Firms' Financial Income

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate firms' financial income.

Variable	Method	Mean	Sum	Std. Dev.	MSE
FI		6986190.74	553704519650	148867201	
pFI	LLN	4367514.80	346156120724	181969258	214714482.93
	LLG	4236290.02	335755637941	179752661	214702313.06
	Tobit 1	4531258.90	359133986802	46269929.61	137601904.08

The three estimation methods used give similar predictions of firms' financial income. However, the simulation results indicate that the LLG method is best suited to the simulation model. In what follows, we will outline the estimation results for this method.

Firms' financial income is estimated in two steps. First (in section 7.13.1), we investigate whether firms have any financial income. This is done using a logistic model with which we find the probability that firms' financial income is positive. Second (in section 7.13.2), we estimate the level of firms' financial income for those firms that have such income. This is done with the Huber-Schweppes robust estimation method.

7.13.1 The Probability that Firms' Financial Income is Positive

A look at firms' financial income indicates that 69360 firms (of 79257 firms) have financial income. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\begin{aligned} &DFI = 1 \qquad if \qquad FI_i > 0 \\ &DFI = 0 \qquad if \qquad FI_i \le 0 \end{aligned} \tag{7.12}$$

Hence, the dependent variable can take only two values: 1 if firms have financial income and 0 if firms do not have any financial income. Table 13 concludes the estimated logistic model with a complementary log-log distribution function. Let us now interpret the results in Table 13a. As the value of explanatory variables change, the value of the index DFI varies. The larger the value of FI the greater the incentive for firms *i* to choose the option $FI_i > 0$. Thus, the greater the value of DFI, the greater will be P_i, the probability that firms *i* choose the option $FI_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $FI_i > 0$ increases with the net investment in buildings, the economic depreciation of machinery and equipment (a decreasing increase), the net sales of machinery and equipment, the net investment in machinery and equipment (the increase is decreasing), the economic depreciation of buildings (at a decreasing rate), the net change in current assets (at an increasing rate), the net change in other fixed assets, the level of other fixed assets, the level of current assets, the level of buildings, and the real interest rate.

(b) The probability that firms choose the option $FI_i > 0$ decreases with the level of machinery and equipment.

(c) Moreover, the magnitude of the probability that firms choose the option $FI_i > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $FI_i > 0$ is higher for firms with a market concentration index close to 1. The probability is lower for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have any financial income is 12.41 percent, while the frequency for predicting positive financial income is 87.59 percent. This coincides approximately with the observed responses (12.49 percent and 87.51 percent, respectively).

7.13.2 The Level of Firms' Financial Income

Influence diagnostics show that 400 observations have severe influence on the response variable. Further, 2235 observations have problems with leverage points. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of financial income. Table 13b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Firms' financial income increases with the net investment in buildings, the net sales of machinery and equipment, the net change in current assets (the increase is decreasing), the net change in other fixed assets, the level of other fixed assets, the level of current assets, the level of buildings, and the real interest rate.

(b) Firms' financial income decreases with the economic depreciation of machinery and equipment (at an increasing rate), the net investment in machinery and equipment (at an increasing rate), the economic depreciation of buildings (at an increasing rate), and the level of machinery and equipment.

(c) Moreover, the level of firms' financial income is higher for those firms that are closed or public firms. On the other hand, the level of firms' financial income is lower for firms that are located in large cities and rural districts.

(d) Finally, firms' financial income is lower for firms with a market concentration index close to 1. This means that firms' financial income is higher for firms that operate in more competitive markets. Firms' financial income is lower for firms that have high market shares.

7.14 Estimating Firms' Financial Expenses

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate firms' financial expenses.

Variable	Method	Mean	Sum	Std. Dev.	MSE
FE		4385541.59	347584870063	55500816.84	
pFE	LLN	3866932.27	306481450888	184547082	189070374.32
	LLG	3793435.50	300656317453	184482554	189245863.40
	Tobit 1	3593695.07	284825490255	29052503.36	44338332.348

The three estimation methods used give similar predictions of firms' financial expenses. However, the simulation results indicate that the LLG method is best suited to the simulation model. In what follows, we will outline the estimation results for this method.

Firms' financial expenses are estimated in two steps. First (in section 7.14.1), we investigate whether firms have any financial expenses. This is done using a logistic model with which we find the probability that firms' financial expenses are positive. Second (in section 7.14.2), we estimate the level of firms' financial expenses for those firms that have such expenses. This is done with the Huber-Schweppes robust estimation method.

7.14.1 The Probability that Firms' Financial Expenses are Positive

A look at firms' financial expenses indicates that 69360 firms (of 79257 firms) have financial expenses. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.13)} \quad \begin{array}{lll} DFE = 1 & if & FE_i > 0 \\ DFE = 0 & if & FE_i \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms have financial expenses and 0 if firms do not have any financial expenses. Table 14 concludes the estimated logistic model with a complementary log-log distribution function. Let us now interpret the results in Table 14a. As the value of explanatory variables change, the value of the index DFE varies. The larger the value of DFE the greater the incentive for firms i to choose the option $FE_i > 0$. Thus, the greater the value of DFE, the greater will be P_i, the probability that firms i choose the option $FE_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $FE_i > 0$ increases with the net investment in buildings, the economic depreciation of machinery and equipment, the net investment in machinery and equipment, the economic depreciation of buildings, the level of buildings, the level of long-term liabilities, the level of current assets, the level of current liabilities, the net change in current liabilities, and the net change in long-term liabilities.

(b) The probability that firms choose the option $FE_i > 0$ decreases with the net sales of machinery and equipment, the level of other fixed assets, the level of machinery and equipment, the net change in current assets, the net change in other fixed assets, and the real interest rate.

(c) Moreover, the magnitude of the probability that firms choose the option $FE_i > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $FE_i > 0$ is higher for firms with a market concentration index close to 1. The probability is higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have any financial expenses is 12.41 percent, while the frequency for predicting positive financial expenses is 87.59 percent. This coincides approximately with the observed responses (12.43 percent and 87.57 percent, respectively).

7.14.2 The Level of Firms' Financial Expenses

Influence diagnostics show that there are 583 observations that have severe influence on the response variable. Further, 2171 observations have problems with leverage points. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of financial expenses. Table 14b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Firms' financial expenses increase with the net investment in buildings, the economic depreciation of machinery and equipment, the economic depreciation of buildings, the level of other fixed assets, the level of machinery and equipment, the level of buildings, the level of long-term liabilities, the level of current liabilities, the net change in current liabilities, the net change in other fixed assets, the net change in long-term liabilities, and the real interest rate.

(b) Firms' financial expenses decrease with the net sales of machinery and equipment, the net investment in machinery and equipment, the level of current assets, and the net change in current assets.

(c) Moreover, the level of firms' financial expenses is higher for those firms that are closed or public firms as well as for firms that are located in large cities and rural districts.

(d) Finally, firms' financial expenses are lower for firms with a market concentration index close to 1. This means that firms' financial expenses are higher for firms that operate in more competitive markets. Firms' financial expenses are lower for firms that have high market shares.

7.15 Estimating Tax Depreciation of Machinery and Equipment

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate the tax depreciation of machinery and equipment.

Variable	Method	Mean	Sum	Std. Dev.	MSE
$TDEP^{MA}$		2607266.98	206644158878	45776744.55	
$pTDEP^{MA}$	LLN	2267581.07	179721672851	31669923.98	30330545.27
	LLG	2364062.75	187368521260	47329077.25	14227763.65
	Tobit 1	2180559.54	172824607488	17205036.19	36574793.38

As is evident, the LLG method gives the best prediction of firms' tax depreciation of machinery and equipment. However, our simulations indicate that this method is not best suited to the simulation model. Instead, the Tobit 1 method gives a better prediction and, in what follows, we will outline the estimation results for the Tobit 1 method.

7.15.1 The Tobit Model for Positive Tax Depreciation of Machinery and Equipment

Let us generate a dependent variable, which has the following characteristics

$$(7.14)\quad \begin{array}{lll} LOWER = EDEP_i^{MA} & if & EDEP_i^{MA} > 0 \\ LOWER = . & if & EDEP_i^{MA} \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: $EDEP_i^{MA}$ if firms carry out tax depreciation of machinery and equipment, and missing values otherwise. Table 15 concludes the estimated Tobit 1 model with a logistic distribution function. Let us now interpret the results in Table 15a. Direct interpretation of the various regression coefficients given in Table 15a is not easy. But from the estimated coefficients one can assess the marginal effects. As Table 15a shows:

(a) An increase in net sales of machinery and equipment, economic depreciation of machinery and equipment (the increase is decreasing), net investment in machinery and equipment (at an increasing rate), and the change in the utilization of tax rules regarding allocations to periodical reserves (at an increasing rate) will increase firms' tax depreciation of machinery and equipment.

(b) An increase in the change in cash flow, the change of firms' dividend policy so that they come closer to the legal constraint on dividend payments, and the real interest rate, will decrease firms' tax depreciation of machinery and equipment.

(c) Moreover, tax depreciation of machinery and equipment is higher for closed firms and firms that are located in rural districts. On the other hand, tax depreciation is lower for public firms and firms that are located in large cities.

(d) Finally, tax depreciation of machinery and equipment is lower for firms with a market concentration index close to 1. This means that tax depreciation of machinery and equipment is higher for firms that operate in more competitive markets. Tax depreciation of machinery and equipment is lower for firms that have high market shares.

7.16 Estimating Reversals from Periodical Reserves

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate reversals from periodical reserves.

Variable	Method	Mean	Sum	Std. Dev.	MSE
zpf		244860.44	19406904251	29191276.08	
pzpf	LLN	41124.24	3259383803	2600970.25	27118183.15
	LLG	41092.92	3256901193	2600952.31	27118216.89
	Tobit 1	37426.74	2966331457	5361914.26	26400777.47

The three estimation methods used give similar predictions of firms' reversals from periodical reserves. However, the simulation results indicate that the LLG method is best suited to the simulation model. In what follows, we will outline the estimation results for this method.

The reversals from periodical reserves are estimated in two steps. First (in section 7.16.1), we investigate whether firms have made any reversals from periodical reserves. This is done using a logistic model with which we find the probability that the reversals from periodical reserves are positive. Second (in section 7.16.2), we estimate the level of the reversals from periodical reserves for those firms that show any change at all. This is done with the Huber-Schweppes robust estimation method.

7.16.1 The Probability that Reversals from Periodical Reserves are Positive

A look at the reversals from periodical reserves indicates that 6917 firms (of 79257 firms) have made changes to periodical reserves. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$(7.15)\quad \begin{array}{lll} DZPF = 1 & if & ZPF_i > 0 \\ DZPF = 0 & if & ZPF_i \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms report reversals from periodical reserves and 0 if firms do not report any change in reserves. Table 16 concludes the estimated logistic model with a complementary log-log distribution function. Let us now interpret the results in Table 16a. As the value of explanatory variables change, the value of the index DFI varies. The larger the value of $DZPF$ the greater the incentive for firms *i* to choose the option $ZPF_i > 0$. Thus, the greater the value of $DZPF$, the greater will be P_i, the probability that firms i choose the option $ZPF_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $ZPF_i > 0$ increases with the change in the utilization of tax rules regarding allocations to periodical reserves (at a decreasing rate), the allocations to periodical reserves, and the possible tax depreciation of machinery and equipment.

(b) The probability that firms choose the option $ZPF_i > 0$ decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, and the real interest rate.

(c) Moreover, the magnitude of the probability that firms choose the option $ZPF_i > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $ZPF_i > 0$ is lower for firms with a market concentration index close to 1. The probability is higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have made any reversals from periodical reserves is 91.29 percent, while the frequency for predicting positive changes in periodical reserves is 8.71 percent. This coincides approximately with the observed responses (91.27 percent and 8.73 percent, respectively).

7.16.2 The Level of Reversals from Periodical Reserves

Influence diagnostics show that there are 8 observations that have severe influence on the response variable. Further, 145 observations have problems with leverage points. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the reversals from periodical reserves. Table 16b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The reversals from periodical reserves increase with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the allocations to periodical reserves, and the possible tax depreciation of machinery and equipment.

(b) The reversals from periodical reserves decrease with the change in the utilization of tax rules regarding allocations to periodical reserves and the real interest rate.

(c) Moreover, the level of the reversals from periodical reserves is lower for closed firms and firms that are located in rural districts. On the other hand, the level of the reversals from periodical reserves is higher for public firms and firms that are located in large cities.

(d) Finally, the reversals from periodical reserves are higher for firms with a market concentration index close to 1. This means that reversals from periodical reserves are lower for firms that operate in more competitive markets. The reversals from periodical reserves are higher for firms that have high market shares.

7.17 Estimating the Net Change in Other Untaxed Reserves

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate the net change in other untaxed reserves.

Variable	Method	Mean	Sum	Std. Dev.	MSE
dour		-56919.11	-4511238208	32177604.04	
pdour	LLN	-64192.31	-5087690167	1075385.15	31791840.06
	LLG	-62167.17	-4927183533	1063217.39	31791908.49
	LSN	-47496.78	-3764452149	3543874.41	30965566.24
	LSG	-45355.03	-3594703456	3545808.76	30962632.73
	MUNO	-46371.94	-3675300733	3862075.22	30993709.69
	Tobit 2	71116.78	5636502368	234487.50	32177947.47

As is evident, the LLG method gives the best prediction of firms' net change in other untaxed reserves. However, our simulations indicate that this method is not best suited to the simulation model. Instead, the MUNO method provides a better prediction and, in what follows, we will outline the estimation results for the LLG method.

The net change in other untaxed reserves is estimated in two steps. First (in section 7.17.1), we use a multinomial logit model to investigate whether the net change in other untaxed reserves is positive, equal to zero, or negative. Second (in section 7.17.2), we estimate the level of the positive net change using the Huber-Schweppes robust estimation method. Third (in section 7.17.3), we estimate the level of the negative net change using the Huber-Schweppes robust estimation method.

7.17.1 The Probability that Firms Make Net Changes in Other Untaxed Reserves

A look at the net change in other untaxed reserves made by firms indicates that 5143 firms account for positive net changes, 15386 firms account for negative net changes (devaluation), while 58728 firms do not account for any net change in other untaxed reserves. To capture this fact we generate a dependent variable, which is dichotomous in nature, as

$$
\begin{aligned}
&Ddour = 1 \quad && if \quad dour_i < 0 \\
(7.16)\quad &Ddour = 2 \quad && if \quad dour_i = 0 \\
&Ddour = 3 \quad && if \quad dour_i > 0
\end{aligned}
$$

Hence, the dependent variable can take three values: 1 if the net change is negative, 2 if firms do not make any net changes, and 3 if the net change is positive. Table 17a concludes the estimated multinomial logit model with a complementary log-log function. As mentioned earlier, the coefficients in this model are difficult to interpret and we will not do that here. Meanwhile, marginal effects can be computed within the statistical program package SAS.

7.17.2 The Level of the Positive Net Change in Other Untaxed Reserves

Influence diagnostics show that 0.23 percent of the observations have severe influence on the response variable. Further, it is also evident that 7.56 percent of the observations have high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality, namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the net change in other untaxed reserves. Table 17b concludes the estimated model. Based on the sign on the estimated coefficients, we infer the following:

(a) First, the positive net change in other untaxed reserves increases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the change in the utilization of tax rules regarding allocations to periodical reserves (this increase is increasing), the possible reversals from periodical reserves, the possible tax depreciation of machinery and equipment, and the real interest rate.

(b) Moreover, the positive net change in other untaxed reserves is lower for closed firms and firms that are located in rural districts. On the other hand, the positive net change is higher for public firms and firms that are located in large cities.

(c) Finally, the positive net change in other untaxed reserves is higher for firms with a market concentration index close to 1. This means that the positive net change is lower for firms that operate in more competitive markets. The positive net change is also higher for firms that have high market shares.

7.17.3 The Level of the Negative Net Change in Other Untaxed Reserves

Influence diagnostics show that 0.86 percent of the observations have severe influence on the response variable. Further, it is also evident that 2.51 percent of the observations have leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality, namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the (net) sales of machinery and equipment. Table 17c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the negative net change in other untaxed reserves increases with the possible reversals from periodical reserves, the possible tax depreciation of machinery and equipment, and the real interest rate.

(b) Second, the negative net change in other untaxed reserves decreases with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the change in the utilization of tax rules regarding allocations to periodical reserves (this decrease is decreasing).

(c) Moreover, the negative net change in other untaxed reserves is lower for public firms and firms that are located in large cities. On the other hand, the negative net change is higher for closed firms and firms that are located in rural areas.

(d) Finally, the negative net change in other untaxed reserves is higher for firms with a market concentration index close to 1. This means that the negative net change is lower for firms that operate in more competitive markets. The negative net change is also lower for firms that have high market shares.

7.18 Estimating the Net Group Contribution

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate the net group contribution.

Variable	Method	Mean	Sum	Std. Dev.	MSE
GC		219327.03	17383202532	59654357.81	
pGC	LLN	-559550.42	-44348287478	34373368.02	51318696.84
	LLG	-576818.28	-45716886119	22861812.53	58016164.11
	LSN	-642883.41	-50953010351	48594137.87	55743827.69
	LSG	-123440.77	-9783545452	43737442.14	53362756.74

As is evident, the LSG method gives the best prediction of the net group contribution. In what follows, we will outline the estimation results for this method.

The net group contribution is estimated in four steps. First (in section 7.18.1), we investigate whether firms have a positive contribution. This is done using a logistic model with which we find the probability that the net group contribution is positive. Second (in section 7.18.2), we investigate those firms that have a negative net group contribution or that decided not to have such a negative contribution. This is done using a logistic model with which we find the probability that the net group contribution is negative. Third (in section 7.18.3), we estimate the level of the positive net group contribution. Finally (in section 7.18.4), we estimate the level of the negative net group contribution. In both latter steps, we use the Huber-Schweppes robust estimation method.

7.18.1 The Probability that the Net Group Contribution is Positive

A look at the database indicates that 6948 firms (of a total of 79257 firms) have a positive net group contribution. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.17)} \quad \begin{array}{lll} DPGC_i = 1 & \quad if & GC_i > 0 \\ DPGC_i = 0 & \quad if & GC_i \leq 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms make a positive net change and 0 otherwise. Table 18 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 18a. As the value of explanatory variables change, the value of the index $DPGC_i$ varies over the real number line. The larger the value of $DPGC_i$ the greater the incentive for firms *i* to choose the option $GC_i > 0$. Thus, the greater the value of $DPGC_i$, the greater will be P_i, that is the probability that firms *i* choose the option $GC_i > 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $GC_i > 0$ increases with financial income, financial expenses, the tax depreciation of machinery and equipment (the increase is decreasing), the economic depreciation of buildings (decreasingly), and the net change in other untaxed reserves. The change in gross national product affects the probability of a positive net group contribution positively.

(b) The probability that firms choose the option $GC_i > 0$ decreases with operating income before depreciation (at an increasing rate) and reversals from periodical reserves.

(c) Moreover, the magnitude of the probability that firms choose the option $GC_i > 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $GC_i > 0$ is lower for firms with a market concentration index close to 1. The probability is higher for firms with high market shares.

7.18.2 The Probability that Investment in Buildings is Negative

A look at the database indicates that 8356 firms (of a total of 72309 firms) have a negative net group contribution. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.18)} \quad \begin{array}{lll} DNGC_i = 1 & \quad if & GC_i < 0 \\ DNGC_i = 0 & \quad if & GC_i = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms make a negative net change and 0 otherwise. Table 18b concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results. As the value of explanatory variables change, the value of the index $DNGC_i$ varies over the real number line. The larger the value of $DNGC_i$ the greater the incentive for firms *i* to choose the option $GC_i < 0$. Thus, the greater the value of $DNGC_i$, the greater will be P_i, the probability that firms *i* choose the option $GC_i < 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $GC_i < 0$ increases with operating income before depreciation, financial income, financial expenses, and reversals from periodical reserves.

(b) The probability that firms choose the option $GC_i < 0$ decreases with tax depreciation of machinery and equipment (at an increasing rate), economic depreciation of buildings (at an increasing rate), and the net change in other untaxed reserves. The change in gross national product also decreases this probability.

(c) Moreover, the magnitude of the probability that firms choose the option $GC_i < 0$ can be different for firms that are closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $GC_i < 0$ is lower for firms with a market concentration index close to 1. The probability is higher for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. The frequency table indicates that the models in section 7.18.1 and 7.18.2 predict that 10.54 percent of firms will have a negative net group contribution, 8.77 percent of firms will have a positive net group contribution, and 80.69 percent will not have any contribution at all. The corresponding frequencies for the observed responses are 10.63 percent, 8.75 percent, and 80.62 percent, respectively. However, it is also evident that there is discordance between the predicted probabilities and the observed responses.

7.18.3 The Level of the Positive Net Group Contribution

Influence diagnostics show that 2.49 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 1.5 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net group contribution. Table 18c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The net group contribution increases with financial income, financial expenses, the tax depreciation of machinery and equipment (at a decreasing rate), the economic depreciation of buildings (the increase is decreasing), the net change in other untaxed reserves, and the change in gross national product.

(b) The net group contribution decreases with operating income before depreciation (at an increasing rate) and reversals from periodical reserves.

(c) Moreover, the net group contribution is lower for closed firms and firms that are located in rural districts. On the other hand, the net group contribution is higher for public firms and firms that are located in large cities.

(d) Finally, the net group contribution is higher for firms with a market concentration index close to 1. This means that the net group contribution is lower for firms that operate in more competitive markets. Meanwhile, the net group contribution decreases when the market share of firms increases.

7.18.4 The level of the negative net group contribution

Influence diagnostics show that 5.28 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 1.96 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of the net group contribution. Table 18d concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) The negative net group contribution increases with financial expenses, tax depreciation of machinery and equipment (at an increasing rate), economic depreciation of buildings (at an increasing rate), and reversals from periodical reserves.

(b) The negative net group contribution decreases with operating income before depreciation (at a decreasing rate), financial income, the net change in other untaxed reserves, and the change in gross national product.

(c) Moreover, the negative net group contribution is lower for public firms and firms that are located in large districts. On the other hand, the negative net group contribution is higher for closed firms and firms that are located in rural areas.

(d) Finally, the net group contribution is lower for firms with a market concentration index close to 1. This means that the net group contribution is higher for firms that operate in more competitive markets. Further, the negative net group contribution increases when the market share of firms increases.

7.19 Estimating Other Allocations

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate other allocations.

Variable	Method	Mean	Sum	Std. Dev.	MSE
OA		512504.50	40619569009	54179761.73	
pOA	LLN	352881.89	27968359660	48852422.19	45852730.91
	LLG	404352.23	32047744379	47898999.61	45434495.13
	LSN	307402.61	24363808588	59347078.06	39767888.48
	LSG	307402.61	24363808588	59347078.06	39767888.48
	MUNO	499497.95	39588709026	56171053.50	42937398.01
	Tobit 2	2155691.94	170853675860	44041160.64	38997586.94

As is evident, the multinomial logit estimation method gives a rather good prediction of other allocations**.** As will be evident in the next section, our simulations indicate that the MUNO method gives a satisfactory prediction as well**.** In what follows, we will outline the estimation results for the MUNO method.

Other allocations are estimated in two steps. First (in section 7.19.1), we use a multinomial logit model to investigate whether other allocations are positive, equal to zero, or negative. Second (in section 7.19.2), we estimate the level of the positive other allocations using the Huber-Schweppes robust estimation method. Third (in section 7.19.3), we estimate the level of the negative other allocations using the Huber-Schweppes robust estimation method.

7.19.1 The Probability that Firms Make Other Allocations

A look at firms' other allocations indicates that 21431 firms have positive other allocations, 8361 firms have negative other allocations, and 49465 firms do not have any other allocations. To capture this fact we generate a dependent variable, which is dichotomous in nature, as

$$
\begin{aligned}
&DOA = 1 \qquad && \text{if} \quad OA_i < 0 \\
(7.19) \qquad &DOA = 2 && \text{if} \quad OA_i = 0 \\
&DOA = 3 && \text{if} \quad OA_i > 0
\end{aligned}
$$

Hence, the dependent variable can take three values: 1 if other allocations are negative, 2 if firms do not make any other allocations, and 3 if other allocations are positive. Table 19a concludes the estimated multinomial logit model with a complementary log-log function. As mentioned earlier, the coefficients in this model are difficult to interpret and we will not do that here. Meanwhile, marginal effects can be computed using the statistical program package SAS.

7.19.2 The Level of the Positive Other Allocations

Influence diagnostics show that 0.56 percent of the observations have severe influence on the response variable. Further, it is also evident that 3.24 percent of the observations have high leverage points. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate other allocations. Table 19b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the positive other allocations increase with the net change in other untaxed reserves, the net group contribution, and the possible tax depreciation of machinery and equipment.

(b) Second, the positive other allocations decrease with the possible reversals from periodical reserves and the real interest rate.

(c) Moreover, the positive other allocations are lower for closed firms. On the other hand, the positive other allocations are higher for public firms, firms that are located in large cities, and firms that are located in rural districts.

(d) Finally, the positive other allocations are higher for firms with a market concentration index close to 1. This means that the positive other allocations are lower for firms that operate in more competitive markets. The positive other allocations are also higher for firms that have high market shares.

7.19.3 The Level of the Negative Other Allocations

Influence diagnostics show that 0.87 percent of the observations have severe influence on the response variable. Further, it is also evident that 5.02 percent of the observations have high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate other allocations. Table 19c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the negative other allocations increase with the net group contribution, the possible tax depreciation of machinery and equipment, the possible reversals from periodical reserves, and the real interest rate.

(b) Second, the negative other allocations decrease with the net change in other untaxed reserves.

(c) Moreover, the negative other allocations are lower for closed firms, public firms, and firms that are located in large cities. On the other hand, the negative other allocations are higher for firms that are located in rural districts.

(d) Finally, the negative other allocations are higher for firms with a market concentration index close to 1. This means that the negative other allocations are lower for firms that operate in more competitive markets. The negative other allocations are also lower for firms that have high market shares.

7.20 Estimating Firms' Tax Liability

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate firms' tax liability.

Variable	Method	Mean	Sum	Std. Dev.	MSE
TL		858500.88	68042204487	13990142.60	
pTL	LLN	627238.64	49713052590	14067781.33	16924074.59
	LLG	564695.84	44756098057	14201793.66	16997564.65
	LSN	605990.10	48028957426	10077990.89	11415533.02
	LSG	619671.79	49113327438	20584804.25	21234198.23
	MUNO	539756.81	42779505388	13096542.59	14680119.66
	Tobit 2	678768.43	53797149697	6393687.17	11529829.56

As is evident, the Tobit method gives the best prediction of firms' tax liability. However, the simulation results indicate that this method is not best suited to the simulation model. Instead, the LLN method gives a better prediction. In what follows, we will outline the estimation results for the LLN method.

Firms' tax liability is estimated in two steps. First (in section 7.20.1), we investigate whether firms have a positive or negative tax liability. This is done using a logistic model with which we find the probability that firms have any tax liability. Second (in section 7.20.2), we estimate the level of firms' tax liability.

7.20.1 The Probability that Firms' Tax Liability is Either Positive or Negative

A look at the database indicates that 50072 firms (of a total of 79257 firms) have a positive tax liability. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.20)} \qquad \begin{array}{lll} DTL = 1 & \quad if & \quad TL_i \neq 0 \\ DTL = 0 & \quad if & \quad TL_i = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms have a tax liability and 0 otherwise. Table 20 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 20a. As the value of explanatory variables change, the value of the index DTL varies. The larger the value of DTL the greater the incentive for firms *i* to choose the option $TL_i \neq 0$. Thus, the greater the value of DTL, the greater will be P_i, the probability that firms *i* choose the option $TL_i \neq 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $TL_i \neq 0$ increases with operating income before depreciation (the increase is decreasing), financial income (at a decreasing rate), reversals from periodical reserves, allocations to periodical reserves, and the change in gross national product.

(b) The probability that firms choose the option $TL_i \neq 0$ decreases with financial expenses (at an increasing rate), tax depreciation of machinery and equipment (at an increasing rate), economic depreciation of buildings (at an increasing rate), and the net change in other untaxed reserves (at an increasing rate).

(c) Moreover, the magnitude of the probability that firms choose the option $TL_i \neq 0$ can be different for closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $TL_i \neq 0$ is higher for firms with a market concentration index close to 1. The probability is also lower for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have any tax liability is 36.86 percent, while the frequency of predicting that firms will have a tax liability is 63.14 percent. These probabilities coincide with the observed responses (36.82 percent and 63.18 percent, respectively).

7.20.2 The Level of the Positive and Negative Tax Liability

Influence diagnostics show that 1290 observations have severe leverage points - points that are outliers in the design matrix. Further, 402 observations have severe influence on the response variable. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality are the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of firms' tax liability. Table 20b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Firms' tax liability increases with operating income before depreciation (at a deceasing rate), financial income (at an increasing rate), the net change in other untaxed reserves (at an increasing rate), allocations to periodical reserves, and the change in gross national product.

(b) Firms' tax liability decreases with financial expenses (at an increasing rate), tax depreciation of machinery and equipment (at an increasing rate), economic depreciation of buildings (at an increasing rate), and reversals from periodical reserves.

(c) Moreover, firms' tax liability is lower for closed firms and firms that are located in rural districts. On the other hand, firms' tax liability is higher for public firms and firms that are located in large cities.

(d) Finally, firms' tax liability is higher for firms with a market concentration index close to 1. This means that firms' tax liability is lower for firms that operate in more competitive markets. On the other hand, firms' tax liability is lower for firms that have high market shares.

7.21 Estimating Other Tax Adjustments

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate other tax adjustments.

Variable	Method	Mean	Sum	Std. Dev.	MSE
OTA		-2853018.20	-2.261217E11	121126109	
pOTA	LSN	-2521631.66	-1.99857E11	110345219	57071905.99
	LSG	-2494956.10	-1.977427E11	110178483	56981847.62
	LSN	-2864352.32	-2.2702E11	115386226	68334830.57
	LSG	-3127258.32	-2.478571E11	116799663	69132047.53
	MUNO	-3173443.69	-2.515176E11	119290963	73486585.65
	Tobit 2	3355930.23	265980962621	15047921.04	120418956.28

As is evident, the LSN method gives the best prediction of firms' other tax adjustments. However, our doubts regarding the normality of the variable distribution and simulation results indicate that this method is not best suited to the simulation model. Instead, the MUNO method provides a better prediction than the LSN method. In what follows, we will outline the estimation results for the MUNO method.

Other tax adjustments are estimated in two steps. First (in section 7.21.1), we use a multinomial logit model to investigate whether other tax adjustments are positive, equal to zero, or negative. Second (in section 7.21.2), we estimate the level of the positive other tax adjustments using the Huber-Schweppes robust estimation method. Third (in section 7.21.3), we estimate the level of the negative other tax adjustments using the Huber-Schweppes robust estimation method.

7.21.1 The Probability that Firms Make Other Tax Adjustments

A look at firms' other tax adjustments indicates that 61241 firms have positive other tax adjustments, 11253 firms have negative other tax adjustments, and 6763 firms have not made other tax adjustments. To capture this fact we generate a dependent variable, which is dichotomous in nature, as

$$\begin{aligned}
&DOTA = 1 \qquad && if \quad OTA_i < 0\\
(7.21)\quad &DOTA = 2 \qquad && if \quad OTA_i = 0\\
&DOTA = 3 \qquad && if \quad OTA_i > 0
\end{aligned}$$

Hence, the dependent variable can take three values: 1 if other tax adjustments are negative, 2 if firms do not make any other tax adjustments, and 3 if other tax adjustments are positive. Table 21a concludes the estimated multinomial logit model with a complementary log-log function. As mentioned earlier, the coefficients in this model are difficult to interpret and we will not do that here. Meanwhile, marginal effects can be computed using the statistical program package SAS.

7.21.2 The Level of the Positive Other Tax Adjustments

Influence diagnostics show that 0.67 percent of the observations have severe influence on the response variable. Further, it is also evident that 2.71 percent of the observations have high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality, namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate other tax adjustments. Table 21b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the positive other tax adjustments increase with allocations to periodical reserves, tax depreciation of machinery and equipment (this increase is decreasing), economic depreciation of buildings (this increase is increasing), the net change in other untaxed reserves, the tax liability, financial expenses, and the change in gross national product.

(b) Second, the positive other tax adjustments decrease with reversals from periodical reserves, operating income before depreciation (this decrease is increasing), and financial income.

(c) Moreover, the positive other tax adjustments are lower for closed firms and firms that are located in rural districts. On the other hand, the positive other tax adjustments are higher for public firms and firms that are located in large cities.

(d) Finally, the positive other tax adjustments are higher for firms with a market concentration index close to 1. This means that the positive other tax adjustments are lower for firms that operate in more competitive markets. The positive other tax adjustments are also higher for firms that have high market shares.

7.21.3 The Level of the Negative Other Tax Adjustments

Influence diagnostics show that 1.25 percent of the observations have severe influence on the response variable. Further, it is also evident that 2.37 percent of the observations have high leverage points - points that are outliers in the design matrix. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality are the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate other tax adjustments. Table 21c concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) First, the negative other tax adjustments increase with allocations to periodical reserves, tax depreciation of machinery and equipment, the net change in other untaxed reserves, and financial expenses.

(b) Second, the negative other tax adjustments decrease with reversals from periodical reserves, operating income before depreciation (at a decreasing rate), economic depreciation of buildings (at an increasing rate), the tax liability, financial income, and the change in gross national product.

(c) Moreover, the negative other tax adjustments are higher for closed firms and public firms. On the other hand, the negative other tax adjustments are lower for firms that are located in large cities and firms that are located in rural districts.

(d) Finally, the negative other tax adjustments are lower for firms with a market concentration index close to 1. This means that the negative other tax adjustments are higher for firms that operate in more competitive markets. The negative other tax adjustments are higher for firms that have high market shares.

7.22 Estimating Tax Depreciation of Buildings

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate tax depreciation of buildings.

Variable	Method	Mean	Sum	Std. Dev.	MSE
$TDEP^{BU}$		398612.07	31592796445	8500272.16	
$pTDEP^{BU}$	LLN	163653.78	12970707645	2669744.65	7848299.82
	LLG	194095.90	15383458588	2934613.97	8204514.84
	Tobit 1	269651.15	21371741299	3271974.57	7696154.25

As is evident, the Tobit 1 method makes the best prediction compared with all the other estimation methods. This method provides also the best simulation results. In what follows, we will outline the estimation results for the Tobit 1 method.

7.22.1 The Tobit Model for Positive Tax Depreciation of Buildings

Let us generate a dependent variable, which has the following characteristics

$$
(7.22) \quad \begin{array}{lll} LOWER = TDEP_i^{BU} & if & TDEP_i^{BU} > 0 \\ LOWER = . & if & TDEP_i^{BU} \leq 0 \end{array}
$$

Hence, the dependent variable can take only two values: $TDEP_i^{BU}$ if firms make tax depreciation of buildings and missing values otherwise. Table 22 concludes the estimated Tobit 1 model with a logistic distribution function. Let us now interpret the results in Table 22a. Direct interpretation of various regression coefficients given in Table 22a is not easy. But from the estimated coefficients one can assess the marginal effects. As Table 22a shows:

(a) An increase in economic depreciation of machinery and equipment (the increase is decreasing), net investment in machinery and equipment, the level of buildings (at a decreasing rate), the net change in current assets (at a decreasing rate), the net change in current liabilities, the change in the utilization of tax rules regarding allocations to periodical reserves (at a decreasing rate), and the change in gross national product, will increase firms' tax depreciation of buildings.

(b) An increase in the change in cash flow, the change of firms' dividend policy so that they come closer to the legal constraint on dividend payments, net sales of machinery and equipment, and the real interest rate, will decrease firms' tax depreciation of buildings.

(c) Moreover, tax depreciation of buildings is lower for closed firms as well as firms that are located in large cities. On the other hand, tax depreciation is higher for public firms and firms that are located in rural districts.

(d) Finally, tax depreciation of buildings is higher for firms with a market concentration index close to 1. This means that tax depreciation of buildings is lower for firms that operate in more competitive markets. Tax depreciation of buildings is lower for firms that have high market shares.

7.23 Estimating Allocations to Periodical Reserves

Let us begin by looking at the predictive ability when estimating allocations to periodical reserves.

Variable	Method	Mean	Sum	Std. Dev.	MSE
P^{allo}		745856.28	59114331155	30745988.93	
pP^{allo}	HS	878936.47	69661867807	31187554.54	5693774.24
	LLN	768007.51	60869971084	31194649.08	5588646.73
	LLG	521949.72	41368169045	30735076.57	6197954.65
	Tobit 1	586462.87	46481288072	6224160.53	28306952.93

As is evident, the LLN method gives the best prediction when estimating allocations to periodical reserves. Despite this fact, we have chosen the HS method.

7.23.1 Estimation Results for Allocations to Periodical Reserves

Influence diagnostics show that 2.37 percent of the observations have severe leverage points - points that are outliers in the design matrix. Further, 0.3 percent of the observations have severe influence on the response variable. An ordinary least square estimation and the corresponding tests for residual normal error distributions show that the normality assumption is not fulfilled (the three test statistics for detecting the presence of non-normality, namely, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). Moreover, the residual distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of allocations to periodical reserves. Table 23a concludes the estimated model. Based on the sign on the estimated coefficients, we infer the following:

(a) Allocations to periodical reserves increase with the change in cash flow, the indicator of whether or not firms reach their dividend constraint faster, the reversals from periodical reserves, the maximum allocations to periodical reserves, and the real interest rate.

(b) Allocations to periodical reserves decrease with the change in the maximum allocation to periodical reserves.

(c) Moreover, allocations to periodical reserves are higher for closed firms and firms that are located in rural districts. On the other hand, allocations to periodical reserves are lower for public firms and firms that are located in large cities.

(d) Finally, allocations to periodical reserves are lower for firms with a market concentration index close to 1. This means that allocations to periodical reserves are higher for firms that operate in more competitive markets. Meanwhile, allocations to periodical reserves decrease when the market share of firms increases.

7.24 Estimating Firms' Reduction of Taxes

Let us begin by looking closer at the predictive ability of the different methods that we have used to estimate firms' reduction of taxes.

Variable	Method	Mean	Sum	Std. Dev.	MSE
ROT		37410.36	2965032745	3650762.25	
pROT	LLN	22476.56	1781424323	1683903.66	3993814.31
	LLG	16271.27	1289612195	1187711.57	3815872.86
	Tobit 1	3627.74	287523857	264553.88	3578229.62

The simulation results indicate that LLG is best suited to the simulation model. In what follows, we will outline the estimation results for this method.

Firms' reduction of taxes is estimated in two steps. First (in section 7.24.1), we investigate whether firms make a positive or negative reduction of taxes. This is done using a logistic model with which we find the probability that firms make any reduction of taxes. Second (in section 7.24.2), we estimate the level of firms' reduction of taxes.

7.24.1 The Probability that Firms' Reduction of Taxes is Either Positive or Negative

A look at the database indicates that 1104 firms (of a total of 79257 firms) have made a positive reduction of taxes. To capture this fact we generate a dependent variable, which is dichotomous in nature, taking a 1 or 0 value, as

$$\text{(7.23)}\qquad \begin{array}{llll} DROT = 1 & \quad if & \quad ROT_i \neq 0 \\ DROT = 0 & \quad if & \quad ROT_i = 0 \end{array}$$

Hence, the dependent variable can take only two values: 1 if firms have reduced taxes and 0 otherwise. Table 24 concludes the estimated logistic model with a complementary log-log function. Let us now interpret the results in Table 24a. As the value of explanatory variables change, the value of the index $DROT$ varies. The larger the value of $DROT$ the greater the incentive for firms *i* to choose the option $ROT_i \neq 0$. Thus, the greater the value of $DROT$, the greater will be P_i, the probability that firms i choose the option $ROT_i \neq 0$. Based on the sign of the estimated coefficients, we infer the following:

(a) The probability that firms choose the option $ROT_i \neq 0$ increases with allocations to periodical reserves, reversals from periodical reserves, tax depreciation of machinery and equipment (at a decreasing rate), operating income before depreciation (at a decreasing rate), the net change in other untaxed reserves, the tax liability, financial income, and the change in gross national product.

(b) The probability that firms choose the option $ROT_i \neq 0$ decreases with economic depreciation of buildings (at an increasing rate), other tax adjustments (at a decreasing rate), tax depreciation of machinery and equipment (at a decreasing rate), and financial expenses.

(c) Moreover, the magnitude of the probability that firms choose the option $ROT_i \neq 0$ can be different for closed firms, public firms, and firms that are located in large cities and rural districts.

(d) Finally, the probability that firms choose the option $ROT_i \neq 0$ is lower for firms with a market concentration index close to 1. The probability is also lower for firms with high market shares.

Let us now concentrate on the association of the predicted probabilities and the observed responses. As is evident, the frequency of predicting that firms will not have reduced taxes is 98.61 percent, while the frequency of predicting that firms will have reduced taxes is 1.39 percent. These probabilities coincide with the observed responses (98.61 percent and 1.39 percent, respectively).

7.24.2 The Level of the Positive and Negative Reduction of Taxes

Influence diagnostics show that 49 observations have severe leverage points - points that are outliers in the design matrix. Further, 25 observations have severe influence on the response variable. An ordinary least square estimation and corresponding tests for normal error distributions show that the normality assumption is not fulfilled (the four test statistics for detecting the presence of non-normality, namely, the Shapiro-Wilk test, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply to our data). The distribution shows that the distribution is both askew and tapering (see skewness and kurtosis).

Because of influential observations and non-normality, we use the Huber-Schweppes robust estimation method to estimate the level of firms’ reduction of taxes. Table 24b concludes the estimated model. Based on the sign of the estimated coefficients, we infer the following:

(a) Firms’ reduction of taxes increases with reversals from periodical reserves, economic depreciation of buildings, the net change in other untaxed reserves, the tax liability, financial expenses, and the change in gross national product.

(b) Firms’ reduction of taxes decreases with allocations to periodical reserves, operating income before depreciation, tax depreciation of buildings, and financial income.

(c) Moreover, firms’ reduction of taxes is lower for closed and public firms. On the other hand, firms’ reduction of taxes is higher for firms that are located in large cities or rural districts.

(d) Finally, firms’ reduction of taxes is lower for firms with a market concentration index close to 1. This means that firms’ reduction of taxes is higher for firms that operate in more competitive markets. On the other hand, firms’ reduction of taxes is lower for firms that have high market shares.

8 The Evaluation of the Swedish Tax Model for Incorporated Businesses

In this chapter, we present the results of the simulation model using current tax rules and the economic environment (see section 8.2). The forecasting accuracy is discussed in section 8.3. For evaluation purposes, we also present the simulation results for a proposed corporate tax rate reduction of three per cent (see section 8.4) and for a hypothetical change in the firms' macroeconomic environment (see section 8.5). Finally, in order to summarize the results, we present simulation results for different variables that can be used to assess firms' financial performance (see section 8.1). In seeking such variables, we use ratio analysis. This technique helps us to interpret relationships between the figures of two or more comparable sets of financial statements for different periods of time or different firms. In this section, we explore the role of ratio analysis in the financial analysis.

8.1 Financial Ratio Analysis

Financial ratio analysis is a relatively simple yet powerful tool for assessing the financial condition of firms. It enables us to relate two pieces of financial data to each other. Ratio analysis is useful only as a comparative tool. Two types of comparisons can be made. First, the ratios of a firm can be compared over time to detect significant improvement or deterioration in the firms' financial or competitive position. Second, the ratios can be compared with ratios for other firms.

The various stakeholders in a firm have different concerns with regard to its financial performance. Stockholders and bondholders are primarily interested in the firm's long-term profitability and, accordingly, how well firms are being run. Moreover, they are interested in the firm's level of operating and financial risk. Management is interested in the firm's ability to pay its bills on time (as well as long-run aspects of firms operations), as it is responsible for running the firm's day-to-day operations and for developing a strategy that will enable the firm to earn the required risk-adjusted rate of return on its assets. Five major categories of financial ratios have been developed, each designed to address an important aspect of a firm's financial condition: liquidity ratios, activity ratios, leverage ratios, profitability ratios, and market value ratios. It is important to mention that we are unable to define activity ratio variables with the variables specified in chapters 1 and 2.

8.1.1 The Liquidity Ratio

The liquidity ratio measures the quality and adequacy of current assets to meet current liabilities as they fall due. This ratio refers to a firm's ability to meet its short-term obligations and is closely related to the size and composition of the firm's working capital position. Other things being equal, a higher working capital position implies a more liquid position. This is because the firm's current assets are the easiest to convert into cash, making them the main source of cash to meet maturing obligations. The current ratio equals current assets divided by current liabilities, or

$$(7.1)\quad CR_t = \frac{CA_t}{CL_t}$$

A current ratio of 2 is generally considered to be a safe ratio.[57]

8.1.2 Leverage Ratios

Leverage ratios measure a firm's ability to service its debt. These ratios are designed to measure firms' risk. Highly leveraged firms are more vulnerable to business downturns and therefore are more likely to default on their contractual and non-contractual obligations than those with lower debt-to-equity ratios. In addition, because shareholders receive only the residual cash flow, the riskiness of the returns to equity increases with leverage. There are two approaches to measuring leverage. One is to examine various balance sheet ratios to see the extent to which a firm uses debt to finance its assets. The other approach uses income statement data to develop coverage ratios that measure a firm's ability to service its debt. The balance sheet ratios include the debt/equity ratio, the equity/capital ratio, and the equity/fixed asset ratio, whereas the coverage ratios include times interest earned (interest coverage ratio).

8.1.2.1 Leverage Ratios Based on the Balance Sheet

Debt ratio: The debt ratio measures the extent to which borrowed funds have been used to finance a firm's total assets. As employed here, debt includes all the firm's liabilities: current liabilities, long-term liabilities, and the tax liability of untaxed reserves. The debt ratio is defined as

$$(7.2)\quad DR_t = \frac{CL_t + LL_t + \tau(ASD_t + \hat{PF}_t + OUR_t)}{K_t}$$

This ratio is a measure of the rate of protection provided by the owners for the creditors. The higher the debt ratio is, the more risk that is assumed by the firm's creditors. Shareholders bear more risk as well, because their residual returns are subject to greater variation. In addition, a high debt ratio impairs the firm's ability to borrow in the future and thereby reduces its financial flexibility.

Debt/equity ratio: This ratio provides a close look at the relationship between the capital contributed by creditors and that contributed by owners. It is defined as

$$(7.3)\quad DER_t = \frac{CL_t + LL_t + \tau(ASD_t + \hat{PF}_t + OUR_t)}{SC_t + RR_t + URE_t + (1-\tau)(ASD_t + \hat{PF}_t + OUR_t)}$$

The debt/equity ratio can be derived from the debt ratio as below

$$(7.4)\quad DER_t = \frac{DR_t}{1-DR_t}$$

[57] This is a rule of thumb and it is important to mention that there is substantial variation within and between industries.

Equity/capital ratio: The equity/capital ratio equals the ratio of owners' equity to total capital, or

(7.5) $$ECR_t = \frac{SC_t + RR_t + URE_t + (1-\tau)(ASD_t + \hat{PF}_t + OUR_t)}{K_t}$$

The equity/capital ratio can be derived from the debt ratio and the debt/equity ratio as below

(7.6) $$ECR_t = \frac{DR_t}{DER_t}$$

A ratio equal to or higher than 0.15 is generally considered to be a safe ratio.[58]

8.1.2.2 Leverage Ratios Based on the Income Statement

Times interest earned or interest coverage ratio: This ratio, a commonly used financial ratio, measures a firm's ability to meet its interest payments through its annual operating earnings. It is defined as

(7.7) $$ICR_t = \frac{OIBD_t - EDEP_t^{MA} - EDEP_t^{BU} + FI_t}{FE_t}$$

This ratio indicates the number of times that a firm's operating earnings can cover its interest expenses. The ratio is generally considered to be a safe ratio if it is above unity.

8.1.3 Profitability Ratios

Profitability ratios or operating ratios measure management's effectiveness as indicated by the return on assets and owners' equity. These ratios are designed to evaluate management performance as regards operating the business as well as how productively corporate assets are employed. They are also used to evaluate the economic performance of the business. It is necessary to distinguish between the two, as a business can be doing quite well despite the poor performance of management, and vice versa. There are two types of profitability ratios: profit margins on sales and return on assets employed. Profit margin ratios try to measure a

---

[58] Gilchrist and Himmelberg (1998) introduced a variable measuring the financial Q as $FQ_t = \frac{CA_t - CL_t - LL_t - \tau(ASD_t + \hat{PF}_t + OUR_t)}{K_t}$. This ratio measures the leverage position of the firm, net of current liquid assets. Gilchrist and Himmelberg (1998) show that the leverage ratio specified in this way has a significant effect on a firm's investment. The financial Q can be derived from both the debt ratio and the debt/equity ratio as follows: $FQ_t = \frac{CA_t}{K_t} - DR_t = \frac{CA_t}{K_t} - \frac{DER_t}{1+DER_t}$.

firm’s ability to control expenses relative to sales.[59] The second set of ratios tries to measure a firm’s ability to employ its assets profitably. They include return on total assets and return on equity.

Return on total assets: The return on total assets, also known as the return on investment, is defined as

$$(7.8)\quad ROA_t = \frac{OIBD_t - EDEP_t^{MA} - EDEP_t^{BU} + FI_t + TL_t}{K_t}$$

In order to focus on the earnings of ongoing operations, the income figure in the numerator usually excludes income or losses from transactions outside the ordinary course of business. In our case, we only exclude interest payments. A decline in this ratio reflects a combination of factors. In what follows, we use a technique known as Du Pont analysis to evaluate the effect of these factors simultaneously.

Return on equity: The return on equity expresses the rate of return on shareholders’ equity, where equity is measured at book value (i.e. the actual cash outlay by shareholders). The return on equity can be defined using a Du Pont analysis. The primary objective of Du Pont analysis is to separate the various factors that determine the return on assets. A modified Du Pont formula seeks to determine the factors influencing the return on equity. The modified Du Pont equation is

$$(7.9)\quad ROE_t = ROA_t + (ROA_t - DI_t)DER_t$$

where the weighted average debt interest (DI) is defined as

$$(7.10)\quad DI_t = \frac{FE_t}{CL_t + LL_t + \tau(ASD_t + \hat{PF}_t + OUR_t)}$$

Inserting (7.8) and (7.10) into (7.9), we are able to derive the following expression for the return on equity

$$(7.11)\quad ROE_t = \frac{OIBD_t - EDEP_t^{MA} - EDEP_t^{BU} + FI_t - FE_t + TL_t}{SC_t + RR_t + URE_t + (1-\tau)(ASD_t + \hat{PF}_t + OUR_t)}$$

It is not uncommon for two firms to have identical figures for a given ratio. But these figures may have quite different implications when viewed together with other ratios. For example, two firms may have the same return on equity but very different capital structures. As a result, the more highly leveraged firm is likely to have a lower return on assets and to be riskier. Return on equity can serve as an indicator of management performance. A high return on equity, normally associated with effective management, could indicate an overly leveraged

---

[59] We cannot generate these sets of ratios with the defined variables in chapters 1 and 2.

firm, whereas a low return on equity, usually a sign of ineffective management, may indicate a conservatively financed firm. A ratio equal to or higher than 0.15 is generally considered to be a safe ratio.

8.1.4 Market Value Ratios

A successful management is one that is able to create positive net present value projects. These are projects whose value, as reflected in the market price of the firm's securities, exceeds the cost of the assets needed to undertake them. The market-to-book ratio tries to measure this price/cost relationship, but it does so imperfectly. Inflation and other price changes alter the replacement cost of assets, which makes the measurement of book value per share difficult. Tobin's q deals with the problem of inflation by using the replacement cost of assets when measuring the value created by the firm. It compares the market value of the firm (debt plus equity) with the replacement cost of the firm's assets. The greater is the real return on investment (ROI) relative to the required return on investment ($RROI$), the higher will be the value of q.[60]

Return on investment (ROI): Firms often use some measure of return on investment to guide the resource allocation process. This practice is appropriate to the extent that prospective investments are comparable to existing ones. The relevant investment base in ROI analysis equals the incremental value of all capital employed. The relevant return on an investment is incremental cash flow net of funds needed to maintain the productive capacity of business. Often, firms use return on total assets as an approximation of return on investment. Therefore, ROI should be defined as follows

$$(7.12) \qquad ROI_t = ROA_t = \frac{OIBD_t - EDEP_t^{MA} - EDEP_t^{BU} + FI_t + TL_t}{K_t}$$

Required return on investment ($RROI$): The required return $RROI$ is usually assumed to be the investors' discount rate. The question is whether it is possible to approximate the investors' real discount rate. The required return (or the cost of capital) is the minimum acceptable rate of return required by shareholders of the firm for undertaking an investment project. Unless the investment generates sufficient funds to repay suppliers of capital, the firm's value will suffer. An important question is whether corporate and capital income taxation affects the optimal employment of capital. Sweden is regarded as having a small open economy. This implies that the required return (or the investors' real discount rate) should be equal to the required return in the international capital market, because investors have access to this market (but they cannot affect it). This means that the domestic capital income taxation should not have any impact on investors' required return. However, this is not the case for corporate taxation, which has an impact on the required return. It is assumed that the investors' required return coincides with the pre-tax interest rate on a government loan.[61]

---

[60] The ability to generate a high real ROI depends on the firm's sustainable competitive advantage. We expect a strong correlation between competitive advantage and q.

[61] The interested reader is referred to the survey by Shahnazarian and Stoltz (1999), available on request, about different views on the impact of double taxation of corporate income. This short survey was based on Apel and Södersten (1996), Bhattacharya (1979), Black and Scholes (1974), Boadway and Bruce (1992), Gordon (1959), Graham and Dodd (1951), Miller and Modigliani (1961), Miller and Scholes (1982), Poterba and Summers (1985), Ross (1977), Shahnazarian (1996), Sinn (1990), Stoltz (1997), and Zodrow (1991).

(7.13) $$RROI_t = \frac{i_t}{1-\tau_t^{eff}}$$

where the effective corporate tax rate is defined as

(7.14) $$\tau_t^{eff} = \frac{FTAX_t}{EBA_t}$$

As mentioned above, the greater is the real return on investment (ROI) relative to the required return on investment ($RROI$), the higher will be the value of q. To be able to capture this, we define a variable, which is the difference between the return on investment and the required return on investment as follows

(7.15) $$ER_t = ROI_t - RROI_t$$

where ER stands for excess return on investment.

8.2 The Simulation Results Using Current Tax Rules

As mentioned in Chapter 1, the estimated functions from the statistical module in Chapter 7 together with the difference equation system of Chapter 2 give us the base structure for our simulations. The difference equation system is solved numerically to be able to simulate the future values of firms' decision variables, their income statements and balance sheets, and their tax payments. Table 25 summarizes the simulation results from 2000-2004 that were obtained by solving the difference equation system numerically. Let us penetrate these results. For this purpose we look at firms' financial flows.

Firms' operating income before depreciation increases between 2000 and 2002. However, in 2003, this income decreases temporarily. But in 2004, the increase continues.

Operating income before depreciation is split into two elements: provision for economic depreciation of machinery and equipment and economic depreciation of buildings. The economic depreciation of machinery and equipment shows the same development as operating income before depreciation. The depreciation increases between 2000 and 2004, with a temporary decrease in 2003. On the other hand, the economic depreciation of buildings has a fluctuating development. It decreases in 2000, increases in 2001, decreases then two years in succession, and finally decreases once again in 2004.

The remainder - operating income after depreciation - has the same development as operating income before depreciation. Operating income after depreciation is one part of the earnings before allocations. Before we have a look at the development for earnings before allocations, we must check the additional parts of these earnings: financial income and financial expenses.

Financial income increases in 2000. However, this income decreases in 2001, after which it once again increases constantly for three years. Financial expenses, on the other hand, increase during the whole simulation period. The development of financial income and financial expenses, together with operating income after depreciation, implies that earnings before allocations decrease in 2000, after which there is an increase four years in a row.

Next, we analyze the development of net allocations. This gives us the necessary information to understand the development of earnings before taxes. As is evident, the allocations to accumulated supplementary depreciation show a decreasing development over the entire simulation period. On the other hand, the removals from periodical reserves show an increasing development. On the contrary, both allocations to periodical reserves and other allocations have a fluctuating development. The developments of these allocations/removals, together with the development of earnings before allocations, imply that earnings before taxes first decrease in 2001 and then increase four years in a row.

Firms' tax liability (which is a prediction of their tax payments) decreases in 2002 before increasing once again for two years a row. In 2004, the tax liabilities decrease moderately. Net income, which is the difference between earnings before taxes and the tax liability, decreases in 2001, before increasing two years a row, and then decreases in 2003 before finally increasing again.

In the tax return forms, firms account for adjustments of net income for tax purposes. Different parts of tax adjustments have following developments. Both the tax depreciation of buildings and other tax adjustments have fluctuating developments. On the other hand, losses from previous years (which are deductible from net income) increase during the entire simulation period. These adjustments imply that the calculated tax payments increase in 2000. In 2003-2004, the calculated tax payments are very stable before increasing once again.

However, the amount of taxes that firms finally pay is not the calculated taxes. Firms receive tax reductions. These reductions have a fluctuating development between 2000 and 2004. This implies that the final taxes paid by firms increase in 2000. These tax payments then decrease for two years. From 2003, the corporate tax payments increase once again.

The development of corporate tax payments and earnings before taxes gives rise to the following development of firms' net business income. Net business income decreases the first simulation year, but increases from 2001, and this development continues during the rest of simulation period.

The first financial decision - the proportion of funds to be retained - is then made by dividing net business income into dividends paid to shareholders, the maximum amount available for dividends (the cash flow), allocations to restricted reserves, and retained earnings (the change in other unrestricted equity). The cash flow as well as dividends paid to shareholders have a fluctuating development. Dividends increase in 2000 and 2001. In 2002, the dividend payment decreases temporarily before increasing once again. On the other hand, the allocations to restricted reserves show a stable development. The development of net business income, dividends paid to shareholders, allocations to restricted reserves, and the cash flow, imply that the earnings retained fluctuate too.

Adding in new long-term debt finance, new short-term debt finance, and new equity issues in proportions determined by the firms' decision on gearing, we obtain the total funds available to the firms. Firms decrease their borrowing in 2000 before increasing it in 2002. Firms' decrease their borrowing once again from 2003. On the other hand, new short-term debt shows a fluctuating development. At the same time, new equity issues show a stable development.

These funds go into four uses: net investment in machinery and equipment, net investment in buildings, net investment in other fixed assets, and new current assets. Investment in machinery and equipment increases the first simulation year. In 2001, it decreases slightly, after which it increases moderately from 2002-2004. On the other hand, investment in buildings decreases in 2000 before beginning to increase in 2001. From 2002, it decreases two years in a row before increasing once again. The current ratio increases during the first two simulation years, after which it decreases tree years in a row. Net investment in other fixed assets fluctuate during the entire simulation period.

Let us now look at the financial performance by using ratio analysis. This technique helps us interpret relationships between the figures of two or more comparable sets of financial statements for different periods of time.

The current ratio refers to firms' ability to meet their short-term obligations. A current ratio of 2 is generally considered to be a safe ratio. The weighted average current ratio[62] for firms increases in 2000. During the rest of the simulation period the weighted average current ratio decreases. This indicates that the liquid position for firms will increase over time.

The weighted average debt ratio decreases in 2000, after which it begins to increase. This increase continues during the rest of the simulation period. This indicates that the extent to which firms use borrowed funds to finance their total assets increases.

The weighted average debt/equity ratio increases during the entire simulation period except in 2000. Our results indicate that the capital contributed by creditors increases compared with the capital contributed by owners.

An equity/capital ratio equal to or higher than 0.15 is generally considered to be a safe ratio. Our results indicate that the weighted average equity/capital ratio is not in the safe range during the simulation period.

The weighted average interest coverage ratio is within the safe range (a ratio higher than 1) during the entire period. This means that firms seem to meet their interest payments out of their operating earnings.

The return on equity expresses the rate of return on shareholders' equity. This ratio has a stable development. The weighted average return on equity is high, which is normally associated with effective management of firms.

Return on investment (or return on total assets) focuses on the earnings power of ongoing operations. This return must be compared with the required return on investment to be able to draw conclusions about the value of TOBIN'S q. The weighted average excess return decreases in 2000. However, the weighted average excess return increases in 2001, after which it decreases once again for three years in a row. This indicates that return on investment is constantly reduced relative to the required return on investment. This indicates that the value of TOBIN'S q becomes lower and lower, which indicates that the value of the firms compared to the replacement costs of the firms' assets decreases. Hence, the market's prediction of the value of the returns generated per SEK 1 of additional investment becomes lower.

8.3 Forecasting Accuracy

It has not been an easy task to evaluate the forecasting accuracy of the simulated results. In Table 26, we present a matched pairs test of the hypothesis that the weighted mean of different variables in the sample for the year 2000 coincides with predicted mean for the same variables. The table shows that t-values for $EDEP_t^{MA}$, S_t^{MA}, I_t^{MA}, , I_t^{BU}, $dofa_t$, dca_t, dll_t, dcl_t, dsc_t, $OIBD_t$, FI_t, FE_t, $TDEP_t^{MA}$, zpf_t, GC_t, OA_t, TL_t, OTA_t, $TDEP_t^{BU}$, p_t^{allo}, and

---

[62] Assume that the current ratio for firm 1 and firm 2 is $CR_1 = CA_1 / CL_1$ and $CR_2 = CA_2 / CL_2$. The average current ratio is then calculated as follows. $\overline{CR} = (CA_1 + CA_2)/(CL_1 + CL_2)$. This can be rewritten as: $\overline{CR} = (CA_1 / CL_1)*(CL_1)/(CL_1 + CL_2) + (CA_2 / CL_2)*(CL_2)/(CL_1 + CL_2)$. It is now easy to see that the average current ratio is a weighted average. The weights are simply calculated as firms' current liabilities related to the total liabilities.

ROT_t lie within the accepted range. However, this is not the case for $EDEP_t^{BU}$, drr_t, and $dour_t$. One should therefore be careful with the simulation results concerning these variables.

The most crucial variable in the simulation model and for the Ministry of Finance is the sum of the corporate taxes that firms pay to the government. The weighted sum of taxes paid by all firms in 2000 equals MSEK 48026. Our simulated tax payments for 2000 are MSEK 47735. The difference is MSEK 291, which indicates an underestimation of the tax payments by 0.6 percent. Table 26 reinforces the forecasting accuracy of the firms' tax payments when we use both the information about the mean and the standard deviation. The matched pairs test indicates that we cannot reject the hypothesis that the weighted mean of tax payments in the sample for the year 2000 coincides with the predicted mean for the same variable.

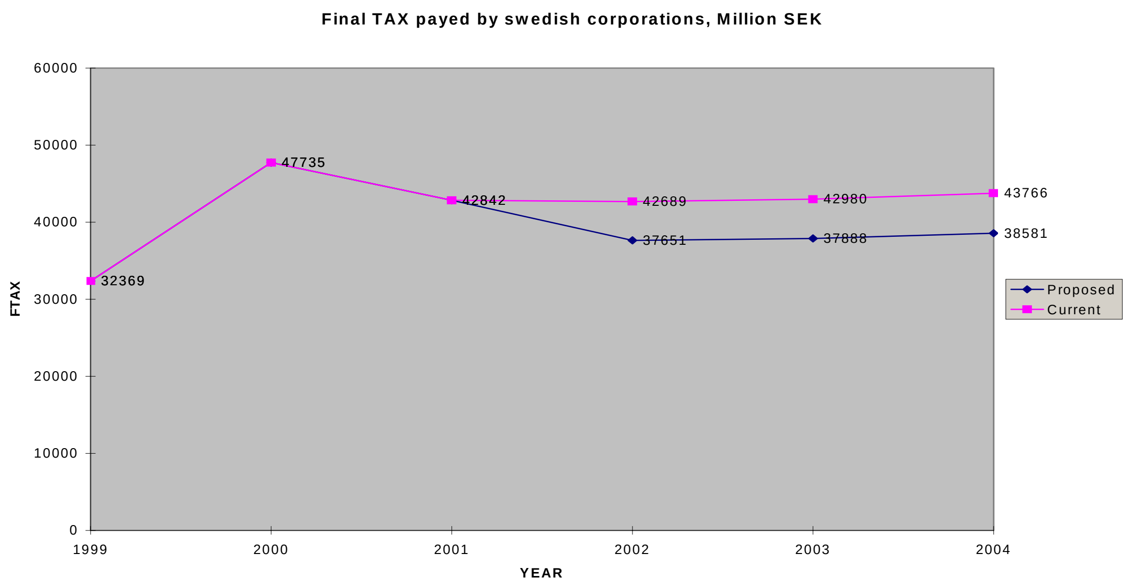
Another way of evaluating the forecasting accuracy is to compare the predicted distribution of tax payments with the actual distribution. This can be done by looking at the mean, the standard deviation, the skewness, the curtosis, and the median of the distributions.

	The actual distribution	*The predicted distribution*
Mean	202425.364	220221.588
Standard Deviation	16057919.6	22966100
Skewness	51.3292645	-155.64327
Curtosis	3517.11279	25481.8173
Median	4076	182961

As is evident, the predicted distribution is more skewed on the left side and more tapering compared to the actual distribution. It is also evident that the median of the predicted distribution is much higher than the median for the actual distribution.

8.4 Simulation Results for a Proposed Corporate Tax Rate Reduction of Three Per Cent

The simulation results for the amount of taxes that firms finally pay and the simulated results for the weighted average financial ratios are summarized in Table 27. The reduction of the corporate tax rate by 3 per cent (from 2002) implies that the final taxes paid by the firms decrease. The cost of the proposed tax rule is about MSEK 5038 in 2002, MSEK 5092 in 2003, and MSEK 5185 in 2004. The simulations indicate that the cost of the proposed tax rule increases over time. The comparison of the final taxes paid by the Swedish firms for the current and proposed rule is summarized in the following figure:



The best way of analyzing the implication of the new rules for the corporations is to look at the development of weighted average financial ratios with the proposed tax rule and compare them with the weighted average financial ratios with the current tax rule. The current ratio is not affected by the new rule. This means that the new tax rule does not have any impact on firms' ability to meet their short-term obligations, which means that the tax change does not improve the liquidity position of firms.

The new tax rule causes a small decrease in the weighted average debt ratio. The tax decrease has a small impact on the extent to which firms use borrowed funds to finance their total assets.

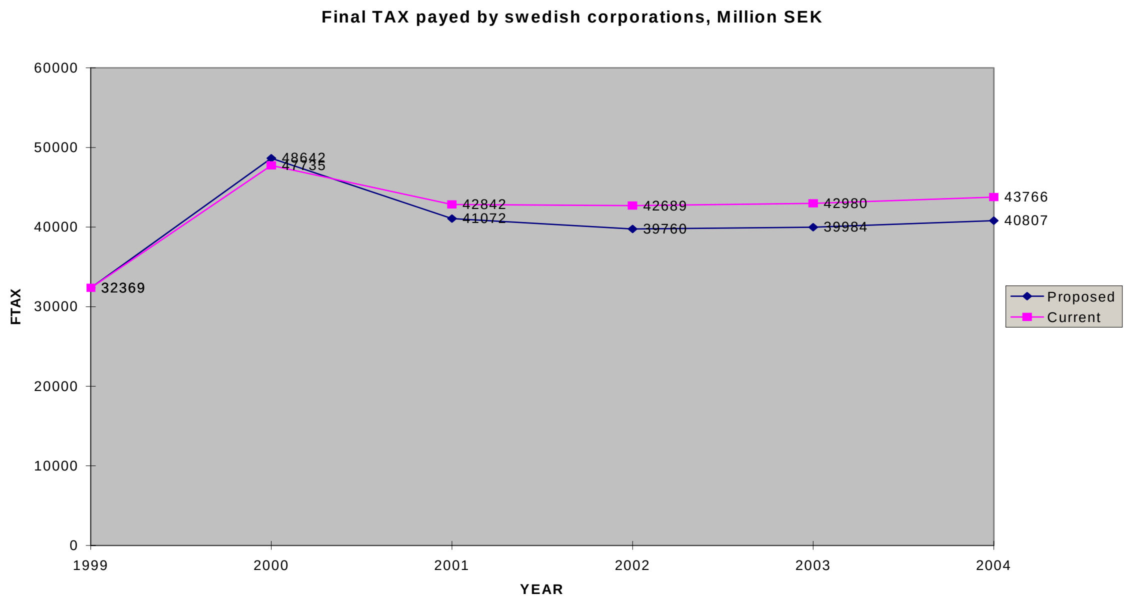
The mean debt/equity ratio decreases as a result of the new tax rule. This result indicates that, because of the new tax rule, the capital contributed by creditors decreases compared with the capital contributed by owners. This is also evident from the increased equity/capital ratio.

The return on shareholders' equity is affected marginally, which indicates that the tax decrease does not have any effect on firms' management.

The required return on investment decreases because of the tax decrease, and hence the weighted average excess return increases. This indicates that the value of TOBIN'S q becomes higher, indicating that the value of the firms compared with the replacement cost of the firms' assets increases. Therefore, the market's prediction of the value of the returns generated per SEK 1 of additional investment becomes higher.

8.5 Simulation Results for a Hypothetical Change in Macroeconomic Developments

The simulation results for a hypothetical change in macroeconomic developments are summarized in Table 28. In this case, we assume that the change in gross national product jumps to a higher level during the entire simulation model. Furthermore, we assume that the interest rate has increased during the same period and that inflation is reduced during the simulation period. These alternative macroeconomic developments reduce the total assets of firms compared with the base assumptions. The same is true for the total liabilities. In this new macroeconomic environment, the government receives higher taxes from the firms during 2000. However, from 2001, firms' tax payments to government decrease, which is apparent from the figure below.



The development of corporate tax payments and earnings before taxes implies that firms' net business income increases in 2000.

The development of net business income, dividends paid to shareholders, the allocations to restricted reserves, and the cash flow, imply that retained earnings decrease in 2000, while they increase in 2001 and 2002 before decreasing once again two years in a row. This is the reason why firms increase their borrowing in 2000, decrease them in 2001, and increase them again in 2004. Short-term borrowing increases in 2000, decreases in 2001 to 2003, and increases once again in 2004. The new equity issues have the opposite development as retained earnings. The simulation results indicate that firms use external financing to compensate for fluctuations in their retained earnings.

The new assumptions about the macroeconomic development do not seem to have any dramatic implications for the firms' investment in machinery and equipment. This is not the case for firms' investment in buildings. This investment increases in 2000, and decreases for the rest of the simulation period. Firms' investment in other fixed assets decreases during the entire period except 2000. Investment in current assets increases in 2000, and then decreases for three years. During the last year of the simulation, investment in these assets increases again. The developments in investment in buildings, other fixed assets and current assets seem to follow the same structure.

Let us now look at the financial performance by using ratio analysis. This technique helps us to interpret relationships between the figures of two or more comparable sets of financial statements for different periods of time.

The current ratio refers to firms' ability to meet their short-term obligations. A ratio of 2 is generally considered to be a safe ratio. The weighted average current ratio for firms increases in 2000. During the rest of the simulation period the weighted average current ratio decreases. This indicates that the liquidity position of firms will improve over time.

The weighted average debt ratio decreases in 2000, after which it begins to increase. This increase continues over the rest of the simulation period. This indicates that the extent to which firms use borrowed funds to finance their total assets increases.

The weighted average debt/equity ratio increases during the entire simulation period except for 2000. Our results indicate that the capital contribution by creditors increases compared with the capital contribution by owners.

An equity/capital ratio equal to or higher than 0.15 is generally considered to be a safe ratio. Our results indicate that the weighted average equity/capital ratio is not in the safe range during the simulation period.

The weighted average interest coverage ratio is within the safe range (a ratio higher than 1) during the entire period. Hence, firms are still able to meet their interest payments out of their operating earnings.

The return on equity expresses the rate of return on shareholders' equity. This ratio has a stable development. The weighted average return on equity is high, which is normally associated with effective management of firms.

The return on investment (or return on total assets) focuses on the earnings power of ongoing operations. This return must be compared to the required return on investment to be able to draw conclusions about the value of TOBIN'S q. However, the weighted average excess return increases in 2001 before decreasing once again for three years in a row. This indicates that the return on investment becomes lower and lower relative to the required return on investment, which indicates that the value of TOBIN'S q becomes lower and lower, which indicates that the value of the firms compared with the replacement cost of the firms' assets decreases. This means that the market's prediction of the value of the returns generated per SEK 1 of additional investment becomes lower.

The weighted average excess return, which is the difference between the return on investment and the required return on investment, decreases in 2000.

9 Future Developments and Improvements

The simulation model presented in this book can both be improved and developed in different directions. In this chapter, we would like to focus on some of these directions.

1) The simulation model: Let us begin with the simulation module. This module was presented in Chapters 1 and 2. We find two ways to improve this module:

a) The current simulation model has four assets (current assets, machinery and equipment, buildings, and other fixed assets) and two major groups of liabilities (long-term liabilities and short-term liabilities). One way to improve the simulation model is to disaggregate both assets and liabilities. Defining *goodwill and other intangible assets* as separate assets, machinery and equipment may be disaggregated. *Fixed assets other than machinery and equipment* could be a separate asset and not a part of firms' buildings. Other fixed assets could be divided into three parts: *Shares and other participations in domestic and foreign companies*, *bonds and other securities*, and *other fixed assets* (which include loans to partners or related people, other long-term receivables, advances to suppliers, deferred charges and other non-depreciable assets). Finally, the best way to define long-term liabilities is to let allocations for pensions be a separate variable.

b) We also believe that the simulation model could be improved by disaggregating different variables in the income statement. Operating income before depreciation is, for example, the sum of firms' *operating revenues*, *labour costs*, *payroll taxes*, and *other operating costs*. Financial income could be divided into *interest income*, *dividends received* (dividends received on stocks and participations, income from affiliated undertakings, income from participating interests, income from other investments and loans forming part of the fixed assets), and *capital gains* (profits on operations disposed of or closed, gains from sales of assets of business lines or facilitates, profits on sales of capital assets, profits on sales of equipment, facilities, etc., and exchange rate profits). Financial expenses can be disaggregated into *interest costs* (on all types of short- and long-term borrowing), *the value adjustments of fixed assets*, and *capital losses* (losses on operations disposed of or closed, losses from sales of assets of business lines or facilitates, losses on sales of capital assets, losses on sales of equipment, facilities, etc., and exchange rate losses). Finally, the simulation model would improve from disaggregation of other tax adjustments as follows: *the taxable revenues which are not booked in the income statement*, *deductible expenses which are not accounted for in the income statement*, *accounted revenues which are not taxable*, *accounted expenses which are not deductible*, *the tax adjustment because of firms' sales of shares*, *the tax adjustment because of firms' sales of buildings*, *the tax adjustments if the firms are partners in a partnership,* and *other tax adjustments* (which in this case include deductions for sales of (disposals of) forest products and deductions for depletion).

Disaggregating variables as outlined above has both advantages and disadvantages. The advantage is that more detailed and qualified assessments become possible. The disadvantage is twofold. First, the dynamic optimisation model must be re-optimised to catch the economic relationships for those new variables, which is very time-consuming work. Second, the propositions above imply that the simulation model will include 43 flow variables instead of 24, which is currently the case. This means that we need to estimate 19 additional flow variables. Our experience from the estimation work is that this work is very time-consuming as well. It should also be noted that there are no guarantees that disaggregating variables will give better forecasts of firms' tax payments.

2) The estimation strategy: The estimations can always be improved by experimenting with other estimation methods that better capture the distributions of the flow variables. There is also a need for more extensive testing of the results obtained by the different estimation methods. Another issue is associated with the bounded-influence estimations. In our estimations, we apply a bounded influence technique suggested by Handschin et al. (1975). There exist several other bounded influence estimators that give higher degrees of protection. The most well-known one is the technique presented by Krasker and Welsch (1982). Whether this technique is better than the one we have chosen is debatable.

The problem with the suggestions above is that they are time-consuming and difficult to make operational. It should be borne in mind that the simulation model must be updated each year as new databases become available. This means that all flow variables must be re-estimated every year. The estimation programs available today include necessary tests and estimation methods. This means that the model can be updated quite easily. New estimation methods and more extensive tests will increase the time needed to update the simulation model.

3) The recursive system: We estimate our model in a recursive way. Recursive models are never under-identified. Crucial for this model, however, is the fact that it is assumed that disturbance terms for the endogenous variables are uncorrelated, that is, the assumption of zero contemporaneous correlation must be fulfilled. When disturbances are correlated across equations, it may be advisable to consider the Seemingly Unrelated Regression (SURE) approach to cope with this assumption. In this case, even if the assumption of no correlation among error terms is not satisfied, we still acquire consistent estimates, and there is thus no need for recursivity assumptions. This is a suggestion that may improve the model estimation.

Klevmarken (2001) argues that, in practice, even the maximum likelihood technique will meet several barriers and become unfeasible. In order to overcome certain obstacles, the Simulated Method of Moments (SMM) is suggested and a discussion of an alternative approach, namely the “moment calibrated” estimator, is in place. However, these approaches are computationally demanding and have not yet been used in practice.

4) The simulation results: Many modellers advocate the use of empirical residuals. The idea behind empirical residuals is the fact that there is no estimation technique that can fully replicate the underlying distribution of a variable. Residuals are therefore inevitable. One way to replicate the distribution (even though the estimation method does not) is a random sample from the empirical residuals, which is added to the prediction made by the estimation method. We have developed programs for such techniques. However, we do not use them in the current simulation model. The reason is the fact that our estimations are based on the samples for 1997-1999, while the base for our simulations is the 1999 sample. This means that the forecasting accuracy of our simulations cannot be improved by using empirical residuals from the estimations, which are based on the samples for 1997-1999. The sum of these empirical residuals does not equal zero for 1999. Our experiments confirm this. It could be a good idea to think through this problem again. However, we believe that the use of empirical residuals for improving the forecasting accuracy may not be the best way to improve the forecasts. In contrast, we believe that it is much better to use other estimation methods to improve the replication of the underlying distribution.

Conclusions

The main conclusion from the work described in this book is that estimation of behavioral factors has an important role in assessing the financial implications of proposals for change in the tax code. The dynamic microeconometric simulation model for incorporated businesses provides a model that both analyzes the behavioral effects induced by changes in the tax code and forecasts the tax revenues. This makes more qualified assessments possible, which in this case facilitates better revenue estimates with regard to corporation tax.

The basic idea behind the simulation model is to combine the dynamic behavior of the corporate system with a statistical model that captures the development and the interrelationships between the firms' different decision variables. *The dynamic behavior of the corporate system* is captured by several difference equations that identify how different variables in the firms' balance sheets change over time. To be able to do this we use the information in the firms' three basic financial statements: the balance sheet, the income statement, and the statement of changes in financial conditions. Furthermore, the difference equations system also incorporates special features of corporate taxation. *The firms' decisions regarding the flow variables* are modeled in a statistical module. From the dynamic optimization problem we derive the relationships between these flows variables and other economic variables. These relationships are then estimated using different robust estimation methods. The estimated functions from the statistical module are inserted into the difference equations system. Finally, this system is solved numerically to be able to simulate the future values of the stock variables in the firms' balance sheets. Let us now look at the main findings in this book:

The dynamic optimisation model includes a technique that we use to derive economic relationships between different decision variables. This is a technique that we like to accentuate. The idea behind this technique is to explore how a change in any parameter or exogenous variable will affect the equilibrium position of the model, which in the present context refers to the optimal values of the decision variables. Changing any parameter or exogenous variable upsets the initial equilibrium. As a result, the various endogenous variables (decision variables) must undergo certain adjustments. The rate of change of the equilibrium values of the endogenous variable are found by differentiating the first order conditions in equilibrium. The result is a set of equations involving the differentials. These differentials include changes in different endogenous variables (which are known) and changes in shadow prices (which are not known). The changes in the shadow prices are obtained by differentiating the complementary-slackness conditions. These differentiations make it possible for us to find an expression for the changes in shadow prices. These expressions are known quantities. The advantage of this approach is that we can derive economic relationships for different variables without relying on specifications of production functions, adjustments cost functions, and bankruptcy costs.

From *the estimation results*, we have obtained several insights. Firstly, we found that the predictions made by different estimation methods do not give us information about the prediction accuracy within the simulation model. The reason for this is the fact that estimations are made on pooled data between 1997 and 1999. However, the starting point for the simulation model is 1999, which of course has fewer observations. Thus, a method that happens to give the best prediction may fail to provide a good forecast in the simulation model. Moreover, in the simulation model, the forecast made by a decision variable is dependent on forecasts made by other variables. This means that we may choose an estimation method that does not give the best sample prediction.

Secondly, the data on different variables indicate that we have problems with unusually large influence on the response variable. Influence diagnostics show further that we also have problems with leverage points - points that are the outliers in the design matrix. The presence of such observations implies that the normality assumption is not fulfilled. The distributions of the variables are both askew and tapering. Because of this we use robust estimation methods to estimate the flow variables. These estimation methods have been shown to replicate the underlying distribution of the variables in the best way. This knowledge is one of the main results in this book, which we like to emphasize.

The simulation results give many interesting insights. Firstly, the simulation results indicate that the forecasting accuracy of the model regarding firms' final tax payments is satisfactory. Secondly, the simulation results for a proposed corporate tax rate reduction of three per cent give us a very important insight. A corporate tax decrease does not seem to have a major impact on firms' investment and financial behavior. This is, however, true given the delimitation of our model. The simulation model does not include the shareholders of the firms. Had this been the case, we would also be able to simulate the market value of the firms.[63] The value of the firms could then be used as an additional explanatory variable in the estimations of flow variables. There are no guarantees that this approach will imply that a corporate tax rate reduction (or reductions of personal tax rates, dividend tax rates and tax rates on realized capital income) will have a greater impact on firms' financial and investment behavior.[64] Third, we use financial ratio analysis to analyze the impact of a corporate tax reduction. This gives us a summary of the financial conditions. As we mentioned, this information is important for shareholders' overall valuation of firms. Finally, it has become clear that macroeconomic developments have a major impact on firms' financial and investment behavior. This may not be surprising. However, the interesting part is that we use three different macroeconomic variables as explanatory variables (the change in gross national product, the interest rate, and inflation) in our estimations. The simulation model gives us the opportunity to examine the impact of combined changes in the macroeconomic variables.

We do not assert that the simulation model presented in this book is a complete tool for all kinds of analysis. In Chapter 9, we point out different parts of the simulation model that can be developed and improved. These should be considered as directions for future research to improve the functioning of the current simulation model.

[63] The shareholders use the information about the market value of the firm, together with the financial ratio analysis, to decide whether they would provide new financing to firms.

[64] We need to complete our database with data about the shareholders of these firms to be able to include the value of the firms in the simulation model.

Appendix A: List of Variables

Balance Sheet Variables

K	**–**	***Assets***
CA	–	Current Assets
FA	–	Fixed Assets
MA	–	Machinery and Equipment
BU	–	Buildings
OFA	–	Other Fixed Assets
CMA	–	The Taxable Residual Value of Machinery and Equipment
WC	–	Working Capital
B	**–**	***Liabilities***
CL	–	Current Liabilities
LL	–	Long-Term Liabilities
UR	–	Untaxed Reserves
ASD	–	Accumulated Supplementary Depreciation
OUR	–	Other Untaxed Reserves
PF_t^{t-5}	–	Remaining Periodical Reserves From t-5 in period t
PF_t^{t-4}	–	Remaining Periodical Reserves From t-5 in period t
PF_t^{t-3}	–	Remaining Periodical Reserves From t-5 in period t
PF_t^{t-2}	–	Remaining Periodical Reserves From t-5 in period t
PF_t^{t-1}	–	Remaining Periodical Reserves From t-5 in period t
PF_t	–	Periodical Reserves in Current Period t
EC	–	Equity Capital
SC	–	Share Capital
RR	–	Restricted Reserves
URE	–	Unrestricted Equity

Income statement variables

OIBD	–	Operating Income Before Depreciation (Accounting Income)
$EDEP^{MA}$	–	Economic Depreciation of Machinery and Equipment
$EDEP^{BU}$	–	Economic Depreciation of Buildings
OIAD	–	Operating Income after Economic Depreciation
FI	–	Financial Income
FE	–	Financial Expenses
EBA	–	Earnings Before Allocations
$TDEP^{MA}$	–	Tax Depreciation of Machinery and Equipment
OA	–	Other Allocations
zpf	–	Change in Periodical Reserves
p^{allo}	–	Allocations to Periodical Reserves
EBT	–	Earnings Before Taxes
TL	–	Tax Liability
NI	–	Net Income
TA	–	Tax Adjustments

OTA – Other Tax Adjustments
$TDEP^{BU}$ – Tax Depreciation of Buildings
OL_t^{t-1} – Losses From Previous Years
TAX – Calculated Tax Payments
ROT – Reduction Of Taxes
$FTAX$ – Final Tax Payments
OL_t – The Stock of Old Losses
NBI – Net Business Income

Flow variables

I^{MA} – Net Investment in Machinery and Equipment
I^{BU} – Net Investment in Buildings
dca – Net Change in Current Assets
$dofa$ – Net Change in Other Fixed Assets
dcl – Net Change in Current Liabilities
dll – Net Change in Long-Term Liabilities
$dour$ – Net Change in Other Untaxed Reserves
dsc – Net Change in Share Capital
drr – Net Changes in Restricted Reserves
$dURE$ – Net Change in Unrestricted Equity (Retained Earnings)
$cashfl$ – Cash flow
S^{MA} – Sales of Machinery and Equipment
IG – Investment Grant
DIV – Dividends Paid to Shareholders
GC – Net Group Contribution

Financial Ratios

CR – The Current Ratio
DR – The Debt Ratio
DER – The Debt/Equity Ratio
ECR – The Equity Capital Ratio
FQ – The Financial Q
ICR – The Interest Coverage Ratio
ROA – Return on Total Assets
ROE – Return on Equity
DI – Average Debt Interest
ROI – Return on Investment
$RROI$ – Required Return on Investment
τ^{eff} – Effective Corporate Tax Rate
ER – Excess Return on Investment

Legal Constraints

$TDDB^{MA}$ – Tax Depreciation of Machinery and Equipment (Declining Balance Method)
$TDSL^{MA}$ – Tax Depreciation of Machinery and Equipment (Straight-Line Method)
$TDRV^{MA}$ – Tax Depreciation of Machinery and Equipment (Rest Value Method)

MTDM – Maximum Amount of Tax Depreciation that Firms May Deduct from Their Taxable Income
dmtdm – Difference Between MTDM and $TDEP^{MA}$
ddmtdm – Indicator Whether Firms Change Their Utilization of Depreciation Allowances
MPA – The Maximum Amount of Allocations that Firms are Allowed to Make to Periodical Reserves
dmpa – Difference Between the Maximum Amount of Allocations to Periodical Reserves and the Allocations Made by the Firms
ddmpa – The Change in the Utilization of Tax Rules Regarding Allocations to Periodical Reserves
dcashfl – The Change in Cash Flow
mcash – Maximum Dividends Firms Can Pay to Their Shareholders
dmcash – The Difference Between the Maximum Dividends Firms Can Pay to Their Shareholders and the Amount of Dividends They Actually Pay
ddmcash – The Change of Firms' Dividend Policy so that they Come Closer to the Legal Constraint on Dividend Payments

Parameters

δ^{DB} – The Depreciation Rate (Declining Balance Method)
δ^{S} – The Depreciation Rate (Straight-Line Method)
δ^{RV} – The Depreciation Rate (Rest Value Method)
M – Number of Months in the Firms' Income Year
τ – Corporate Tax Rate

Macroeconomic Variables

BNP – Gross National Product
ranta10 – The Interest Rate on a Government Bond with a Maturity of 10 Years
inf – Inflation
realr – Real Interest Rate

Other Variables

Public – Indicator for Firms that are Public Companies
FAAB – Indicator for Firms that are Closed Companies
Largcity – Indicator for Firms Located in Large Cities
Ruralare – Indicator for Firms Located in Rural Areas
Market – Index for Competition in the Market
MarketW – The Market Share of the Firm

Variables Controlling for the Decisions Made by the Firms

DI^{MA} – Possible Investment in Machinery and Equipment
DI^{BU} – Possible Investment in Buildings
Ddofa – Possible Net Change in Other Fixed Assets
Ddll – Possible Net Change in Long-Term Liabilities
Ddsc – Possible Net Change in Share Capital
Dzpf – Possible Change in Periodical Reserves

$Ddour$ – Possible Net Change in Other Untaxed Reserves
$DTDEP^{MA}$ – Possible Tax Depreciation of Machinery and Equipment

Appendix B: The Framework of the Simulation Model within a Simple Model with Two Assets

In this appendix, we will present the framework of the simulation model within a simple model with two assets. The idea behind this appendix is to give the reader the opportunity to understand the building blocks of the simulation model.

The framework is presented in three steps. In section B1, we present a simple model with two assets. In this section, we also outline the way the simulation model works. In section B2, we present the dynamic optimization model and the approach we use to find the economic relationships for firms' decision variables. In this section, we also discuss why we use a recursion system to estimate the behavior of the firms. Finally, in section B3, we insert the estimated equations for different decision variables (from section B2) into the simulation model described in section B1. These summarize the system that is used for simulation purposes.

B1 The simulation model

The asset side of the balance sheet is divided into two major components: current assets and fixed assets (machinery and equipment in this simple model). The liability side of the balance sheet includes only equity capital, which in this simple model is equal to the firms' unrestricted equity. The closing balance sheet for firms at the end of year *t* is summarized below.

The closing balance sheet for an incorporated firm in period *t*

Assets		**Liabilities**	
Fixed assets	MA_t	*Equity capital*	URE_t
Current assets	CA_t		

In the balance sheet the value of the firm's assets must be equal to the firm's liabilities, so that

(B.1) $$CA_t + MA_t = URE_t$$

The income statement in this simple model includes only the operating income before depreciation ($OIBD_t$). This implies that firm's earnings before taxes coincides with the operating income before depreciation as follows

(B.2) $$EBT_t = OIBD_t$$

Firms pay corporate taxes based on their earnings before taxes

(B.3) $$TAX_t = \tau EBT_t$$

where τ is the corporate tax rate. We can thus write net business income as

(B.4) $NBI_t = EBT_t - TAX_t = (1-\tau)OIBD_t$

Net business income, NBI_t, increases unrestricted equity. However, unrestricted equity decreases also because of the maximum amount available for dividends in the current period (the so-called net cash flow, $cashfl_t$). Thus, unrestricted equity in period t can be derived from

(B.5) $URE_t = URE_{t-1} + NBI_t - cashfl_t$

Let us now summarize the dynamic characteristic of the balance sheet. The stock (state) variables of this model are URE_t, CA_t, and MA_t. In what follows we will specify equations of motions that hold for these state variables.

The level of current assets at the end of time *t* equals the level of current assets at the end of time *t-1 plus* the net change in current assets

(B.6) $CA_t = CA_{t-1} + dca_t$

The level of the firm's machinery and equipment at the end of time *t* equals the level of machinery and equipment at the end of time *t-1 plus* the net investment in new machinery and equipment

(B.7) $MA_t = MA_{t-1} + I_t^{MA}$

Inserting the difference equations for URE_t, CA_t, and MA_t (from (B.5)-(B.7)) into equation (B.1) we obtain the cash flow constraint

(B.8) $cashfl_t = (1-\tau)OIBD_t - I_t^{MA} - dca_t$

Equations (B.5)-(B.8) are the major equations in our simulation model. We use equations (B.5)-(B.7) to simulate the values of different balance sheet items in the next period, while equation (B.8) makes sure that we achieve balance between the asset and the liability sides of the balance sheet during the simulation.

As is evident, we need initial values for I_t^{MA}, dca_t, and $OIBD_t$ to be able to solve the difference equations system in (B.5)-(B.8). These initial values are obtained by estimating these variables. In the next section, we present the economic model that we use to find economic relationships for firms' behavior regarding I_t^{MA}, dca_t, and $OIBD_t$.

B2 The Dynamic Optimization Model

Let us begin by defining the value of the firms at time t

(B.9) $\mathrm{V}(t) = \int_{u=t}^{\infty} cashfl\; e^{-i(u-t)} du$

where i is the shareholders' return on holding bonds and $cashfl = (1-\tau)OIBD - I^{MA} - dca$. We define operating income before depreciation as

(B.10) $OIBD = f(MA, CA)$

where $f(MA, CA)$ is the production function[1] with $f_{MA} > 0$, $f_{CA} > 0$, $f_{MAMA} < 0$, and $f_{CACA} < 0$. As we mentioned in the previous section the changes in current assets and machinery and equipment are

(B.11) $\dot{MA} = I^{MA}$

(B.12) $\dot{CA} = dca$

We also impose non-negativity constraints on each and every balance sheet item

(B.13) $MA \geq 0$

(B.14) $CA \geq 0$

We assume that the interest rate is exogenously given. Firms will choose the time path of investment in machinery and equipment (I^{MA}) and net change of current assets (dca) so that the market value of their shares is maximized. The state variables of the optimization problem are the stock of machinery and equipment (MA), and the stock of current assets (CA). The equations of motion for these state variables were specified in equations (B.11) and (B.12).

Firms' objective is then to maximize (B.9) subject to (B.11)-(B.14), and the initial conditions $MA = MA_0$, and $CA = CA_0$. The current-value Hamiltonian for this problem is:

$$F = (1-\tau) f(MA, CA) - I^{MA} - dca + \mu_{MA} I^{MA} + \mu_{CA}\, dca + n_{MA} MAK + n_{CA}\, CA$$

where μ_{MA} and μ_{CA} are the shadow prices or co-state variables of the stock of machinery and equipment and the stock of current assets. n_{MA} and n_{CA} are the Khun-Tucker shadow-price of constraints (B.13) and (B.14). The first order necessary conditions are:

(B.15) $I^{MA}: \; -1 + \mu_{MA} = 0$

(B.16) $dca: \; -1 + \mu_{CA} = 0$

(B.17) $MA: \; \dot{\mu}_{MA} = i\mu_{MA} - [(1-\tau) f_{MA} + n_{MA}]$

[1] The firms' production function is equal to the firms' revenue, because the product price is assumed to be equal to one.

(B.18) $CA: \ \dot{\mu}_{CA} = i\mu_{CA} - [(1-\tau) f_{CA} + n_{CA}]$

(B.19) $n_{MA}: \ \frac{\delta F}{\delta n_{MA}} \geq 0 \, , \ n_{MA} \geq 0 \, , \ n_{MA} \frac{\delta F}{\delta n_{MA}} = 0$

(B.20) $n_{CA}: \ \frac{\delta F}{\delta n_{CA}} \geq 0 \, , \ n_{CA} \geq 0 \, , \ n_{CA} \frac{\delta F}{\delta n_{CA}} = 0$

We impose stationary constraints on (B.17)-(B.18), i.e.

(B.21) $\dot{\mu}_{MA} = \dot{\mu}_{CA} = 0$

Equation (B.21) together with equations (B.17) and (B.18) give the values of μ_{MA} and μ_{CA}

(B.22) $\mu_{MA} = \frac{1}{i}[(1-\tau) f_{MA} + n_{MA}]$

(B.23) $\mu_{CA} = \frac{1}{i}[(1-\tau) f_{CA} + n_{CA}]$

Using the steady state solutions for μ_{MA} and μ_{CA} in equations (B.22) and (B.23), the first order conditions in steady state become

(B.24) $I^{MA}: \ -i + (1-\tau) f_{MA} + n_{MA} = 0$

(B.25) $dca: \ -i + (1-\tau) f_{CA} + n_{CA} = 0$

(B.26) $n_{MA}: \ n_{MA} MA = 0$

(B.27) $n_{CA}: \ n_{CA} CA = 0$

In the next step, we will subject the optimization to investigations of the comparative-static sort. We are concerned with the magnitude of the change in decision variables equilibrium values resulting from a given change in parameters or exogenous variables. The rate of change of the equilibrium values of the endogenous variable are found by total differentiating the first order conditions in equilibrium (B.24)-(B.27).

(B.28) $(1-\tau) f_{MACA} dCA + (1-\tau) f_{MAMA} dMA + dn_{MA} = di + f_{MA} d\tau$

(B.29) $(1-\tau) f_{CACA} dCA + (1-\tau) f_{CAMA} dMA + dn_{CA} = di + f_{CA} d\tau$

(B.30) $dn_{MA} MA + n_{MA} dMA = 0$

(B.31) $dn_{CA} CA + n_{CA} dCA = 0$

Solving (B.30) for dn_{MA} and (B.31) for dn_{CA} and substituting the obtained solution into equations (B.28) and (B.29) yields

(B.32) $\beta_1^{MA} dCA + \beta_2^{MA} dMA = \beta_0^{MA}$

(B.33) $\beta_1^{CA} dCA + \beta_2^{CA} dMA = \beta_0^{CA}$

where $\beta_0^{MA} = di + f_{MA} d\tau$[2], $\beta_0^{CA} = di + f_{CA} d\tau$, $\beta_1^{MA} = (1-\tau)\, f_{MACA}$, $\beta_2^{CA} = (1-\tau)\, f_{CAMA}$, $\beta_2^{MA} = [(1-\tau)\, f_{MA} - (n_{MA} / MA)]$, and $\beta_1^{CA} = [(1-\tau)\, f_{CA} - (n_{CA} / CA)]$ are evaluated in the initial values and are therefore constants. Let us now solve equation (B.32) for $dMA = I^{MA}$ and equation (B.33) for dCA. This yields

(B.34) $I^{MA} = (\beta_0^{MA} / \beta_2^{MA}) - (\beta_1^{MA} / \beta_2^{MA})\, dca = \gamma_0^{MA} + \gamma_1^{MA}\, dca$

(B.35) $dca = (\beta_0^{CA} / \beta_1^{CA}) + (\beta_2^{CA} / \beta_1^{CA})\, I^{MA} = \gamma_0^{CA} + \gamma_1^{CA}\, I^{MA}$

As is evident from the equation system of (B.34) and (B.35), we have an interrelationship between I^{MA} and dca. This means that the equation system must be solved or estimated simultaneously. In this simple model, this may not seem to be a complicated task. However, in the simulation model presented in the main text, which includes 24 decision variables, this is rather complicated. The complication has to do with the nature of the data in hand. There are many different statistical difficulties that must be solved. The most important one is the presence of extreme observation on both dependent and explanatory variables. Solving this statistical problem within the context of simultaneous estimation is very difficult.[3] Therefore, we have chosen a recursive system to estimate each and every decision variable within a firm. Let us now introduce the time index to make the recursive system more straightforward.

(B.36) $I_t^{MA} = \gamma_0^{MA} + \gamma_1^{MA}\, dca_{t-1}$

(B.37) $dca_t = \gamma_0^{CA} + \gamma_1^{CA}\, I_t^{MA}$

In this recursion system, we assume that firms first make decisions about their fixed assets before making decisions about net investment in current assets. After the investment decisions, firms undertake different financial decisions. For these decisions, it is important to estimate the operating income before depreciation. We

---

[2] The corporate tax rate has not been changed during 1997-1999, which is the estimation period in our case. The interest rate did change during the same period. In our estimations of some of the decision variables, we use the real interest rate as an explanatory variable.

[3] Peracchi (1991) introduced bounded-influence estimators for the SURE model. There are two practical problems with this approach. Firstly, these bounded-influence estimators are difficult to implement. Most importantly, we are not sure whether these estimators will improve our simulation results or not. Secondly, as we mentioned in Chapter 1, there is no functioning microeconometric simulation model for firms. Therefore, we believe that it is better to use simpler estimators (bounded-influence estimators for each and every model) in a first step to be able to deliver a simulation model that works satisfactory. Using bounded-influence estimators for the SURE model is a natural forthcoming development of the simulation model presented in this book and is therefore left as a suggestion for future research. See further Chapter 9 where we outline the ideas for further research.

know that $OIBD_t = f(MA_t, CA_t)$. By using (B.6) and (B.7), we find the following economic relationship for $OIBD_t$[4]

(B.38) $$OIBD_t = \gamma_0^{OIBD} + \gamma_1^{OIBD} I_t^{MA} + \gamma_2^{OIBD} dca_t + \gamma_3^{OIBD} MA_{t-1} + \gamma_4^{OIBD} CA_{t-1}$$

In statistical models we must realize that economic relationships are not exact and contain both a predictable systematic component and an unobserved and unpredictable random error component. Therefore, we add a disturbance term to each and every equation above. Giving the errors a random interpretation converts our economic models into statistical probability models and gives us a basis for statistical inference.

We estimate our model in a recursive way. As can be seen from the set of the equations, they will all be estimated one-way, with no feedback looping. A recursive model is a situation where OLS can be applied appropriately even in the context of simultaneous equation systems – OLS can be applied to each equation separately. Moreover, recursive models are never under-identified. Crucial for this model, however, is the fact that it is assumed that disturbance terms for the endogenous variables are uncorrelated. By applying OLS in a recursive model we obtain consistent estimates.

We use different robust estimation methods to estimate equations (B.36) to (B.38) because some of the underlying assumptions for the OLS estimations are not fulfilled. In our estimations, we use the information from 1997-1999. For these years, we have two time series observations on different variables for each and every firm. The observations in the samples for 1997-1999 are pooled for estimation purposes.

B3 The dynamic system used for simulation purposes

Let us now insert the estimated equations for different decision variables (from section B2) into the simulation model described in section B1. This yields the following dynamic system for firms

(B.39) $$I_t^{MA} = \hat{\gamma}_0^{MA} + \hat{\gamma}_1^{MA} dca_{t-1}$$

(B.40) $$MA_t = MA_{t-1} + I_t^{MA}$$

(B.41) $$dca_t = \hat{\gamma}_0^{CA} + \hat{\gamma}_1^{CA} I_t^{MA}$$

(B.42) $$CA_t = CA_{t-1} + dca_t$$

(B.43) $$OIBD_t = \hat{\gamma}_0^{OIBD} + \hat{\gamma}_1^{OIBD} I_t^{MA} + \hat{\gamma}_2^{OIBD} dca_t + \hat{\gamma}_3^{OIBD} MA_{t-1} + \hat{\gamma}_4^{OIBD} CA_{t-1}$$

(B.44) $$cashfl_t = (1-\tau) OIBD_t - I_t^{MA} - dca_t$$

(B.45) $$URE_t = URE_{t-1} + NBI_t - cashfl_t$$

[4] In this simple model, we assume the function for $OIBD_t$ is linear. However, in our estimations, we investigate the functional forms that are best suited to the regression models.

where $\hat{\gamma}_0^{MA}$, $\hat{\gamma}_1^{MA}$, $\hat{\gamma}_0^{CA}$, $\hat{\gamma}_1^{CA}$, $\hat{\gamma}_0^{OIBD}$, $\hat{\gamma}_1^{OIBD}$, $\hat{\gamma}_2^{OIBD}$, $\hat{\gamma}_3^{OIBD}$, and $\hat{\gamma}_4^{OIBD}$ are the estimated coefficients. We begin our simulation from year $t-1=1999$. We use the information for the net investment in current assets in the current period ($t-1=1999$) to draw conclusions about the investment in machinery and equipment in the next period $t=2000$ (equation (B.39)). The investment in machinery and equipment increases the stock of machinery and equipment in the balance sheet for the next period (equation (B.40)). As is evident from equation (B.41), the net investment in current assets in the next period is a function of firms' investment decisions regarding machinery and equipment. This net investment in current assets increases the stock of current assets in the balance sheet for the next period (equation (B.42). Firms' investment must be financed in one way or another. In our case, the financing comes from equity capital, which is a function of operating income before depreciation (or earnings before taxes in this simple model). Operating income before depreciation in the next period is a function of firms' investment in different assets in the next period and the stock of the same assets in the current period (equation (B.43)). Having determined the investment in different assets and earnings before taxes, we can use equation (B.44) to calculate firms' cash flow. Finally, net business income in the next period increases the stock of unrestricted equity. On the other hand, the cash flow in the next period decreases the stock of unrestricted equity in the next period (equation (B.45)).

Appendix C: General Structure and Assumptions Underlying the Classical Regression Models

C.1 General Structure

In this section we discuss different strategies and methods that we use in order to estimate certain relationships and concentrate on how to deal with different assumptions that underlie these relationships. We start the analysis by considering a *k*-variable linear regression model to estimate the dependent variable Y as follows

(C.1) $Y_i = \beta_1 + \beta_2 X_{i2} + ... + \beta_k X_{ik} + u_i$

The regression model in equation (C.1) gives the mean or expected value of Y conditional upon the fixed values of $X_2,...,X_k$. This model is linear in the parameters $\beta_1,...,\beta_k$ and can be estimated by OLS regression. The variable Y is related to (or depends on) a number of explanatory variables $X_2,...,X_k$. The slope coefficients $\beta_1,...,\beta_k$ measure the average percentage change in Y for a one-percent increase in $X_2,...,X_k$. In matrix notation (C.1) can be rewritten as

(C.2) $Y = X\beta + u$

where $\beta = (\beta_1, \beta_2,...,\beta_k)'$, $X = (X_1',...,X_{n_{pi}}')'$, and $X_i' = (1\ X_{i2},...,X_{ik})$. Vector X_i' includes a subset of the explanatory variables (the variables included in the firms' three basic financial statements as well as macroeconomic variables. Our next task is to estimate the unknown parameters of the model. Following least squares estimation, we find estimates for the parameters as

(C.3) $\hat{\beta} = (X'X)^{-1}X'Y$

The least square estimator is unbiased $E[\hat{\beta}] = \beta$ and has a covariance matrix given by $\text{cov}(\hat{\beta}) = \sigma^2(X'X)^{-1}$ where σ^2 is the variance of u_i and $(X'X)^{-1}$ is the inverse matrix appearing in equation (C.3)). An unbiased estimator of σ^2 is given by $\hat{\sigma}^2 = \dfrac{\hat{u}'\hat{u}}{n_{pi} - k}$ with $\hat{u}'\hat{u} = Y'Y - \hat{\beta}X'Y$. After the estimation of the parameters, we undertake hypothesis tests to be able to investigate whether the explanatory factors that we have included have a significant effect, individually or jointly, on Y.

Tests of significance of the individual parameters are based on the assumption that u_i follows a normal distribution with zero mean and constant variance σ^2. Given the normality assumption, it can be shown that $\hat{\beta} \sim N[\beta, \sigma^2(X'X)^{-1}]$, that is, each element of $\hat{\beta}$ is normally distributed with mean to the corresponding element of true β and the variance given by σ^2 times the appropriate diagonal element of the inverse matrix $(X'X)^{-1}$. Since in practice σ^2 is unknown, it is estimated by $\hat{\sigma}^2$. By shifting to the *t* distribution, it follows that each element of $\hat{\beta}$ follows the *t* distribution with $n_{pi} - k$ df, $t = \dfrac{\hat{\beta}_k - \beta_k}{var(\hat{\beta}_k)}$. The *t*-distribution can therefore be used to test the hypothesis about the true β_i. Because of the absence of a known σ^2, we replace $var(\beta_k)$ with $v\hat{a}r(\hat{\beta}_k)$ and use the *t*-distribution instead of the normal distribution. Thus, $t = \dfrac{\hat{\beta}_k - \beta_k}{v\hat{a}r(\hat{\beta}_k)} \sim t_{(n_{pi}-k)}$. The null hypothesis of the non-significant coefficient is rejected if this computed value is greater than t_c (the critical value from the right side of the distribution), or less than $-t_c$ (the critical value from the left side of the distribution).

We can go further than the result on the *t*-test and show that the *F*-statistic for jointly testing any number of linear hypothesis can be written in a general form by appropriately defining a $(J \times k)$ matrix R and a $(J \times 1)$ vector r, where *J* is the number of linear hypotheses. We simultaneously test the J linear combination or hypothesis about the elements in β, where the null and alternative hypothesis can be written as $H_0 \colon R\beta = r$ and $H_1 \colon R\beta \neq r$. Each row in the $(J \times k)$ matrix R defines a linear combination of the elements in β, and r is a $(J \times 1)$ vector of the values hypothesized for the J linear combinations. If the null hypothesis H_0 is true, it can be shown that $F = \dfrac{(R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\beta - r)}{J\hat{\sigma}^2} \sim F_{[J,(n_{pi}-k)]}$. We reject H_0 if $F > F_c$ where, for a given significance level, F_c is the critical value from $F_{[J,(n_{pi}-k)]}$ distribution.

Polynomial Regression

To complete the statistical models in Chapter 3, we need to examine the functional form or algebraic relationship among our economic variables. For the statistical models considered in this book, we assume that the relationships are nonlinear. The relevant form may vary from problem to problem. A number of practical restrictions allow us to transform variables of the model at best as polynomials of the original variables. This poses us the possible difficulty of functional form choice, which is the practical matter of the degree of the polynomials. By creating a set of dummy variables for each explanatory variable, based on retaining the same proportion of data on each interval, we estimate the model as a first step of the analysis. Observing how the estimated coefficients for dummies are related to each other for each explanatory variable, we obtain a crude approximation of the functional form for each independent variable. This, together with other tests of least squares assumptions, lays the ground for how different explanatory variables will enter the model.

C.2 Assumptions of the Multiple Regression Models

Several assumptions should be met in order to apply a valid regression model. Most regression assumptions are concerned with residuals. Here we concentrate on several of these assumptions and consider alternative ways to deal with abnormal observations in our data.

Collinearity

Multicollinearity is the inter-correlation of the independent variables in a multiple regression context. The absence of multicollinearity is essential to the model. Collinearity, which is associated with inaccurate estimates of the regression coefficients, will cause the variances to be large and will produce inaccurate hypothesis tests. These inflated variances are quite damaging to regression because some variables add very little or even no new and independent information to the model.

Collinearity is problematic when the purpose of the analysis is explanation rather than mere prediction. Collinearity makes it more difficult to achieve significance of the collinear parameters. However, if such estimates are statistically significant, they are as reliable as any other variables in a model, and even if they are not significant, the sum of the coefficients is likely to be reliable. In this case, increasing the sample size is a viable remedy for collinearity when prediction instead of explanation is the goal and, in fact, the sample size is relatively large in our study to cover such an aspect.

There exist several other ways for detecting multicollinearity such as a matrix of bivariate correlation, the regression of each independent variable in the equation on all other independent variables, and so on. A better approach, though, is using the Variance Inflation Factor (VIF). Variation inflation is the consequence of multicollinearity. We may say that multicollinearity is the symptom, while variance inflation is the disease. In a regression model we expect a high variation explained (R-square); the higher the variation explained is, the better the model is expected to be. However, if multicollinearity exists, it is probable that the variance, standard error, and parameter estimates are all inflated. In other words, the high R-square is not a result of good independent predictors, but a miss-specified model that carries mutually dependent and thus redundant variables. Variance inflation factor (VIF) is a common way for detecting multicollinearity and we will apply it in our analysis. On this basis we can decide whether to throw out certain variables or not by examining the size of VIF. The general rule is that the VIF should not exceed 10 for the variable to be considered as independent.

Different data transformation in the model may cause additional multicollinearity, besides the one that we could correct through the usage of the VIF. Practical examples are the inclusion of interaction terms in the model or curvilinear cases that involve quadratic and/or cubic terms (polynomial regression). Besides, transformation to higher powers may be a source of artificially creating outliers in the direction of explanatory variables, which may tend to yield unreliable results.

Using centered scores, also known as deviation scores, is one way to avoid collinearity in regression.[5] Techniques such as transforming to orthogonal

[5] When scores cluster about the mean, the deviations from the mean will be small and thus the standard deviation will be small. Deviation scores are obtained by subtracting the mean from the raw scores.

polynomials – partial or full orthogonalization – can also be considered as a solution in order to make variables independent of each other.

Normality assumption

Normal error distributions are not necessary for OLS, but the validity of significance tests and the efficiency of the regression may be reduced in the case of non-normal distributed error terms. Although there is not general agreement regarding the best way to test normality[6], there are four test statistics for detecting the presence of non-normality, namely, the Shapiro-Wilk, the Kolmogorov-Smirnov test, the Cramer von Mises test, and the Anderson-Darling test that we will apply. The latter three tests attempt to calculate goodness-of-fit based on the empirical distribution function (EDF) for a specified distribution.

Most normality tests have small statistical power (probability of detecting non-normal data) unless the sample size is large, that is, the tests' ability to reject the null hypothesis increases with the sample size. As the sample size becomes larger, increasingly smaller departures from normality can be detected. On the other hand, it is argued that if the sample size is over 2000, the latest three mentioned tests should be used, while if the sample size is less than 2000, the Shapiro test performs better.

Although many authors recommended using skewness and kurtosis for examining normality, it is argued that skewness and kurtosis often fail to detect distributional irregularities in the residuals and the test based on these measures may be less useful. Still, we will look at these two parameters as guidance and a tool for comparison between different distributions.

Heteroscedasticity

One of the key assumptions of regression analysis is that the variance of the errors is constant across observations, in which case the errors are called homoscedastic. Standard estimation methods are inefficient when the errors are heteroscedastic or have non-constant variance, leading, thus, to unreliable confidence intervals and inaccurate tests.

The residuals of estimation are used to investigate the heteroscedasticity of the true disturbances and the Breusch-Pagan test is applied. This test assumes that the error variance varies with a set of regressors.

If heteroscedasticity is present in our data we will consider a number of transformations. If more than one variable is the source of heteroscedasticity or we do not know which of the four independent variables causes it, we can use the predicted value of the dependent variable (*Y*-hat) to transform and assume either a linear relationship between the variance of the error term and *Y*-hat, or a quadratic relationship[7]. A second transformation starts with an OLS of the squared residuals on the rest of the regressors and the variables of the model are then divided by the

---

Deviation scores have a zero mean and the same standard deviation as the raw scores. Standard deviation scores (*z*-scores) are obtained by dividing deviation scores by the standard deviation. Standard Scores have a zero mean and a standard deviation equal to one[5]. While using deviation scores centers the data and removes the effect of the mean, converting the deviation scores to *z*-scores removes the effects of the original units of measurement and the degree of dispersion of the variable.

[6] It is important to note that the normality test will be applied to the residuals rather than the raw scores.

[7] This line of thinking may suffer from the fact that other possible functional forms describing the relationship between error variance and *Y*-hat may exist.

residual of this regression. In a third transformation we attempt to make the residuals homoscedastic without making any explicit assumptions on the error variance. By somehow grouping observation – for example by size – and trying to correct for heteroscedasticity through dividing variables in the model with the standard error of each group.

C.3 The Logistic Model

To capture the fact that only a fraction of firms may undertake an action regarding a decision variable, we generate a dependent variable, which is dichotomous in nature and takes the value of 1 or 0, as

$$\text{(C.4)} \quad \begin{aligned} DY_i &= 1 \, if \ y_i > 0 \\ DY_i &= 0 \, if \ y_i = 0 \end{aligned}$$

Hence, the dependent variable can take only two values: 1 if the firm takes an action in the current period and 0 if the firm does not take an action in the current period. We could express DY_i as a linear function of the explanatory variables. This is called a linear probability model. However, this kind of model is plagued by several problems, such as non-normality of unobservable random variable, heteroscedasticty of unobservable random variables, the possibility of DY_i lying outside the [0-1] range, and the generally lower R^2 values. Moreover, a linear probability model is not logically a very attractive model because it assumes that $P_i = E(DY_i = 1)$ increases linearly with the dependent variables. Therefore, what we need is a probability model that fulfills the following conditions. As the explanatory variables increase, $P_i = E(DY_i = 1)$ increases but never steps outside the [0-1] interval, and the relationship between P_i and the explanatory variables is nonlinear, that is one that approaches zero at slower and slower rates as the explanatory variables get small and approach 1 at slower and slower rates as the explanatory variables become very large. The cumulative distribution function chosen in this paper, to capture the functional form described above, is the logistic. This gives rise to the following logit model

$$\text{(C.5)} \quad P_i = E(DY_i = 1 | X_{i2}, \cdots, X_{ik}) = \beta_1 + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} \text{ where } i = 1, \ldots, n$$

where β_1 is the intercept, β_2 to β_k are partial slope coefficients, i is the i:th observation, and $n = n_{97} + n_{98} + n_{98}$ is the size of the sample. The regression model in (C.5) gives the mean or expected value of DY_i conditional upon the fixed (in repeated sampling) values of $X_{i2}, \cdots, X_{ik}$. Equation (C.5) can be rewritten in an alternative but more illuminating way as follows

$$\text{(C.6)} \quad P_i = X_i^{'} \beta$$

where $\beta = (\beta_1, \beta_2, \cdots, \beta_k)^{'}$ and $X_i^{'} = (1, X_{i2}, \cdots, X_{ik})$. $X_i^{'}$ include a subset of the explanatory variables.

Consider now the following representation of the firms' action regarding a decision variable

(C.7) $P_i = F(X_i^{'}\beta) = \dfrac{1}{1+e^{-X_i^{'}\beta}}$

where $F(\cdot)$ is the cumulative distribution of a logistic random variable. If P_i is the probability of undertaking an action, then $1-P_i$, the probability of not undertaking an action, is $1-P_i = \dfrac{1}{1+e^{X_i^{'}\beta}}$. $\dfrac{P_i}{1-P_i}$ is the odds ratio in favor of undertaking an action. Now if we take the natural log of the odds ratio, we obtain $\eta_i = \ln\left(\dfrac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}$. The log of the odds ratio is not only linear in X, but also linear in the parameters. This is the logit model. $\beta_2, \cdots, \beta_k$ are the slopes, and measure the change in L for a unit change in X, that is, it tells us how the log-odds in favor of undertaking an action change as an explanatory variable changes by one unit. The intercept is the value of log-odds in favor of undertaking an action if all explanatory variables are zero. Like most interpretations of intercepts, this interpretation may not have any physical meaning. Equation (C.7)) represents what is known as the cumulative logistic distribution function. We use the maximum likelihood method to estimate the parameters.[8] The likelihood function of the logit model is

(C.8) $l(\beta) = \prod_{DY_i=0}[1-F(X_i^{'}\beta)]\prod_{DY_i=1}[F(X_i^{'}\beta)] = \prod_{i=1}^{n}[F(X_i^{'}\beta)]^{DY_i}[1-F(X_i^{'}\beta)]^{1-DY_i}$

Taking logs, we obtain

(C.9) $\ln l(\beta) = \sum_i DY_i \ln F(X_i^{'}\beta) + (1-DY_i)\ln[1-F(X_i^{'}\beta)]$

The first order condition for maximization requires

(C.10) $\dfrac{\delta \ln l(\beta)}{\delta\beta} = \sum_i \left[\dfrac{DY_i f_i}{F_i} + (1-DY_i)\dfrac{-f_i}{(1-F_i)}\right] x_i = 0$

8 For a discussion of maximum likelihood in the context of the logit model, see Amemiya (1985), Fomby, Hill and Johnson (1984), and Green (1993).

The choice of a particular form for F_i leads to the empirical model. We use the computer program SAS to obtain the maximum likelihood estimates of the logit model. By default, SAS uses the Fischer-scoring method, which is equivalent to fitting by iteratively re-weighted least squares.[9] Using the maximum likelihood estimates of the logit model we obtain the following linear predictor

(C.11) $\hat{\eta}_i = \beta_1 + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_k X_{ik}$

The maximum likelihood estimator of the logit model parameters has an approximate normal distribution, in large samples, with mean β and a covariance matrix given by equation $\operatorname{cov}(\hat{\beta}) = (X'DX)^{-1}$ where X is the $n \times k$ matrix of observations on k explanatory variables for n firms, and $D = diag(d_1, d_2, \cdots, d_n)$ is a diagonal matrix with elements $d_i = F(X_i^{'}\beta)[1 - F(X_i^{'}\beta)]$, where $F(\cdot)$ is the logistic cumulative distribution function in equation (C.7). After the estimation of the logit model parameters, we perform hypothesis tests to be able to investigate whether the explanatory factors that we have included have a significant effect, individually or jointly, on the probability of choice.

Tests of the significance of individual parameters are based on the fact that in large samples the maximum likelihood estimator $\hat{\beta}$ from the logit model has the distribution: $\hat{\beta} \stackrel{a}{\sim} N[\beta, \operatorname{cov}(\hat{\beta})]$. Consequently $t = \dfrac{\hat{\beta}_k - \beta_k}{se(\hat{\beta}_k)} \stackrel{a}{\sim} N(0,1)$. Given the null and alternative hypotheses, $H_0: \beta_k = 0$ and $H_1: \beta_k \neq 0$, the *t*-ratio is $t = \dfrac{\hat{\beta}_k}{se(\hat{\beta}_k)}$. If the null hypothesis is true, the *t*-ratio has a normal distribution in a large sample, and the critical values for the test may be taken from the standard normal distribution or the $t_{(n-k)}$-distribution if the sample is not large.

Within the framework of maximum likelihood estimation, general and joint hypotheses about parameter values may be tested by the likelihood ratio test. Let the null and alternative hypotheses be stated as: $H_0: R\beta = r$ and $H_1: R\beta \neq r$. These linear equations represent J independent hypotheses about the parameters β. In large samples the test procedure is the likelihood ratio test which compares the value of the log-likelihood function, $\ln l$, evaluated at the maximum likelihood estimator $(\hat{\beta}*)$ that results when the log-likelihood function is maximized subject to the restrictions $R\beta = r$ being true. The log-likelihood ratio test statistic $\lambda_{LR} = 2[\ln l(\hat{\beta}) - \ln(\hat{\beta}*)]$ has a $\chi^2_{(J)}$ distribution if the joint null hypothesis is true. If the data do not support the null hypotheses then the value of the test statistic becomes large, and the null hypothesis is rejected if $\lambda_{LR} = \chi^2_{(J)}$.

The estimated parameters of the logit model can be used to predict the behavior of firms who must choose whether to undertake an action or not. This is done for the actual values of the explanatory variables, $X_i^{*'}$, (see equation (C.11)). The predicted

[9] For a description, see the SAS/STAT user guide, version 8, volume 2, pp. 1942-1943.

value for $P(Y \le 1|x)$ is obtained by back-transforming the corresponding measures for the linear predictor as follows

$$\text{(C.12)} \quad \hat{P}_i = \frac{1}{1+e^{-\hat{\eta}_i}}$$

C.4 The Multinomial Model

To capture the fact that a fraction of firms may undertake a positive action regarding a decision variable while other firms may undertake a negative action or not undertake an action at all, we generate a dependent variable, which is dichotomous in nature and take the value of 1, 2, or 3 as

$$\text{(C.13)} \quad \begin{aligned} DY_i &= 1 if \ y_i < 0 \\ DY_i &= 2 \ \ if \ y_i = 0 \\ DY_i &= 3 \ \ if \ y_i > 0 \end{aligned}$$

Hence, the dependent variable can take only three values: 1 if the firm takes a negative action in the current period, 2 if the firm does not take an action in the current period, and 3 if the firm takes a positive action. If there are n observations, then the probability distribution of the number falling into 3 categories (m_1, m_2, m_3) can be modeled by the multinomial distribution

$$\text{(C.14)} \quad f(m_1, m_2, m_3) = \frac{m!}{m_1!m_2!m_3!} P_1^{m_1} P_2^{m_2} P_3^{m_3}$$

where $\sum_{j=1}^{3} m_j = m$.

Let (p_{i0}, p_{i1}, p_{i2}) be the category probabilities. The cumulative category probabilities can be derived by $P_{ic} = \sum_{j=1}^{c} p_{ij}$, $c = 1{,}2$ (note that $P_{i3} = 1$). P_{ic} is the cumulative probability of the c:th or lower category. The ordinal model is

$$\text{(C.15)} \quad g(P_{ir}) = \mu_c + X_i{}'\beta \qquad for \qquad c = 1{,}2$$

where μ_1 and μ_2 are intercept terms that depend only on the categories and X_i is a vector of covariates that does not include an intercept term. The logit and complementary log-log functions g can be used here.[10]

The logit function: The model implies that we can compute c log-odds ratios

[10] For the estimation of multinomial models, the reader is referred to Green (2000).

(C.16) $$\eta_{ic} = \ln\left[\frac{P_{ic}}{1-P_{ic}}\right] = \mu_c + X_i'\beta$$

where $c = 1.2$. The predicted value for $\hat{P}_{i1} = \hat{p}_{i1}$ is obtained by back-transforming the corresponding measures for the linear predictor as follows

(C.17) $$\hat{p}_{i1} = \frac{e^{\hat{\eta}_{i1}}}{1+e^{\hat{\eta}_{i1}}}$$

which is the probability that the firm takes a negative action. The predicted value for $\hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2}$ is obtained by back-transforming the corresponding measures for the linear predictor as follows

(C.18) $$\hat{P}_{i2} = \hat{p}_{i1} + \hat{p}_{i2} = \frac{e^{\hat{\eta}_{i2}}}{1+e^{\hat{\eta}_{i2}}}$$

Using (C.17), we can solve an expression for $\hat{p}_{i2}$ from equation (C.18) as follows

(C.19) $$\hat{p}_{i2} = \frac{e^{\hat{\eta}_{i2}} - e^{\hat{\eta}_{i1}}}{(1+e^{\hat{\eta}_{i1}})(1+e^{\hat{\eta}_{i2}})}$$

which is the probability that the firm does not take any action. We know that $\hat{p}_{i3} = 1 - \hat{p}_{i1} - \hat{p}_{i2}$ so that

(C.20) $$\hat{p}_{i3} = \frac{e^{\hat{\eta}_{i2}}}{1+e^{\hat{\eta}_{i2}}}$$

which is the probability that the firm takes a positive action.

Figure 1: The Structure of the Modules

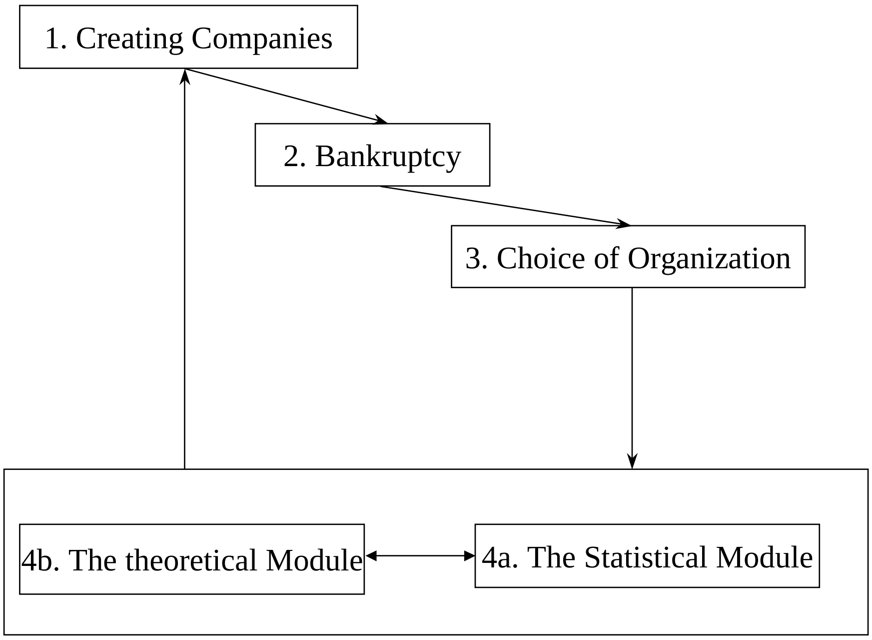


Figure 2: The Structure of the Theoretical Module

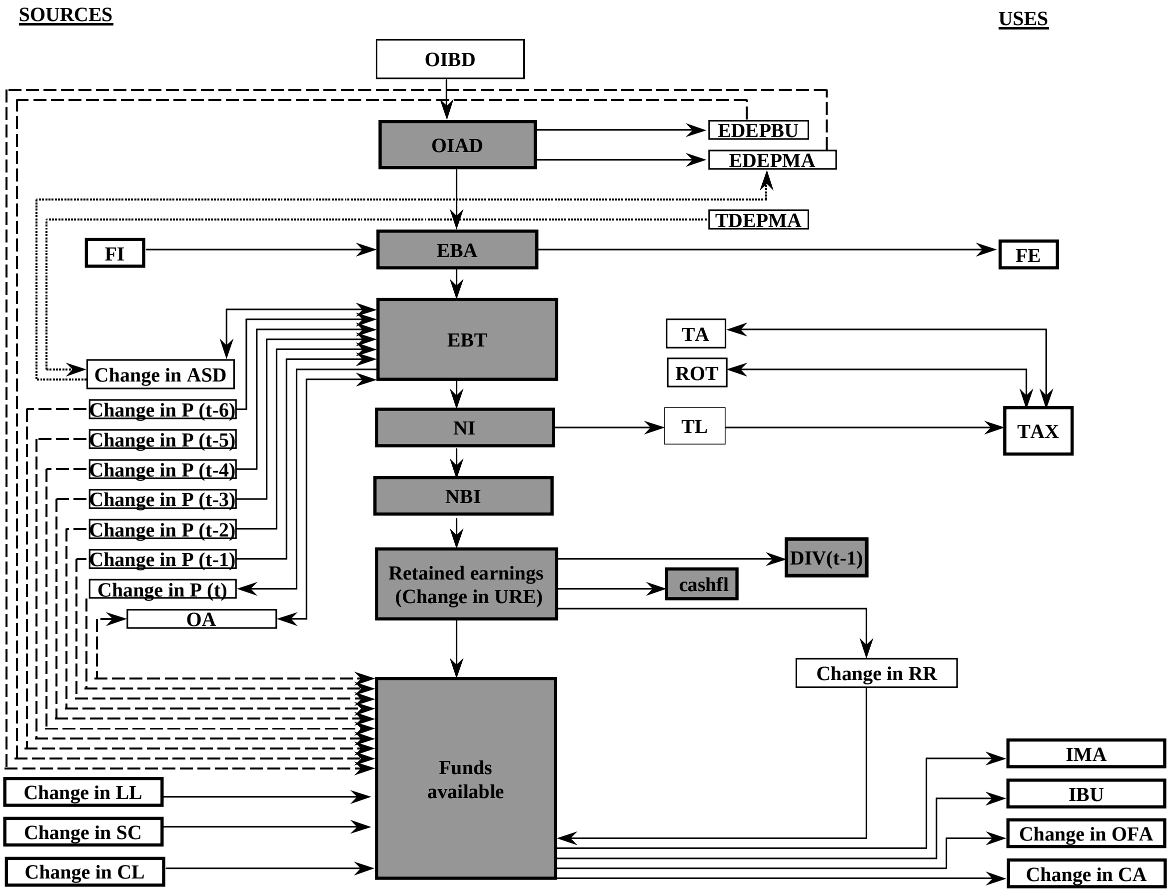


Figure 3: The Structure of the Theoretical Module (cont.)

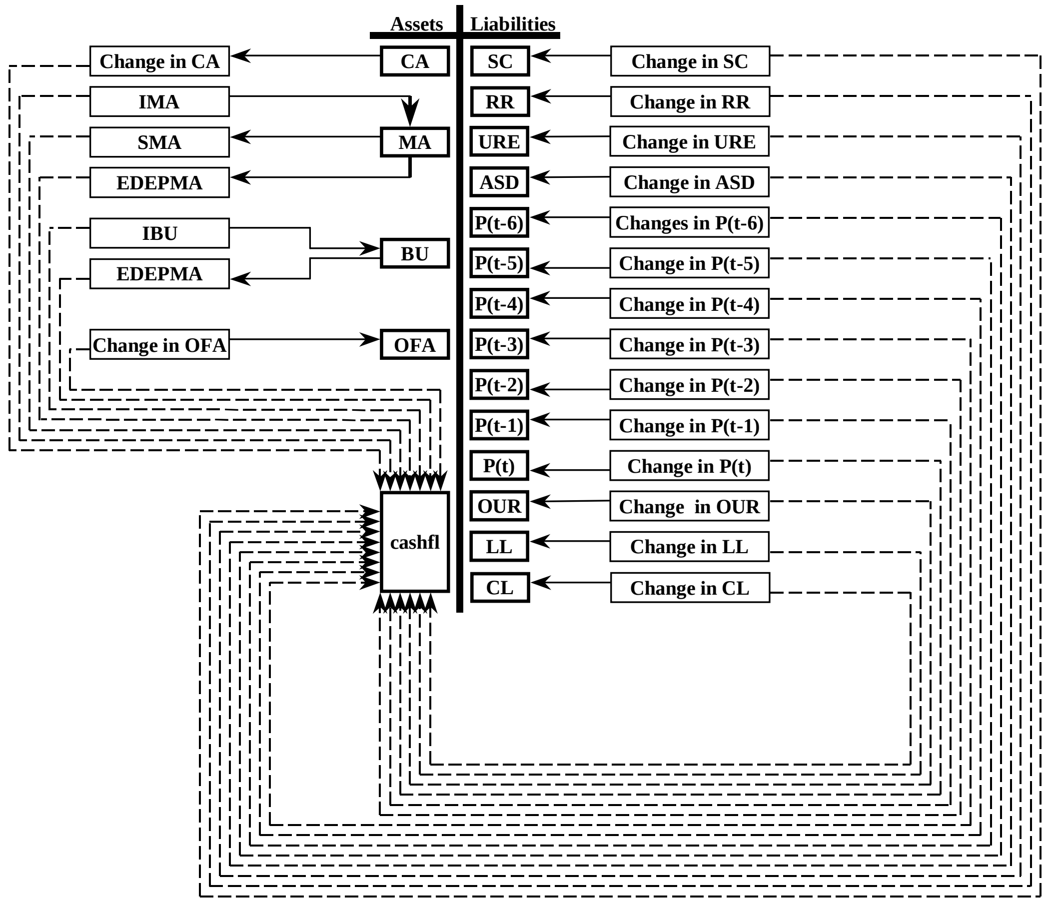


Figure 4: The Structure of the Statistical Module

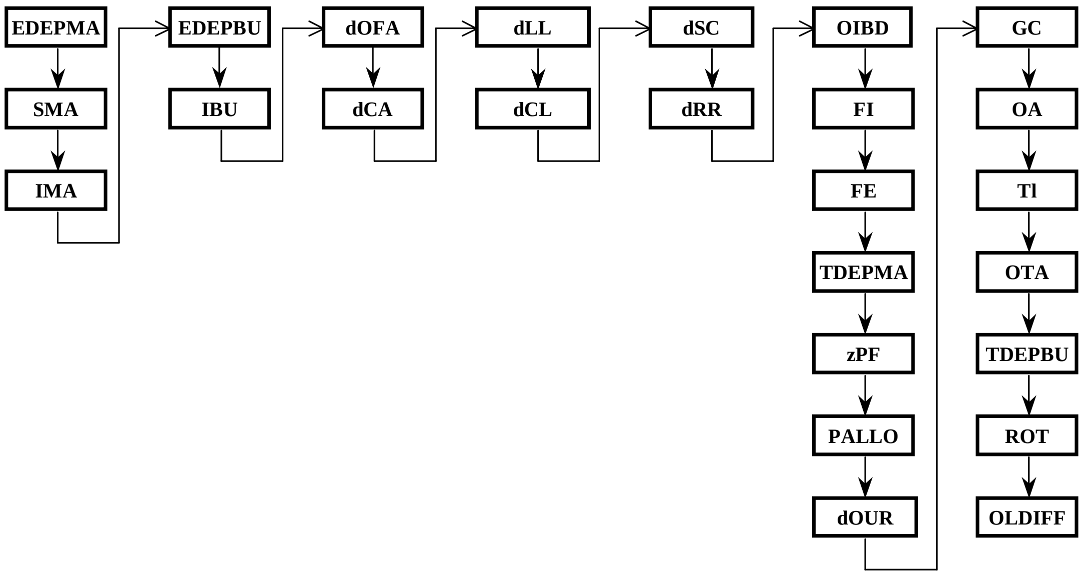


Table 1: Estimation results for $EDEP^{MA}$

Table 1a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	63400
2	0	15857

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	79338.314	77491.353
SC	79347.594	77695.523
-2LogL	79336.314	77447.353

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	1888.9612	21	<.0001
Score	1685.9496	21	<.0001
Wald	1647.2494	21	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	0.3749	0.0331	127.9790	<.0001
sumcasht_1	-48E-12	3.17E-11	2.3005	0.1293
diffcasht_1	-268E-13	7.97E-12	11.3206	0.0008
TDEPMAt_1	7.86E-10	6.93E-10	1.2879	0.2564
MAt_1	8.5E-11	1.2E-10	0.5045	0.4775
I_MAt_1	1.556E-9	4.5E-10	11.9763	0.0005
I_MAt_12	-48E-20	9.69E-20	24.5023	<.0001
EDEPBUt_1	3.043E-8	3.81E-9	63.7926	<.0001
EDEPBUt_12	-116E-18	1.59E-17	52.8386	<.0001
ddmtdmt_1	2.52E-12	3.53E-10	0.0001	0.9943
ddmtdmt_12	-782E-22	7.17E-20	1.1884	0.2757
dcat_1	2.08E-11	1.98E-11	1.1051	0.2932
ddmpat_1	2.71E-11	6.39E-10	0.0018	0.9662
ddmpat_12	9.61E-20	2.98E-19	0.1042	0.7469
dclt_1	-23E-13	1.74E-11	0.0175	0.8949
dgnp	1.79E-13	4.67E-13	0.1464	0.7020
FAAB	0.2807	0.00918	934.7207	<.0001
Public	0.2681	0.0351	58.3047	<.0001

ruralare	0.1046	0.0145	51.8783	<.0001
largcity	-0.1321	0.00968	186.3178	<.0001
market	48.8895	12.0085	16.5749	<.0001
marketw	8.5185	3.3697	6.3908	0.0115

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-48E-12-268E-13)=ddmcasht_1*(-7.48E-11)
dcasht_1*(-48E-12+268E-13)=dcasht_1*(-2.12E-11)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	58.2
Percent Discordant	33.0
Percent Tied	8.8

$EDEP^{MA}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	15857	20.01	15857	20.01
positive	63400	79.99	79257	100.00

$PEDEP^{MA}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	15820	19.95	15820	19.96
positive	63437	80.05	79257	100.00

Table 1b: Estimating the level of positive $EDEP^{MA}$

Influence diagnostics (observations with positive $EDEP^{MA}$***)***

Outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	62958	99.30	62958	99.30
1	442	0.70	63400	100.00

Leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	61393	96.83	61393	96.83
1	2007	3.17	63400	100.00

Normality tests (observations with positive $EDEP^{MA}$***)***

Moments

N	63400	Sum Weights	63400
Mean	0	Sum Observations	0
Std Deviation	21099432.9	Variance	4.45186E14
Skewness	23.062561	Kurtosis	2601.00122
Uncorrected SS	2.82244E19	Corrected SS	2.82244E19
Coeff Variation	.	Std Error Mean	83796.5529

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.407672	Pr> D	<0.0100
Cramer-von Mises	W-Sq	4106.005	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	19428.74	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive $EDEP^{MA}$***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	18	1.347E18	7.485E16	30128.5	<.0001
Residual	63382	1.625E17	2.564E12		
Uncorrected Total	63400	1.51E18			
Corrected Total	63399	1.476E18			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	374143	49371.0	277374	470912
sumcasht_1	0.000028	0.000044	-0.00006	0.000115
diffcasht_1	-0.00003	0.000011	-0.00005	-0.00001
tdepmt_1	0.5419	0.00200	0.5380	0.5458
MAt_1	0.0288	0.000229	0.0283	0.0292
I_MAt_1	0.0563	0.00109	0.0542	0.0585
I_MAt_12	-232E-13	6.09E-13	-244E-13	-22E-12
EDEPBUt_1	-0.00106	0.00379	-0.00850	0.00638
EDEPBUt_12	1.74E-10	2.25E-11	1.29E-10	2.18E-10
ddmtdmt_1	-0.00065	0.000831	-0.00228	0.000980
ddmtdmt_12	-129E-13	1.46E-13	-132E-13	-126E-13
dcat_1	-0.00005	0.000031	-0.00011	8.631E-6
ddmpat_1	0.00165	0.00157	-0.00143	0.00473
ddmpat_12	-748E-14	1.11E-12	-966E-14	-531E-14
dclt_1	0.000035	0.000027	-0.00002	0.000088
dgnp	-1.2E-6	6.927E-7	-2.56E-6	1.53E-7
FAAB	-241316	13422.9	-267626	-215007
Public	512723	50902.4	412953	612494
ruralare	-9549.0	20486.2	-49702.7	30604.7
largcity	-1144.8	14274.0	-29122.4	26832.9
market	29411097	2927201	23673673	35148521
marketw	1.4677E8	4855163	1.3725E8	1.5629E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(0.000028-0.00003)=ddmcasht_1*(-2.0E-6)
dcasht_1*   (0.000028-0.00003)=dcasht_1*(5.8E-5)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	6434	10.15	6434	10.15
1	56966	89.85	63400	100.00

Table 1c: Association of Predicted and Observed $EDEP^{MA}$

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
$EDEP^{MA}$	2164552.96	171555973575	34911990.83	0	4657107196
$PEDEP^{MA}$	1719706.51	136298779006	28391034.14	-27973519.28	4476478618

Table 2: Estimation results for S^{MA}

Table 2a: Multinomial Logit

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	Positive	21452
2	Zero	38497
3	Negative	19308

Analysis Of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Wald 95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept1	-2.6765	0.03290	-2.7410	-2.6120	6617.12	0
Intercept2	-1.1358	0.03189	-1.1983	-1.0733	1268.93	64E-279
sumcasht_1	706E-13	296E-13	126E-13	129E-12	5.69180	0.01704
diffcasht_1	263E-13	659E-14	134E-13	392E-13	15.9159	0.00007
TDEPMAt_1	9.8E-9	542E-12	8.73E-9	1.09E-8	326.665	512E-75
EDEPMAt	-24E-9	1.14E-9	-27E-9	-22E-9	459.048	77E-103
EDEPMAt2	586E-20	349E-21	518E-20	654E-20	282.033	271E-65
MAt_1	-47E-12	41E-12	-13E-11	338E-13	1.29095	0.25587
I_BUt_1	889E-13	379E-13	147E-13	163E-12	5.50999	0.01891
I_BUt_12	135E-22	609E-23	151E-23	254E-22	4.87496	0.02725
EDEPBUt_1	1.15E-8	2.66E-9	6.31E-9	1.68E-8	18.7287	0.00002
EDEPBUt_12	-86E-18	253E-19	-14E-17	-37E-18	11.7105	0.00062
ddmtdmt_1	4.67E-9	261E-12	4.16E-9	5.18E-9	319.326	203E-73
ddmtdmt_12	-18E-19	188E-21	-21E-19	-14E-19	86.7326	124E-22
dcat_1	648E-14	171E-13	-27E-12	401E-13	0.14297	0.70535
ddmpat_1	208E-12	446E-12	-67E-11	1.08E-9	0.21835	0.64030
ddmpat_12	-28E-20	219E-21	-71E-20	149E-21	1.63063	0.20162
dclt_1	231E-13	156E-13	-74E-13	536E-13	2.20021	0.13799
dgnp	222E-13	446E-15	213E-13	23E-12	2471.66	0
FAAB	0.04094	0.00878	0.02374	0.05815	21.7502	3.11E-6
Public	-0.2894	0.03555	-0.3591	-0.2197	66.2560	396E-18
ruralare	-0.0192	0.01379	-0.0462	0.00784	1.93725	0.16397
largcity	-0.0135	0.00926	-0.0317	0.00463	2.13132	0.14432
market	-1.9939	2.58990	-7.0700	3.08221	0.59270	0.44138
marketw	-3.7895	1.80063	-7.3187	-0.2603	4.42914	0.03533

NOTE: The scale parameter was held fixed.

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(706E-13+263E-13)=ddmcasht_1*(9.69E-11)
dcasht_1*(706E-13+263E-13)=dcasht_1*(4.43E-11)

Table 2b: Estimating the level of positive S^{MA}

Influence diagnostics (observations with positive S^{MA})

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	21177	98.72	21177	98.72
1	275	1.28	21452	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	20575	95.91	20575	95.91
1	877	4.09	21452	100.00

Normality tests (observations with positive S^{MA})

Moments

N	21452	Sum Weights	21452
Mean	0	Sum Observations	0
Std Deviation	61585166.2	Variance	3.79273E15
Skewness	-0.7786936	Kurtosis	637.997519
Uncorrected SS	8.13579E19	Corrected SS	8.13579E19
Coeff Variation	.	Std Error Mean	420476.979

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.406554	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1230.898	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	5872.558	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive S^{MA})

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	22	8.315E20	3.78E19	1186466	<.0001
Residual	21430	7.144E17	3.333E13		
Uncorrected Total	21452	8.323E20			
Corrected Total	21451	8.313E20			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-504825	297714	-1088378	78728.3
sumcasht_1	-0.00005	0.000258	-0.00055	0.000460
diffcasht_1	0.000050	0.000063	-0.00007	0.000175
EDEPMAt	-4.3436	0.0111	-4.3653	-4.3219
MAt_1	0.9754	0.00110	0.9733	0.9776
I_BU_t_1	-0.00042	0.000302	-0.00101	0.000172
I_BU_t_12	-786E-16	4.52E-14	-167E-15	1.01E-14
EDEPBUt_1	0.0509	0.0191	0.0134	0.0883
EDEPBUt_12	3.28E-11	1.17E-10	-197E-12	2.63E-10

ddmtdmt_1	0.0495	0.00384	0.0419	0.0570
dcat_1	0.00176	0.000167	0.00143	0.00209
ddmpat_1	0.00213	0.00849	-0.0145	0.0188
ddmpat_12	-196E-13	1.15E-11	-422E-13	3.02E-12
dclt_1	0.000419	0.000166	0.000093	0.000745
dclt_12	1.54E-13	1.47E-14	1.25E-13	1.83E-13
dgnp	0.000020	4.25E-6	0.000012	0.000029
FAAB	-398468	84054.7	-563225	-233712
Public	562959	303316	-31573.6	1157492
ruralare	-142176	128785	-394608	110256
Largcity	252269	88941.4	77934.5	426604
market	-3.723E8	1.3419E8	-6.353E8	-1.093E8
marketw	-4.499E7	27915063	-9.971E7	9726819

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00005+0.000050)=ddmcasht_1*0
dcasht_1*(-0.00005-0.000050)=dcasht_1*(-1.0E-4)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	2550	11.89	2550	11.89
1	18902	88.11	21452	100.00

Table 2c: Estimating the level of negative SMA^{MA}

Influence diagnostics (observations with negative S^{MA})

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	19291	99.91	19291	99.91
1	17	0.09	19308	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	18818	97.46	18818	97.46
1	490	2.54	19308	100.00

Normality tests (observations with negative S^{MA})

Moments

N	19308	Sum Weights	19308
Mean	0	Sum Observations	0
Std Deviation	731612684	Variance	5.35257E17
Skewness	-133.03053	Kurtosis	18173.7156
Uncorrected SS	1.03342E22	Corrected SS	1.03342E22
Coeff Variation	.	Std Error Mean	5265172.22

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.458626	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1484.892	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	6944.358	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative S^{MA}***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	21	5.137E19	2.446E18	10417.4	<.0001
Residual	19287	4.725E18	2.45E14		
Uncorrected Total	19308	5.61E19			
Corrected Total	19307	5.576E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-3346287	810320	-4934614	-1757959
sumcasht_1	-0.00013	0.000754	-0.00161	0.00135
diffcasht_1	-0.00006	0.000148	-0.00035	0.000229
sumcaclt_1	0.000140	0.000197	-0.00025	0.000526
diffcaclt_1	-0.00039	0.000341	-0.00106	0.000274
TDEPMAt_1	2.0626	0.0269	2.0099	2.1154
EDEPMAt	-3.6128	0.0227	-3.6573	-3.5684
EDEPMAt2	1.47E-10	1.24E-11	1.23E-10	1.71E-10
MAt_1	0.0327	0.00362	0.0256	0.0398
I_BU_t_1	-0.00007	0.000813	-0.00166	0.00152
EDEPBUt_1	0.4997	0.0588	0.3845	0.6148
EDEPBUt_12	-5.56E-9	3.12E-10	-6.17E-9	-4.95E-9
ddmtdmt_1	0.0547	0.0105	0.0341	0.0753
ddmpat_1	0.0254	0.0211	-0.0159	0.0668
dgnp	0.000049	0.000012	0.000025	0.000073
FAAB	360774	237705	-105157	826706
Public	-210165	722860	-1627060	1206730
ruralare	-236983	364630	-951703	477737
Largcity	624165	251204	131773	1116556
market	-5.095E8	2.1518E8	-9.313E8	-8.773E7
marketw	-1223215	32763502	-6.544E7	62997274

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00013-0.00006) =ddmcasht_1*-1.9E-4
dcasht_1*(-0.00013+0.00006)=-dcasht_1*7.0E-5
dcat_1*(0.00014-0.00039)=dcat_1*-2.5E-4
dclt_1*(0.00014+0.00039)=dclt_1*5.3E-4

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	1299	6.73	1299	6.73
1	18009	93.27	19308	100.00

Table 2d: **Association of Predicted and Observed** S^{MA}

Variable	*Mean*	***Sum***	*Std Dev*	*Minimum*	*Maximum*
S^{MA}	-1645164.33	-1.303908E11	387737778	-1.001795E11	27738387911
PS^{MA}	330153.28	26166958492	107218176	-6404227018	13440778197

Table 3: Estimation results for I^{MA}

Table 3a: Tobit Regression

Analysis of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Marginal effects*	*Chi- Square*	*Pr > ChiSq*
Intercept	-1.92E6	143855		177.430	176E-42
sumcasht_1	-0.0003	0.00013		4.69292	0.03029
diffcasht_1	-183E-7	0.00003		0.39840	0.52792
smat	-0.0002	0.00005	-0.000100005	9.50485	0.00205
I_BUt_1	0.00012	0.00015	0.0000600009	0.61623	0.43245
EDEPBUt_1	0.05629	0.01237	0.0282260403	20.6909	5.4E-6
EDEPBUt_12	-68E-11	865E-13	-3.39308E-10	62.3373	289E-17
EDEPMAt	1.07320	0.00590	0.7239437802	33082.2	0
TDEPMAt_1	0.01086	0.00170	0.0054520271	40.7896	17E-11
TDEPMAt_12	-96E-13	566E-15	-4.78457E-12	285.890	391E-66
ddmtdmt_1	0.00085	0.00090	0.0004250618	0.89368	0.34448
dcat_1	-273E-7	0.00007	-0.00001365	0.14701	0.70141
ddmpat_1	0.00134	0.00162	0.0006700059	0.67900	0.40993
ddmpat_12	374E-15	582E-15	1.870036E-13	0.41168	0.52112
dclt_1	0.00002	0.00007	0.0000100001	0.11819	0.73100
dgnp	1.35E-6	2.02E-6	6.8467709E-7	0.44442	0.50500
FAAB	381211	39612.9	195764.37443	92.6100	637E-24
Public	592649	149238	296816.7482	15.7701	0.00007
ruralare	147754	60550.9	74092.469452	5.95436	0.01468
largcity	-352519	41927.9	-171693.8005	70.6902	418E-19
market	1.998E8	1.699E7	100238432.31	138.295	628E-34
marketw	1616473	8836311	808290.00217	0.03347	0.85485
Scale	3186957	14431.6			

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.0003-183E-7)=ddmcasht_1*-3.183E-4
dcasht_1*(-0.0003+183E-7)=-dcasht_1*2.817E-4

The marginal effects of ddmcasht_1 and dcasht_1 are -0.000159113 and -0.000140831 respectively.

Association of Predicted Probabilities and Observed Responses

I^{MA}	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	33366	42.10	33366	42.10
positive	45891	57.90	79257	100.00

PI^{MA}	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	46037	58.09	46037	58.09
Positive	33220	41.91	79257	100.00

Table 3b: *Association of Predicted and Observed* I^{MA}

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
I^{MA}	2985313.01	236606953116	67211295.62	0	15377560354
PI^{MA}	2439182.18	193322261876	28282487.98	0	2230869900

Table 4: Estimation results for $EDEP^{BU}$

Table 4a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	17847
2	0	61410

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	84550.981	79488.471
SC	84560.261	79664.800
-2LogL	84548.981	79450.471

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	5098.5095	18	<.0001
Score	5234.4852	18	<.0001
Wald	3911.7129	18	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.2451	0.0558	498.1982	<.0001
sumcasht_1	1.88E-11	3.84E-11	0.2397	0.6244
diffcasht_1	-133E-13	1E-11	1.7768	0.1825
edep_mat	1.178E-8	8.03E-10	215.2480	<.0001
EDEPMAt2	-242E-20	1.56E-19	240.3745	<.0001
SMAt	-223E-13	1.22E-11	3.3500	0.0672
IMAt	-803E-12	2.78E-10	8.3133	0.0039
BUt_1	1.686E-9	8.05E-11	438.6655	<.0001
BUt_12	-203E-21	1.44E-20	197.5080	<.0001
dcat_1	5.76E-11	2.58E-11	4.9736	0.0257
ddmpat_1	-739E-12	6.85E-10	1.1630	0.2809
dclt_1	3.42E-11	2.41E-11	2.0231	0.1549
dgnp	2.12E-12	7.91E-13	7.1721	0.0074
FAAB	-0.2074	0.0156	176.0553	<.0001
Public	0.1722	0.0561	9.4392	0.0021
ruralare	0.3906	0.0199	383.7119	<.0001
largcity	-0.7356	0.0176	1740.4872	<.0001
market	93.6332	11.5221	66.0387	<.0001

marketw	2.5334	2.1943	1.3329	0.2483

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(1.88E-11-133E-13)=ddmcasht_1*(5.5E-12)
dcasht_1*(1.88E-11+133E-13)=3dcasht_1*(.2E-11)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.7	Somers'D	0.354
Percent Discordant	31.3	Gamma	0.361
Percent Tied	2.0	Tau-a	0.123
Pairs	1095984270	c	0.677

$EDEP^{BU}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
Zero	61410	77.48	61410	77.48
Positive	17847	22.52	79257	100.00

$PEDEP^{BU}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
Zero	61591	77.71	61591	77.71
Positive	17666	22.29	79257	100.00

Table 4b: Estimating the level of positive $EDEP^{BU}$

Influence diagnostics (observations with positive $EDEP^{BU}$ ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	17569	98.44	17569	98.44
1	278	1.56	17847	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	17223	96.50	17223	96.50
1	624	3.50	17847	100.00

Normality tests (observations with positive $EDEP^{BU}$ ***)***

Moments

N	17847	Sum Weights	17847
Mean	0	Sum Observations	0
Std Deviation	5647640.74	Variance	3.18958E13
Skewness	6.50294865	Kurtosis	449.131457
Uncorrected SS	5.69213E17	Corrected SS	5.69213E17
Coeff Variation	.	Std Error Mean	42275.0812

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.373814	Pr> D	<0.0100
Cramer-von Mises	W-Sq	933.0834	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	4469.001	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive $EDEP^{BU}$ ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	19	1.284E17	6.757E15	14060.0	<.0001
Residual	17828	8.743E15	4.904E11		
Uncorrected Total	17847	1.371E17			
Corrected Total	17846	1.329E17			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	335770	41448.4	254526	417015
sumcasht_1	-0.00011	0.000041	-0.00019	-0.00003
diffcasht_1	-0.00002	9.773E-6	-0.00004	-4.55E-6
sumcaclt_1	-0.00004	9.461E-6	-0.00006	-0.00003
diffcaclt_1	-0.00005	0.000021	-0.00009	-6.22E-6
EDEPMAt	0.0472	0.000503	0.0462	0.0482
EDEPMAt2	-114E-13	1.11E-13	-116E-13	-112E-13
SMAt	-0.00276	0.000061	-0.00288	-0.00264
SMAt2	-191E-15	8.31E-15	-207E-15	-175E-15
IMAt	0.00539	0.000080	0.00523	0.00554
BUt_1	0.0219	0.000072	0.0217	0.0220
ddmpat_1	-0.00501	0.00106	-0.00708	-0.00294
Dgnp	-3.8E-6	5.897E-7	-4.96E-6	-2.64E-6
FAAB	-39495.9	11354.4	-61752.1	-17239.7
Public	113901	47010.8	21753.7	206048
ruralare	-13568.7	14412.6	-41819.3	14681.9
largcity	-3824.7	13148.5	-29597.6	21948.2
market	1611088	1012804	-374142	3596317
marketw	10788356	1956467	6953420	14623292

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00011-0.00002)=ddmcasht_1*-1.3E-4
dcasht_1*(-0.00011+0.00002)=dcasht_1*-9.0E-5

dcat_1*(-0.00004-0.00005)=dcat_1*-9.0E-5
dclt_1*(-0.00004+0.00005)=dclt_1*1.0E-5

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	3429	19.21	3429	19.21
1	14418	80.79	17847	100.00

Table 4c: Association of Predicted and Observed $EDEP^{BU}$

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
$EDEP^{BU}$	373828.87	29628554820	4455424.75	0	328088487
$PEDEP^{BU}$	217608.30	17246981021	6786438.77	-1639456960	320940471

Table 5: Estimation results for I^{BU}

Table 5a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	17157
2	0	62100

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	82810.912	73947.609
SC	82820.192	74133.218
-2LogL	82808.912	73907.609

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	8901.3030	19	<.0001
Score	6950.4865	19	<.0001
Wald	5240.2500	19	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	1.2931	0.0571	511.9621	<.0001
sumcasht_1	1.01E-10	3.66E-11	7.5486	0.0060
diffcasht_1	8.53E-13	9.23E-12	0.0085	0.9264
EDEPMAt	1.263E-8	9.34E-10	182.9705	<.0001
EDEPMAt2	-233E-20	1.98E-19	138.9135	<.0001
SMAt	3.57E-10	8.07E-11	19.5452	<.0001
IMAt	-2.06E-9	1.52E-10	183.4013	<.0001
EDEPBUt	2.861E-7	7.766E-9	1356.9108	<.0001
EDEPBUt2	-951E-18	2.52E-17	1426.1260	<.0001
dcat_1	7.32E-11	2.95E-11	6.1593	0.0131
ddmpat_1	6.89E-10	8.65E-10	0.6352	0.4254
ddmpat_12	-818E-21	3.54E-19	5.3351	0.0209
dclt_1	1.72E-12	2.75E-11	0.0039	0.9502
dgnp	-385E-13	8.49E-13	2055.6801	<.0001
FAAB	-0.1246	0.0162	59.2058	<.0001
Public	-0.6302	0.0793	63.1478	<.0001
ruralare	0.2607	0.0216	145.1048	<.0001
largcity	-0.5440	0.0177	942.4840	<.0001

market	2.3473	2.2870	1.0534	0.3047
marketw	11.9200	2.6530	20.1877	<.0001

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(1.01E-10+8.53E-13)=ddmcasht_1*1.01853E-10
dcasht_1*(1.01E-10-8.53E-13)=dcasht_1*1.00147E-10

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.4	Somers'D	0.441
Percent Discordant	27.2	Gamma	0.448
Percent Tied	1.4	Tau-a	0.150
Pairs	1065449700	c	0.721

I^{BU}	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	62100	78.35	62100	78.35
$\neq 0$	17157	21.65	79257	100.00

PI^{BU}	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	62095	78.35	62095	78.35
$\neq 0$	17162	21.65	79257	100.00

Table 5b: Estimating the level of positive I^{BU}

Influence diagnostics (observations with positive I^{BU})

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	16866	98.30	16866	98.30
1	291	1.70	17157	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	16417	95.69	16417	95.69
1	740	4.31	17157	100.00

Normality tests (observations with positive I^{BU})

Moments

N	17157	Sum Weights	17157
Mean	0	Sum Observations	0
Std Deviation	251567376	Variance	6.32861E16
Skewness	-2.0273255	Kurtosis	529.458839
Uncorrected SS	1.08574E21	Corrected SS	1.08574E21
Coeff Variation	.	Std Error Mean	1920585.49

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.371838	Pr> D	<0.0100
Cramer-von Mises	W-Sq	943.2938	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	4480.941	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive I^{BU})

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	19	5.992E18	3.154E17	541.62	<.0001
Residual	17138	1.041E19	6.077E14		
Uncorrected Total	17157	1.641E19			
Corrected Total	17156	1.634E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	4406689	1394837	1672616	7140761
sumcasht_1	-0.00155	0.00137	-0.00425	0.00114
diffcasht_1	-0.00037	0.000286	-0.00094	0.000187
sumcaclt_1	-0.00037	0.000302	-0.00096	0.000226
diffcaclt_1	-0.00210	0.000562	-0.00320	-0.00100
EDEPMAt	-0.2362	0.0132	-0.2620	-0.2104
EDEPMAt2	3.62E-11	2.56E-12	3.12E-11	4.13E-11
SMAt	-0.1643	0.00217	-0.1686	-0.1601
IMAt	0.0973	0.00560	0.0863	0.1082
EDEPBUt	2.9552	0.0713	2.8155	3.0949
EDEPBUt2	-1.38E-8	3.25E-10	-1.45E-8	-1.32E-8
ddmpat_1	0.00783	0.0315	-0.0540	0.0697
dgnp	-0.00005	0.000021	-0.00009	-0.00001
FAAB	-947456	406170	-1743605	-151308
Public	77806.5	1460661	-2785290	2940903
ruralare	-442724	541380	-1503904	618457
largcity	-711521	448280	-1590212	167170
market	3.222E8	4.4098E8	-5.422E8	1.1866E9
marketw	-1.607E8	59028936	-2.764E8	-4.501E7

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00155-0.00037)=ddmcasht_1*-1.92E-3
dcasht_1*(-0.00155+0.00037)=dcasht_1*-1.18E-3

dcat_1*(-0.00037-0.00210)=dcat_1*-2.47E-3
dclt_1*(-0.00037+0.00210)=dclt_1*1.73E-3

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	2612	15.22	2612	15.22
1	14545	84.78	17157	100.00

Table 5c: Association of Predicted and Observed I^{BU}

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
I^{BU}	1111386.90	88085191902	154251931	-23953208569	14374105758
PI^{BU}	615033.70	48745726307	25966895.91	-5086955062	1286936974

Table 6: Estimation results for *dofa*

Table 6a: Logistic model with complementary log-log function (Gompertz) (observations with positive *dofa*)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	25951
2	0	53306

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	100237.15	97374.353
SC	100246.43	97513.560
-2LogL	100235.15	97344.353

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	2890.7952	14	<.0001
Score	3015.0396	14	<.0001
Wald	3133.2627	14	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.2435	0.0380	1068.6748	<.0001
sumcasht_1	-213E-13	3.2E-11	0.4430	0.5057
diffcasht_1	-225E-13	6.79E-12	11.0255	0.0009
ddmpat_1	6.18E-10	8.41E-10	0.5394	0.4627
ddmpat_12	-248E-21	5.2E-19	0.2268	0.6339
ddmpat_13	-832E-31	1.96E-28	0.1810	0.6706
DIMA	0.1374	0.0131	110.3992	<.0001
DIBU	0.6276	0.0139	2035.2534	<.0001
realr	2.0145	0.7748	6.7608	0.0093
FAAB	-0.1309	0.0129	102.2693	<.0001
Public	0.9012	0.0363	616.7370	<.0001
ruralare	-0.0134	0.0202	0.4412	0.5065
largcity	0.0604	0.0137	19.3173	<.0001
market	14.5222	12.3659	1.3792	0.2402
marketw	-4.0786	1.8176	5.0352	0.0248

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-213E-13-225E-13) =ddmcasht_1*-4.38E-11
dcasht_1*(-213E-13+225E-13) = dcasht_1*1.2E-12

Association of Predicted Probabilities and Observed Responses

Percent Concordant	59.2	Somers'D	0.205
Percent Discordant	38.6	Gamma	0.210
Percent Tied	2.2	Tau-a	0.090
Pairs	1383344006	c	0.603

Table 6b: Logistic model with complementary log-log function (Gompertz) (observations with negative and zero *dofa*)

Number of Response Levels:	2
Number of Observations:	53306
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	18146
2	0	35160

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	68373.166	61431.523
SC	68382.050	61555.896
-2LogL	68371.166	61403.523

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	6967.6430	13	<.0001
Score	7202.0413	13	<.0001
Wald	7875.4733	13	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.4525	0.0459	1001.0052	<.0001
sumcasht_1	2.02E-10	6.15E-11	10.7607	0.0010
diffcasht_1	2.85E-11	1.09E-11	6.8084	0.0091
ddmpat_1	-7.94E-9	1.904E-9	17.3898	<.0001
ddmpat_12	1.13E-17	2.87E-18	15.5899	<.0001
DIMA	0.1216	0.0157	59.7371	<.0001
DIBU	1.4135	0.0166	7254.8394	<.0001
realr	2.8988	0.9338	9.6368	0.0019
FAAB	0.0272	0.0156	3.0393	0.0813
Public	0.8366	0.0653	164.3235	<.0001
ruralare	-0.0403	0.0244	2.7343	0.0982
largcity	0.0978	0.0167	34.3516	<.0001
market	3.1548	16.7428	0.0355	0.8505
marketw	12.8676	3.0591	17.6933	<.0001

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(2.02E-10+2.85E-11) = ddmcasht_1*2.305E-10
dcasht_1*(2.02E-10-2.85E-11) = dcasht_1*1.735E-10

Association of Predicted Probabilities and Observed Responses

Percent Concordant	65.3	Somers'D	0.340
Percent Discordant	31.4	Gamma	0.351
Percent Tied	3.3	Tau-a	0.152
Pairs	638013360	c	0.670

dofat	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	18146	22.90	18146	22.90
Zero	35160	44.36	53306	67.26
positive	25951	32.74	79257	100.00

pdofat	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	18253	23.03	18253	23.03
zero	35038	44.21	53291	67.24
Positive	25966	32.76	79257	100.00

Table 6c: Estimating the level of positive *dOFA*

Influence diagnostics (observations with positive dOFA)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	25937	99.95	25937	99.95
1	14	0.05	25951	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	24905	95.97	24905	95.97
1	1046	4.03	25951	100.00

Normality tests (observations with positive** dOFA**)

Moments

N	25951	Sum Weights	25951
Mean	0	Sum Observations	0
Std Deviation	7016039171	Variance	4.92248E19
Skewness	151.992423	Kurtosis	23820.9964
Uncorrected SS	1.27738E24	Corrected SS	1.27738E24
Coeff Variation	.	Std Error Mean	43552687.2

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.460984	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1985.074	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	9300.079	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** dOFA**)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	15	1.088E20	7.251E18	321.30	<.0001
Residual	25936	5.808E20	2.239E16		
Uncorrected Total	25951	6.896E20			
Corrected Total	25950	6.815E20			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	21954687	5900918	10388349	33521025
sumcasht_1	0.1873	0.00750	0.1726	0.2020
diffcasht_1	0.0222	0.00142	0.0194	0.0250
ddmpat_1	0.8256	0.3596	0.1207	1.5305
ddmpat_12	2.487E-8	6.38E-10	2.362E-8	2.612E-8
ddmpat_13	-571E-19	2.15E-18	-613E-19	-529E-19
DIMA	-8357507	1965808	-1.221E7	-4504343
DIBU	7714556	2087091	3623668	11805445
realr	-6.399E7	1.2047E8	-3.001E8	1.7214E8
FAAB	-1.488E7	1930454	-1.866E7	-1.11E7
Public	81934251	5520681	71113212	92755291
ruralare	-3418444	3012336	-9322897	2486009
largcity	9059620	2069878	5002470	13116770
market	-5.107E7	2.1592E8	-4.743E8	3.7215E8
marketw	1.124E9	2.6669E8	6.013E8	1.6468E9

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(0.1873+0.0222) =ddmcasht_1*0.2095
dcasht_1*(0.1873-0.0222) = dcasht_1*0.1651

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	909	3.50	909	3.50
1	25042	96.50	25951	100.00

Table 6d: Estimating the level of negative *dOFA*

Influence diagnostics (observations with negative** dOFA**)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	18018	99.29	18018	99.29
1	128	0.71	18146	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	17660	97.32	17660	97.32
1	486	2.68	18146	100.00

Normality tests (observations with negative** dOFA**)

Moments

N	18146	Sum Weights	18146
Mean	0	Sum Observations	0
Std Deviation	678804274	Variance	4.60775E17
Skewness	-42.351732	Kurtosis	2647.12688
Uncorrected SS	8.36077E21	Corrected SS	8.36077E21
Coeff Variation	.	Std Error Mean	5039113.2

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.422701	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1238.932	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	5825.842	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative** dOFA**)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	14	8.935E20	6.382E19	48751.6	<.0001
Residual	18132	2.553E19	1.408E15		
Uncorrected Total	18146	9.19E20			
Corrected Total	18145	9.18E20			

Parameter	Estimate	Approx Std Error	Approx 95% Confidence Limits	
Intercept	-5528631	1865502	-9185258	-1872003
sumcasht_1	-0.3116	0.00686	-0.3251	-0.2982
diffcasht_1	-0.0904	0.000548	-0.0915	-0.0893
ddmpat_1	0.3269	0.0217	0.2844	0.3695
ddmpat_12	6.07E-10	6.35E-12	5.95E-10	6.2E-10
DIMA	1317718	600543	140574	2494861
DIBU	-1074584	621136	-2292092	142924
realr	45094253	38242045	-2.987E7	1.2005E8
FAAB	-81077.9	590756	-1239037	1076881
Public	-4261200	2495083	-9151889	629489
ruralare	250037	919009	-1551340	2051414
largcity	-1184036	634995	-2428710	60637.7
market	4.9093E9	8.6765E8	3.2085E9	6.61E9
marketw	-6.6E9	1.7651E8	-6.946E9	-6.254E9

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.3116-0.0904) = ddmcasht_1*-0.402
dcasht_1*(-0.3116+0.0904) = dcasht_1*-0.2212

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	2365	13.03	2365	13.03
1	15781	86.97	18146	100.00

Table 6e: **Association of Predicted and Observed** *dOFA*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dofat	23846605.85	1.8900104E12	4035157262	-54031464978	1.1064462E12
pdofat	2594347.53	205620201844	165284005	-29713860970	4490855663

Table 7: **Estimation results for** *dca*

Table 7a: Estimating the level of *dca*

Influence diagnostics (observations with positive *dca* ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	78652	99.24	78652	99.24
1	605	0.76	79257	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	76931	97.07	76931	97.07
1	2326	2.93	79257	100.00

Normality tests (observations with positive *dca* ***)***

Moments

N	79257	Sum Weights	79257
Mean	0	Sum Observations	0
Std Deviation	464900513	Variance	2.16132E17
Skewness	9.60259674	Kurtosis	3178.93571
Uncorrected SS	1.71298E22	Corrected SS	1.71298E22
Coeff Variation	.	Std Error Mean	1651357.91

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.403584	Pr> D	<0.0100
Cramer-von Mises	W-Sq	5252.372	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	24765.17	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *dca* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	22	2.816E19	1.28E18	2278.49	<.0001
Residual	79235	4.631E19	5.845E14		
Uncorrected Total	79257	7.447E19			
Corrected Total	79256	7.428E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	2364307	679153	1033148	3695467
sumcasht_1	-0.00614	0.00129	-0.00867	-0.00362
diffcasht_1	0.000392	0.000211	-0.00002	0.000806
EDEPMAt	0.1248	0.0188	0.0879	0.1616
EDEPMAt2	8.92E-11	3.36E-12	8.26E-11	9.58E-11
SMAt	-0.5029	0.00398	-0.5107	-0.4951
IMAt	0.4168	0.0104	0.3965	0.4372
EDEPBUt	0.4543	0.0713	0.3145	0.5942

EDEPBUt2	-7.32E-9	6.37E-10	-8.56E-9	-6.07E-9
IBUt	0.0192	0.00195	0.0154	0.0231
IBUt2	2.73E-12	4.3E-13	1.88E-12	3.57E-12
IBUt3	1.38E-21	9.02E-23	1.2E-21	1.56E-21
dclt_1	-0.00598	0.000508	-0.00697	-0.00498
ddmpat_1	-0.0176	0.0149	-0.0467	0.0115
ddmpat_12	-219E-13	1.64E-11	-54E-12	1.02E-11
dgnp	-0.00002	9.515E-6	-0.00004	1.142E-6
FAAB	-1532929	184479	-1894513	-1171345
Public	10225598	818377	8621555	11829642
ruralare	-386018	286128	-946837	174801
largcity	715615	194897	333612	1097618
market	1.2121E9	47791492	1.1184E9	1.3057E9
marketw	-1.991E9	52169075	-2.094E9	-1.889E9

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00614+0.000392) = ddmcasht_1*-5.748E-3
dcasht_1*(-0.00614-0.000392) = dcasht_1*-6.532E-3

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	11186	14.11	11186	14.11
1	68071	85.89	79257	100.00

Table 7c: Association of Predicted and Observed *dca*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dcat	8093669.95	641479998924	1090458896	-1.399818E11	200112041817
pdcat	2862281.33	226855831152	229821137	-33473561880	50397492609

Table 8: Estimation results for *dll*

Table 8a: Logistic model with complementary log-log function (Normal)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	56852
2	0	22405

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	94392.983	88036.817
SC	94402.264	88185.304
-2LogL	94390.983	88004.817

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	6386.1668	15	<.0001
Score	5760.8053	15	<.0001
Wald	5219.4346	15	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	0.6135	0.0484	160.7127	<.0001
sumcasht_1	2.98E-11	6.55E-11	0.2076	0.6487
diffcasht_1	5.02E-12	1.23E-11	0.1658	0.6839
ddmpat_1	3.66E-10	2.482E-9	0.0217	0.8829
ddmpat_12	2.29E-18	2.52E-18	0.8269	0.3632
ddmpat_13	6.46E-29	5.34E-27	0.0001	0.9903
DIMA	0.4924	0.0165	890.6469	<.0001
DIBU	1.0961	0.0268	1672.8169	<.0001
Ddofa	0.3626	0.0170	454.6333	<.0001
realr	-8.5213	0.9861	74.6774	<.0001
FAAB	0.4011	0.0171	552.7391	<.0001
Public	0.1323	0.0675	3.8452	0.0499
ruralare	0.2004	0.0290	47.7102	<.0001
largcity	-0.2784	0.0178	245.4264	<.0001
market	1.7242	3.9601	0.1896	0.6633
marketw	97.0519	14.4729	44.9675	<.0001

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(2.98E-11+5.02E-12) = ddmcasht_1*(3.482E-11
dcasht_1*(2.98E-11-5.02E-12) = dcasht_1*(2.478E-11

Association of Predicted Probabilities and Observed Responses

Percent Concordant	67.0	Somers'D	0.355
Percent Discordant	31.6	Gamma	0.360
Percent Tied	1.4	Tau-a	0.144
Pairs	1273769060	c	0.677

dllt	Frequency	Percent	Cumulative Frequency	Cumulative Percent
zero	22405	28.27	22405	28.27
positive	56852	71.73	79257	100.00

pdllt	Frequency	Percent	Cumulative Frequency	Cumulative Percent
zero	22325	28.17	22325	28.17
Positive	56932	71.83	79257	100.00

Table 8b: Estimating the level of positive *dll*

Influence diagnostics (observations with positive *dll* ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	56824	99.95	56824	99.95
1	28	0.05	56852	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	55483	97.59	55483	97.59
1	1369	2.41	56852	100.00

Normality tests (observations with positive *dll* ***)***

Moments

N	56852	SumWeights	56852
Mean	0	SumObservations	0
Std Deviation	4158047476	Variance	1.72894E19
Skewness	235.072236	Kurtosis	55781.86
Uncorrected SS	9.82917E23	CorrectedSS	9.82917E23
Coeff Variation	.	StdErrorMean	17438798.9

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.466017	Pr> D	<0.0100
Cramer-von Mises	W-Sq	4421.273	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	20663.78	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** dll **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	16	3.536E18	2.21E17	44.18	<.0001
Residual	56836	2.852E20	5.018E15		
Uncorrected Total	56852	2.888E20			
Corrected Total	56851	2.885E20			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-1.11E7	1828690	-1.468E7	-7514234
sumcasht_1	-0.00024	0.00269	-0.00551	0.00503
diffcasht_1	-0.00123	0.000469	-0.00215	-0.00031
ddmpat_1	0.1907	0.0419	0.1086	0.2729
ddmpat_12	1.16E-11	2.34E-11	-342E-13	5.75E-11
ddmpat_13	-412E-22	7.99E-21	-569E-22	-256E-22
DIMA	1013777	627187	-215536	2243090
DIBU	-659890	735908	-2102302	782521
Ddofa	951614	649024	-320500	2223728
Realr	3.247E8	37105895	2.5197E8	3.9743E8
FAAB	-6590658	616287	-7798606	-5382709
Public	14494032	2368139	9852380	19135683
ruralare	-344517	926130	-2159770	1470736
largcity	245538	662574	-1053136	1544211
market	8.334E8	1.3259E8	5.7352E8	1.0933E9
marketw	-663913	1.1168E8	-2.196E8	2.1824E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00024-0.00123) = ddmcasht_1*-1.47E-3
dcasht_1*(-0.00024-0.00123) = dcasht_1*9.9E-4

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	2965	5.22	2965	5.22
1	53887	94.78	56852	100.00

Table 8c: Association of Predicted and Observed *dll*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dllt	16822279.61	1.3332834E12	3523901883	-37680257866	986926777678
pdllt	1449731.78	114901391510	6807263.00	-336207889	1352490628

Table 9: **Estimation results for** *dcl*

Table 9a: Estimating the level of *dcl*

Influence diagnostics (observations with positive *dcl* ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	78681	99.27	78681	99.27
1	576	0.73	79257	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	76966	97.11	76966	97.11
1	2291	2.89	79257	100.00

Normality tests (observations with positive *dcl* ***)***

Moments

N	79257	Sum Weights	79257
Mean	0	Sum Observations	0
Std Deviation	356392982	Variance	1.27016E17
Skewness	46.3665716	Kurtosis	8547.95364
Uncorrected SS	1.00668E22	Corrected SS	1.00668E22
Coeff Variation	.	Std Error Mean	1265931.86

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.419865	Pr> D	<0.0100
Cramer-von Mises	W-Sq	5442.903	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	25555.26	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *dcl* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	22	2.672E22	1.214E21	3881865	<.0001
Residual	79235	2.597E19	3.277E14		
Uncorrected Total	79257	2.674E22			
Corrected Total	79256	2.674E22			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	4175407	511558	3172738	5178076
sumcasht_1	-0.00240	0.000924	-0.00421	-0.00059
diffcasht_1	-0.00034	0.000142	-0.00062	-0.00006
EDEPMAt	-0.1182	0.0177	-0.1528	-0.0836
EDEPMAt2	-167E-12	3.19E-12	-173E-12	-161E-12
SMAt	-0.1765	0.00238	-0.1812	-0.1719
IMAt	0.1569	0.00810	0.1410	0.1728
EDEPBUt	0.2457	0.0920	0.0653	0.4261

EDEPBUt2	-3.85E-8	1.093E-9	-4.06E-8	-3.64E-8
IBUt	0.4451	0.00304	0.4391	0.4510
IBUt2	1.04E-10	5.98E-13	1.03E-10	1.05E-10
ddmpat_1	-0.1212	0.0236	-0.1674	-0.0749
ddmpat_12	-376E-12	3.15E-12	-382E-12	-369E-12
ddmpat_13	1.61E-19	2.47E-21	1.56E-19	1.66E-19
dcat	0.6642	0.000651	0.6629	0.6654
dgnp	-0.00006	7.162E-6	-0.00008	-0.00005
FAAB	-226048	139352	-499183	47086.5
Public	-911302	643623	-2172823	350218
ruralare	-60452.5	216435	-484671	363766
largcity	177176	147482	-111894	466246
market	-3.093E7	1.2052E8	-2.672E8	2.053E8
marketw	-1.076E9	65751293	-1.205E9	-9.474E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00240-0.00034) = ddmcasht_1*-2.74E-3
dcasht_1*(-0.00240+0.00034) = dcasht_1*-2.06E-3

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	12817	16.17	12817	16.17
1	66440	83.83	79257	100.00

Table 9c: Association of Predicted and Observed *dcl*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dclt	7839932.89	621369560928	871769194	-24889920428	162947438839
pdclt	7389177.18	585644015925	785209320	-49860658462	162945822985

Table 10: **Estimation results for** *dsc*

Table 10a: Logistic model with complementary log-log function (Gompertz) (observations with positive *dsc*)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	15502
2	0	63755

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	78344.367	75010.173
SC	78353.647	75167.941
-2LogL	78342.367	74976.173

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	3366.1935	16	<.0001
Score	3348.5109	16	<.0001
Wald	3485.2284	16	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-2.1160	0.0520	1653.3357	<.0001
sumcasht_1	-125E-12	5.1E-11	6.0427	0.0140
diffcasht_1	-365E-13	9.44E-12	14.9478	0.0001
ddmpat_1	3.022E-9	2.309E-9	1.7136	0.1905
ddmpat_12	-339E-20	7.48E-18	0.2051	0.6507
ddmpat_13	-49E-29	9.85E-27	0.0025	0.9603
DIMA	0.1965	0.0171	132.7174	<.0001
DIBU	0.3457	0.0191	328.6485	<.0001
Ddofa	0.6536	0.0188	1210.9969	<.0001
Ddll	-0.1320	0.0188	49.3455	<.0001
realr	1.5237	1.0100	2.2757	0.1314
FAAB	-0.1120	0.0167	45.0441	<.0001
Public	1.1631	0.0393	875.3258	<.0001
ruralare	-0.0727	0.0270	7.2616	0.0070
largcity	0.0940	0.0177	28.2837	<.0001
market	3.1124	1.8134	2.9458	0.0861
marketw	-113.7	11.6636	95.0880	<.0001

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-125E-12-365E-13) = ddmcasht_1*-1.615E-10
dcasht_1*(-125E-12+365E-13) = dcasht_1* -8-85E-11

Association of Predicted Probabilities and Observed Responses

Percent Concordant	64.1	Somers'D	0.294
Percent Discordant	34.8	Gamma	0.297
Percent Tied	1.1	Tau-a	0.092
Pairs	988330010	c	0.647

Table 10b: Logistic model with complementary log-log function (Gompertz) (observations with negative and zero *dsc*)

Number of Response Levels:	2
Number of Observations:	63755
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	3717
2	0	60038

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	28343.355	24439.425
SC	28352.418	24593.493
-2LogL	28341.355	24405.425

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	3935.9298	16	<.0001
Score	4142.2291	16	<.0001
Wald	3150.1762	16	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-4.8439	0.1128	1843.0479	<.0001
sumcasht_1	1.53E-10	6.37E-11	5.7746	0.0163
diffcasht_1	7.72E-11	1.57E-11	24.2147	<.0001
ddmpat_1	-1.22E-8	4.422E-9	7.5547	0.0060
ddmpat_12	-319E-19	1.48E-17	4.6206	0.0316

ddmpat_13	4.19E-26	1.9E-26	4.8367	0.0279
DIMA	-0.0460	0.0344	1.7846	0.1816
DIBU	1.1420	0.0353	1047.8281	<.0001
Ddofa	1.6523	0.0535	953.5460	<.0001
Ddll	0.3261	0.0476	46.9416	<.0001
realr	-0.2649	1.9671	0.0181	0.8929
FAAB	0.4363	0.0337	167.4095	<.0001
Public	-0.3393	0.1497	5.1413	0.0234
ruralare	-0.1496	0.0516	8.3873	0.0038
largcity	0.0830	0.0361	5.2856	0.0215
market	-251.9	71.2786	12.4870	0.0004
marketw	8.9469	2.5324	12.4821	0.0004

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(1.53E-10+7.72E-11) =ddmcasht_1*2.302E-10
dcasht_1*(1.53E-10-7.72E-11) = dcasht_1*7.58E-11

Association of Predicted Probabilities and Observed Responses

Percent Concordant	77.6	Somers'D	0.568
Percent Discordant	20.7	Gamma	0.578
Percent Tied	1.7	Tau-a	0.062
Pairs	223161246	c	0.784

dsct	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	3717	4.69	3717	4.69
Zero	60038	75.75	63755	80.44
positive	15502	19.56	79257	100.00

pdsct	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	3682	4.65	3682	4.65
zero	59952	75.64	63634	80.29
Positive	15623	19.71	79257	100.00

Table 10c: Estimating the level of positive *dsc*

Influence diagnostics (observations with positive** *dsc* **)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	15461	99.74	15461	99.74
1	41	0.26	15502	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	15169	97.85	15169	97.85
1	333	2.15	15502	100.00

Normality tests (observations with positive *dsc* ***)***

Moments

N	15502	Sum Weights	15502
Mean	0	Sum Observations	0
Std Deviation	454619543	Variance	2.06679E17
Skewness	90.9526555	Kurtosis	9566.03173
Uncorrected SS	3.20373E21	Corrected SS	3.20373E21
Coeff Variation	.	Std Error Mean	3651356.48

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.438397	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1078.228	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	5104.103	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *dsc* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	16	3.957E17	2.473E16	27.01	<.0001
Residual	15486	8.025E18	5.182E14		
Uncorrected Total	15502	8.42E18			
Corrected Total	15501	8.235E18			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	978406	2162395	-3260219	5217031
sumcasht_1	0.00230	0.00163	-0.00090	0.00551
diffcasht_1	0.000645	0.000301	0.000055	0.00124
ddmpat_1	0.0741	0.0821	-0.0868	0.2351
ddmpat_12	-139E-12	2.18E-10	-567E-12	2.89E-10
DIMA	-9809.5	391404	-777022	757403
DIBU	2452522	439441	1591150	3313893
Ddofa	2184568	430052	1341600	3027536
Ddll	878469	429537	36511.7	1720427
realr	15982915	45427186	-7.306E7	1.0503E8
FAAB	-4293344	381123	-5040403	-3546285
Public	9963197	922574	8154812	11771583
ruralare	-518435	616492	-1726854	689984
largcity	699912	406947	-97766.6	1497591
market	1.9802E8	46271652	1.0732E8	2.8872E8
marketw	19733601	51875669	-8.195E7	1.2142E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(0.0023+0.000645) = ddmcasht_1*2.94E-3
dcasht_1*(0.0023-0.000645) = dcasht_1*1.655E-3

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	663	4.28	663	4.28
1	14839	95.72	15502	100.00

Table 10d: Estimating the level of negative dsc

Influence diagnostics (observations with negative dsc ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	3643	98.01	3643	98.01
1	74	1.99	3717	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	3527	94.89	3527	94.89
1	190	5.11	3717	100.00

Normality tests (observations with negative dsc ***)***

Moments

N	3717	Sum Weights	3717
Mean	0	Sum Observations	0
Std Deviation	183743443	Variance	3.37617E16
Skewness	-13.833062	Kurtosis	325.549531
Uncorrected SS	1.25458E20	Corrected SS	1.25458E20
Coeff Variation	.	Std Error Mean	3013807.92

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.365384	Pr> D	<0.0100
Cramer-von Mises	W-Sq	197.974	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	948.6574	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative dSC ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	15	1.182E18	7.881E16	123.57	<.0001
Residual	3702	2.15E18	5.807E14		
Uncorrected Total	3717	3.332E18			
Corrected Total	3716	3.154E18			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-7898247	9121644	-2.578E7	9986014
sumcasht_1	-0.0732	0.00546	-0.0839	-0.0625
diffcasht_1	-0.0245	0.000911	-0.0263	-0.0227

ddmpat_1	0.1962	0.0742	0.0506	0.3417
DIMA	-170817	855053	-1847268	1505633
DIBU	-1436539	865020	-3132533	259454
Ddofa	-1596384	1320278	-4184974	992206
Ddll	-655005	1185800	-2979932	1669923
realr	1.2084E8	1.9528E8	-2.62E8	5.0373E8
FAAB	1438075	846889	-222369	3098519
Public	-750417	3753915	-8110493	6609659
ruralare	-379241	1290347	-2909147	2150665
largcity	-1167779	898192	-2928809	593252
market	3.9413E9	1.5386E9	9.2466E8	6.9579E9
marketw	-1.916E9	89349912	-2.091E9	-1.741E9

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.0732-0.0245) = ddmcasht_1*-0.096045
dcasht_1*(-0.0732+0.0245) = dcasht_1*-0.144745

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	503	13.53	503	13.53
1	3214	86.47	3717	100.00

Table 10e: **Association of Predicted and Observed** *dsc*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dsct	1836259.98	145536457269	218866630	-19980000000	49997500000
pdsct	110033.85	8720953132	31972705.77	-8161117933	117881274

Table 11: **Estimation results for** *drr*

Table 11a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	35649
2	0	43608

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	109074.94	93835.177
SC	109084.22	93974.384
-2LogL	109072.94	93805.177

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	15267.7621	14	<.0001
Score	14100.8154	14	<.0001
Wald	15404.5306	14	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.6031	0.0378	1794.4957	<.0001
ddmcasht_1	-658E-14	1.2E-11	0.3014	0.5830
ddmcasht_12	1.11E-21	5.44E-22	4.1341	0.0420
DIMA	0.0490	0.0115	18.0928	<.0001
DIBU	0.3412	0.0137	621.6900	<.0001
Ddofa	0.2590	0.0122	449.8682	<.0001
Ddll	0.1141	0.0132	74.8879	<.0001
Ddsc	1.2387	0.0112	12151.7782	<.0001
realr	2.7568	0.7395	13.8958	0.0002
FAAB	0.2314	0.0114	412.1329	<.0001
Public	0.3820	0.0374	104.3883	<.0001
ruralare	-0.00198	0.0178	0.0123	0.9115
largcity	0.00792	0.0122	0.4216	0.5161
market	-1.4597	2.1852	0.4462	0.5041
marketw	-4.8003	1.6744	8.2188	0.0041

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.2	Somers'D	0.450
Percent Discordant	27.1	Gamma	0.454
Percent Tied	0.7	Tau-a	0.223
Pairs	1554581592	c	0.725

drrt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	43608	55.02	43608	55.02
positive	35649	44.98	79257	100.00

pdrrt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	43351	54.70	43351	54.70
positive	35906	45.30	79257	100.00

Table 11b: Estimating the level of positive *drr*

Influence diagnostics (observations with positive* drr *)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	35621	99.92	35621	99.92
1	28	0.08	35649	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	34587	97.02	34587	97.02
1	1062	2.98	35649	100.00

Normality tests (observations with positive* drr *)

Moments

N	35649	Sum Weights	35649
Mean	0	Sum Observations	0
Std Deviation	1342502302	Variance	1.80231E18
Skewness	177.121014	Kurtosis	32603.8253
Uncorrected SS	6.42488E22	Corrected SS	6.42488E22
Coeff Variation	.	Std Error Mean	7110356.29

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.471829	Pr> D	<0.0100
Cramer-von Mises	W-Sq	2761.251	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	12908.06	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** drr **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	15	1.641E19	1.094E18	1787.27	<.0001
Residual	35634	2.336E19	6.556E14		
Uncorrected Total	35649	3.978E19			
Corrected Total	35648	3.977E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-388638	1029706	-2406930	1629654
ddmcasht_1	-0.00398	0.000347	-0.00466	-0.00330
ddmcasht_12	-181E-15	3.36E-15	-188E-15	-174E-15
DIMA	319810	283380	-235634	875254
DIBU	-66042.4	337800	-728153	596068
Ddofa	233849	306650	-367205	834904
Ddll	-261543	326681	-901860	378773
Ddsc	500336	288511	-65165.4	1065837
realr	17095738	20470192	-2.303E7	57218676
FAAB	-1440296	282310	-1993643	-886949
Public	19922935	972164	18017429	21828442
ruralare	-110096	434697	-962132	741940
largcity	279199	301698	-312150	870548
market	5.5484E8	2.4073E8	83000521	1.0267E9
marketw	-2.181E8	48759785	-3.137E8	-1.225E8

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	1239	3.48	1239	3.48
1	34410	96.52	35649	100.00

Table 11c: Association of Predicted and Observed *drr*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
drrt	4937664.41	391344468325	900834626	-13815449999	247908737750
pdrrt	255213.98	20227494096	14634401.95	-3997055830	271864737

Table 12: Estimation results for *OIBD*

Table 12a: Estimating the level of *OIBD*

Influence diagnostics (observations with positive** OIBD**)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	78487	99.03	78487	99.03
1	770	0.97	79257	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	76868	96.99	76868	96.99
1	2389	3.01	79257	100.00

Normality tests (observations with positive** OIBD**)

Moments

N	79257	Sum Weights	79257
Mean	0	Sum Observations	0
Std Deviation	48838769.9	Variance	2.38523E15
Skewness	7.07454062	Kurtosis	2322.62411
Uncorrected SS	1.89043E20	Corrected SS	1.89043E20
Coeff Variation	.	Std Error Mean	173478.597

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.383839	Pr> D	<0.0100
Cramer-von Mises	W-Sq	4895.571	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	23185.47	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** OIBD**)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	27	1.653E19	6.122E17	59970.9	<.0001
Residual	79230	8.326E17	1.051E13		
Uncorrected Total	79257	1.736E19			
Corrected Total	79256	1.722E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	346731	93265.7	163928	529535
sumcaclt_1	0.000305	0.000058	0.000191	0.000419
diffcaclt_1	0.000872	0.000107	0.000661	0.00108
MAt_1	0.0426	0.000693	0.0412	0.0439
I_MAt	0.0658	0.00244	0.0610	0.0706
SMAt	-0.0426	0.000698	-0.0440	-0.0412
EDEPMAt	1.6231	0.00590	1.6116	1.6347
EDEPMAt2	-295E-12	6.98E-13	-296E-12	-294E-12

BUt_1	0.0511	0.000515	0.0501	0.0521
I_BUt	0.0510	0.000516	0.0500	0.0520
EDEPBUt	1.3327	0.0257	1.2822	1.3831
EDEPBUt2	1.495E-9	1.14E-10	1.273E-9	1.718E-9
dcat	0.000746	0.000113	0.000524	0.000967
dcat2	-334E-16	6.11E-16	-346E-16	-322E-16
ddmpat_1	0.00103	0.00204	-0.00298	0.00503
ddmpat_12	-477E-14	3.13E-12	-109E-13	1.37E-12
ddmpat_13	1.62E-21	1.16E-21	-666E-24	3.9E-21
dcasht_1	0.000016	0.000041	-0.00006	0.000095
dcasht_12	1.45E-15	2.1E-15	-267E-17	5.56E-15
dclt	-0.00024	0.000083	-0.00040	-0.00008
dgnp	-2.29E-6	1.31E-6	-4.86E-6	2.798E-7
FAAB	-40563.6	25634.8	-90808.5	9681.3
Public	-4662610	144356	-4945552	-4379667
ruralare	-111537	39382.4	-188728	-34346.4
largcity	22275.8	27183.8	-31005.2	75556.9
market	3.2605E8	29582012	2.6807E8	3.8403E8
marketw	-1.47E7	9572636	-3.347E7	4059674

To be able to interpret the coefficient in front of sumCACLt_1 and diffCACLt_1, we transform these variables to

CAt_1*(0.000305+0.000872) = CAt_1*(1.77E-3)
CLt_1*(0.000305-0.000872) = CLt_1*(-5.67E-4)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	17799	22.46	17799	22.46
1	61458	77.54	79257	100.00

Table 12c: Association of Predicted and Observed *OIBD*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
OIBDt	6550522.15	519174733773	75916333.37	-3064033931	7005478256
pOIBDt	5312982.55	421091058308	51678146.59	-572680489	4305132366

Table 13: Estimation results for *FI*

Table 13a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	69360
2	0	9897

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	59685.929	57664.689
SC	59695.209	57878.140
-2LogL	59683.929	57618.689

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	2065.2395	22	<.0001
Score	532.7658	22	<.0001
Wald	883.8287	22	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	0.4339	0.0270	257.6338	<.0001
I_BUt	1.645E-9	2.8E-10	34.5366	<.0001
EDEPMAt	2.391E-8	3.191E-9	56.1607	<.0001
EDEPMAt2	-614E-20	4.8E-19	163.1547	<.0001
SMAt	1.235E-9	2.11E-10	34.4181	<.0001
I_MAt	2.254E-9	1.125E-9	4.0148	0.0451
I_MAt2	-304E-21	6.46E-20	22.1000	<.0001
EDEPBUt	2.87E-8	1.057E-8	7.3696	0.0066
EDEPBUt2	-417E-18	4.61E-17	81.6531	<.0001
dcat	1.43E-9	1.37E-10	109.4186	<.0001
dcat2	7.35E-21	1.42E-20	0.2683	0.6045
dofat	7.37E-12	2.03E-11	0.1318	0.7165
OFAt_1	6.66E-11	2.63E-11	6.4003	0.0114
CAt_1	1.422E-9	1.33E-10	113.5575	<.0001
MAt_1	-875E-12	2.08E-10	17.6312	<.0001
BUt_1	2.1E-9	2.68E-10	61.2567	<.0001
realr	3.1024	0.5662	30.0264	<.0001
FAAB	0.1562	0.00968	260.2169	<.0001

Public	0.6142	0.0575	114.1134	<.0001
ruralare	0.0185	0.0148	1.5555	0.2123
largcity	0.0502	0.0101	24.8785	<.0001
market	36.9863	13.2654	7.7739	0.0053
marketw	-1.5806	3.2920	0.2305	0.6311

Association of Predicted Probabilities and Observed Responses

Percent Concordant	64.2	Somers'D	0.316
Percent Discordant	32.6	Gamma	0.326
Percent Tied	3.2	Tau-a	0.069
Pairs	686455920	c	0.658

FIt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	9897	12.49	9897	12.49
positive	69360	87.51	79257	100.00

pFIt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	9837	12.41	9837	12.41
positive	69420	87.59	79257	100.00

Table 13b: Estimating the level of positive *FI*

Influence diagnostics (observations with positive FI)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	68960	99.42	68960	99.42
1	400	0.58	69360	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	67125	96.78	67125	96.78
1	2235	3.22	69360	100.00

Normality tests (observations with positive FI)

Moments

N	69360	Sum Weights	69360
Mean	0	Sum Observations	0
Std Deviation	125780535	Variance	1.58207E16
Skewness	56.7535155	Kurtosis	6434.69514
Uncorrected SS	1.09731E21	Corrected SS	1.09731E21
Coeff Variation	.	Std Error Mean	477594.035

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.422514	Pr> D	<0.0100

Cramer-von Mises	W-Sq	4718.127	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	22195.36	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive FI ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	23	1.604E20	6.973E18	259278	<.0001
Residual	69337	1.948E18	2.81E13		
Uncorrected Total	69360	1.623E20			
Corrected Total	69359	1.622E20			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-14640.6	121117	-252035	222753
I_BUt	0.00945	0.000545	0.00839	0.0105
EDEPMAt	-0.0546	0.00605	-0.0664	-0.0427
EDEPMAt2	4.19E-10	8.13E-13	4.17E-10	4.2E-10
SMAt	0.0138	0.000663	0.0125	0.0151
I_MAt	-0.0111	0.00183	-0.0147	-0.00748
I_MAt2	1.29E-11	1.07E-13	1.27E-11	1.31E-11
EDEPBUt	-0.3063	0.0245	-0.3543	-0.2583
EDEPBUt2	3.2E-9	8.01E-11	3.043E-9	3.357E-9
dcat	0.0245	0.000241	0.0240	0.0250
dcat2	-128E-15	1.24E-15	-13E-14	-125E-15
dofat	0.0414	0.000184	0.0411	0.0418
OFAt_1	0.0418	0.000182	0.0414	0.0421
CAt_1	0.0251	0.000236	0.0246	0.0256
MAt_1	-0.0134	0.000650	-0.0147	-0.0121
BUt_1	0.00916	0.000531	0.00812	0.0102
realr	1122468	2535530	-3847257	6092193
FAAB	110718	42594.5	27231.4	194205
Public	992812	176694	646487	1339137
ruralare	-26553.7	65888.2	-155697	102589
largcity	-14437.0	44902.6	-102448	73573.7
market	-4225702	9204425	-2.227E7	13815285
marketw	-2189313	8646646	-1.914E7	14758409

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	5940	8.56	5940	8.56
1	63420	91.44	69360	100.00

Table 13c: Association of Predicted and Observed FI

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
FIt	6986190.74	553704519650	148867201	0	21656674938
pFIt	4236290.02	335755637941	179752661	-2405820765	45814748433

Table 14: Estimation results for *FE*

Table 14a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	69405
2	0	9852

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	59510.458	57361.749
SC	59519.738	57603.041
-2LogL	59508.458	57309.749

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	2198.7082	25	<.0001
Score	933.2619	25	<.0001
Wald	1244.2202	25	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	0.7403	0.0271	746.6314	<.0001
I_BUt	1.516E-9	3.18E-10	22.7028	<.0001
EDEPMAt	2.567E-8	3.35E-9	58.7356	<.0001
SMAt	-769E-13	4.49E-10	0.0293	0.8641
I_MAt	7.18E-10	8.47E-10	0.7174	0.3970
EDEPBUt	1.508E-8	1.113E-8	1.8365	0.1754
OFAt_1	-159E-12	2.5E-11	40.6404	<.0001
MAt_1	-153E-13	4.47E-10	0.0012	0.9728
BUt_1	1.946E-9	3.18E-10	37.3972	<.0001
LLt_1	5.71E-10	8.72E-11	42.9018	<.0001
sumcaclt_1	2.68E-10	3.96E-11	45.9394	<.0001
diffcaclt_1	-26E-11	4.81E-11	29.2427	<.0001
sumdcadclt	1.51E-10	3.44E-11	19.2813	<.0001
diffdcadclt	-175E-12	4.45E-11	15.5218	<.0001
sumdofadllt	1.7E-10	3.76E-11	20.5510	<.0001
diffdofadllt	-261E-12	4.53E-11	33.1787	<.0001
realr	-1.9037	0.5672	11.2649	0.0008
FAAB	0.1900	0.00967	386.0749	<.0001

Public	0.4724	0.0466	102.5980	<.0001
ruralare	0.0953	0.0154	38.3988	<.0001
largcity	-0.1258	0.0100	156.8897	<.0001
market	12.2057	11.6345	1.1006	0.2941
marketw	0.8544	3.7710	0.0513	0.8208

To be able to interpret the coefficient in front of the sums and diffs, we transform these variables to

CAt_1*(2.68E-10-268E-11)) = CAt_1*(2.412E-9)
CLt_1*(2.68E-10+268E-11) = CLt_1*(2.948E-9)

dCAt_1*(1.51E-10-175E-12) =dCAt_1*(-2.4E-11)
dCLt_1*(1.51E-10+175E-12) =dCLt_1*(3.26E-10)

dOFAt_1*(1.7E-10-261E-12) = dOFAt_1*(-9.1E-11)
dLLt_1*(1.7E-10+261E-12) = dLLt_1*(4.31E-10)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.9	Somers'D	0.311
Percent Discordant	32.8	Gamma	0.322
Percent Tied	3.4	Tau-a	0.068
Pairs	683778060	c	0.655

FEt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	9837	12.41	9837	12.41
positive	69420	87.59	79257	100.00

pFEt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	9837	12.41	9837	12.41
positive	69420	87.59	79257	100.00

Table 14b: Estimating the level of positive FE

Influence diagnostics (observations with positive FE***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	68822	99.16	68822	99.16
1	583	0.84	69405	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	67234	96.87	67234	96.87
1	2171	3.13	69405	100.00

Normality tests (observations with positive *FE* ***)***

Moments

N	69405	Sum Weights	69405
Mean	0	Sum Observations	0
Std Deviation	42713783.9	Variance	1.82447E15
Skewness	46.4426449	Kurtosis	4138.05629
Uncorrected SS	1.26625E20	Corrected SS	1.26625E20
Coeff Variation	.	Std Error Mean	162133.466

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.419804	Pr> D	<0.0100
Cramer-von Mises	W-Sq	4481.337	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	21127.32	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *FE* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	23	2.809E18	1.221E17	20902.2	<.0001
Residual	69382	4.144E17	5.973E12		
Uncorrected Total	69405	3.223E18			
Corrected Total	69404	3.161E18			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-182379	56031.5	-292203	-72555.5
I_BUt	0.0217	0.000320	0.0211	0.0224
EDEPMAt	0.0839	0.00261	0.0788	0.0890
SMAt	-0.00676	0.000403	-0.00755	-0.00597
I_MAt	-0.0210	0.000997	-0.0229	-0.0190
EDEPBUt	0.1914	0.00985	0.1721	0.2107
OFAt_1	0.0145	0.000120	0.0142	0.0147
MAt_1	0.00653	0.000395	0.00575	0.00730
BUt_1	0.0212	0.000309	0.0206	0.0218
LLt_1	0.0349	0.000197	0.0345	0.0353
sumcaclt_1	0.00815	0.000073	0.00801	0.00829
diffcaclt_1	-0.0183	0.000185	-0.0186	-0.0179
sumdcadclt	0.00795	0.000072	0.00781	0.00809
diffdcadclt	-0.0176	0.000179	-0.0180	-0.0173
sumdofadllt	0.0244	0.000078	0.0243	0.0246
diffdofadllt	-0.0102	0.000144	-0.0105	-0.00990
realr	4025817	1178460	1715994	6335639
FAAB	33749.7	19724.7	-4911.5	72410.9
Public	534338	85644.0	366473	702203
ruralare	3376.6	30016.7	-55457.2	62210.4
largcity	31715.9	20912.2	-9272.7	72704.6
market	-2122201	6154293	-1.418E7	9940424
marketw	-1.027E7	5694985	-2.143E7	896216

To be able to interpret the coefficient in front of sums and diffs, we transform these variables to

CAt_1*(0.00815-0.0183) = CAt_1*(-0.01015)
CLt_1*(0.00815+0.0183) = CLt_1*(0.02645)

dCAt_1*(0.00795-0.0176) = dCAt_1*(-9.65E-3)
dCLt_1*(0.00795+0.0176) = dCLt_1*(0.02555)

dOFAt_1*(0.0244-0.0102) = dOFAt_1*(0.0142)
dLLt_1*(0.0244+0.0102) = dLLt_1*(0.0346)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	7468	10.76	7468	10.76
1	61937	89.24	69405	100.00

Table 14c: Association of Predicted and Observed *FE*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
FEt	4385541.59	347584870063	55500816.84	0	4899875040
pFEt	3754349.94	297558513478	186822081	-212172403	51351615348

Table 15: Estimation results for $TDEP^{MA}$

Table 15a: Tobit Regression

Analysis of Parameter Estimates

Variable	*Estimate*	*Std. Error*	*Marginal effects*	*Chi-Square*	*Pr > ChiSq*
Intercept	-241944.6173	25491.379121		90.0835	<.0001
sumcasht_1	-0.000050009	0.000029176		2.9380	0.0865
diffcasht_1	-7.406171E-6	5.1781376E-6		2.0457	0.1526
SMAt	0.0000618203	0.0000129133	0.0000309082	22.9187	<.0001
EDEPMAt	0.9725199718	0.0021372003	0.9082877614	207064.949	<.0001
EDEPMAt2	-1.5758E-11	3.37353E-13	-7.78342E-12	2181.8995	<.0001
I_MAt	0.124253144	0.0015758228	0.0763689344	6217.2743	<.0001
I_MAt2	7.258052E-12	8.69759E-14	3.704029E-12	6963.7353	<.0001
ddmpat_1	0.0007151182	0.0007616983	0.0003575658	0.8814	0.3478
ddmpat_12	2.270743E-14	1.272161E-13	1.135377E-14	0.0319	0.8583
ddmpat_13	-5.98092E-23	8.355313E-23	-2.9905E-23	0.5124	0.4741
realr	-307511.9312	534787.4293	-152395.3792	0.3306	0.5653
FAAB	122477.19932	8952.7603075	63373.930883	187.1529	<.0001
Public	-49683.53468	34632.258835	-24827.89291	2.0581	0.1514
ruralare	30897.946616	13878.648244	15486.76249	4.9564	0.0260
largcity	-62222.20411	9493.1729025	-30540.56719	42.9604	<.0001
market	-8384350.076	2806072.3429	-4189784.906	8.9277	0.0028
marketw	-2496132.123	1764878.3452	-1247554.415	2.0003	0.1573
Scale	794649.34994	3085.7414692			

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.000050009-7.406171E-6) = ddmcasht_1*(-5.742E-5)
dcasht_1*(-0.000050009+7.406171E-6) = dcasht_1*(-4.26E-5)

The marginal effects of ddmcasht_1 and dcasht_1 are -0.000028703 and -0.0000213 respectively.

Association of Predicted Probabilities and Observed Responses

$TDEP^{MA}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	16013	20.20	16013	20.20
positive	63244	79.80	79257	100.00

$pTDEP^{MA}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	36957	46.63	36957	46.63
positive	42300	53.37	79257	100.00

Table 15b: Association of Predicted and Observed $TDEP^{MA}$

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
$TDEP^{MA}$	2607266.98	206644158878	45776744.55	0	6294501810
$pTDEP^{MA}$	2180559.54	172824607488	17205036.19	0	556254545

Table 16: Estimation results for *ZPF*

Table 16a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	6917
2	0	72340

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	46951.126	46334.055
SC	46960.406	46473.262
-2LogL	46949.126	46304.055

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	645.0707	14	<.0001
Score	629.7685	14	<.0001
Wald	609.4830	14	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.9046	0.0753	640.3104	<.0001
sumcasht_1	-41E-12	8.15E-11	0.2524	0.6154
diffcasht_1	-561E-15	1.21E-11	0.0022	0.9629
PALLOt_1	1.567E-9	1.452E-9	1.1639	0.2807
ddmpat_1	4.254E-9	7.05E-9	0.3642	0.5462
ddmpat_12	-807E-19	7.96E-17	1.0287	0.3105
ddmpat_13	-225E-28	2.21E-26	1.0302	0.3101
DTDEPMA	0.3745	0.0345	118.1515	<.0001
realr	-22.4922	1.4968	225.8144	<.0001
FAAB	0.3961	0.0248	254.9833	<.0001
Public	-0.2431	0.1019	5.6916	0.0170
ruralare	0.0698	0.0375	3.4593	0.0629
largcity	0.0470	0.0263	3.1936	0.0739
market	-84.4650	41.0100	4.2420	0.0394
marketw	0.3594	3.5953	0.0100	0.9204

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-41E-12-561E-15) = ddmcasht_1*(-4.156E-11)
dcasht_1*(-41E-12+561E-15) = dcasht_1*(-4.0439E-11)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	56.4	Somers'D	0.176
Percent Discordant	38.8	Gamma	0.184
Percent Tied	4.8	Tau-a	0.028
Pairs	500375780	c	0.588

ZPFt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	72340	91.27	72340	91.27
positive	6917	8.73	79257	100.00

pZPFt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	72356	91.29	72356	91.29
positive	6901	8.71	79257	100.00

Table 16b: Estimating the level of positive *ZPF*

Influence diagnostics (observations with positive ZPF)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	6909	99.88	6909	99.88
1	8	0.12	6917	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	6772	97.90	6772	97.90
1	145	2.10	6917	100.00

Normality tests (observations with positive ZPF)

Moments

N	6917	SumWeights	6917
Mean	0	SumObservations	0
Std Deviation	55276414.3	Variance	3.05548E15
Skewness	80.1630196	Kurtosis	6596.80288
Uncorrected SS	2.11317E19	CorrectedSS	2.11317E19
Coeff Variation	.	StdErrorMean	664631.584

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.44944	Pr> D	<0.0100

Cramer-von Mises	W-Sq	520.3203	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	2439.726	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** ZPF **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	13	1.361E15	1.047E14	30.78	<.0001
Residual	6904	1.621E16	2.348E12		
Uncorrected Total	6917	1.757E16			
Corrected Total	6916	1.708E16			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	417800	115299	191775	643824
sumcasht_1	0.000464	0.000543	-0.00060	0.00153
diffcasht_1	-0.00011	0.000043	-0.00019	-0.00002
PALLOt_1	0.1053	0.00909	0.0875	0.1231
ddmpat_1	-0.0106	0.0125	-0.0351	0.0140
DTDEPMA	9917.2	53390.4	-94746.3	114581
realr	-3668399	2236211	-8052140	715343
FAAB	-168141	38626.3	-243862	-92420.5
Public	934196	176655	587891	1280501
ruralare	-40800.4	58108.8	-154714	73112.7
largcity	91246.9	40916.7	11036.1	171458
market	1.5706E8	67146208	25433879	2.8869E8
marketw	1.6632E8	20684102	1.2577E8	2.0687E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(0.000464-0.00011) = ddmcasht_1*(3.54E-4)
dcasht_1*(0.000464+0.00011) = dcasht_1*(5.74E-4)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	401	5.80	401	5.80
1	6516	94.20	6917	100.00

Table 16c: Association of Predicted and Observed *ZPF*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
ZPFt	244860.44	19406904251	29191276.08	0	5224770000
pZPFt	41092.92	3256901193	2600952.31	-761564.16	536305090

Table 17: Estimation results for *dour*

Table 17a: Multinomial Logit

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	Positive	5143
2	Zero	58728
3	Negative	15386

Analysis Of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Wald 95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept1	-2.1685	0.02953	-2.2264	-2.1106	5390.92	0
Intercept2	1.06147	0.02692	1.00870	1.11424	1554.29	0
sumcasht_1	-46E-12	321E-13	-11E-11	166E-13	2.07771	0.14946
diffcasht_1	-15E-12	571E-14	-26E-12	-41E-13	7.14213	0.00753
ddmpat_1	1.11E-9	698E-12	-26E-11	2.48E-9	2.51794	0.11256
ddmpat_12	309E-24	27E-20	-53E-20	53E-20	1.31E-6	0.99909
ddmpat_13	-6E-29	117E-30	-29E-29	17E-29	0.25832	0.61128
DTDEPMA	-0.3132	0.01118	-0.3351	-0.2913	785.200	89E-174
DZPF	-0.3576	0.01595	-0.3889	-0.3264	502.617	26E-112
realr	-4.8872	0.52906	-5.9242	-3.8503	85.3313	252E-22
FAAB	-0.1380	0.00903	-0.1557	-0.1203	233.631	963E-55
Public	0.07302	0.03349	0.00739	0.13866	4.75534	0.02921
ruralare	-0.0182	0.01419	-0.0460	0.00961	1.64506	0.19963
largcity	0.03468	0.00953	0.01601	0.05335	13.2538	0.00027
market	1.33087	2.14806	-2.8793	5.54100	0.38386	0.53554
marketw	-10.725	2.04754	-14.738	-6.7118	27.4362	1.62E-7

NOTE: The scale parameter was held fixed.

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-46E-12-15E-12) = ddmcasht_1*(-6.1E-11)
dcasht_1*(-46E-12+15E-12) = dcasht_1*(-3.1E-11)

Table 17b: Estimating the level of positive *dour*

Influence diagnostics (observations with positive *dour* ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	5131	99.77	5131	99.77
1	12	0.23	5143	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*

0	4754	92.44	4754	92.44
1	389	7.56	5143	100.00

Normality tests (observations with positive** dour **)

Moments

N	5143	Sum Weights	5143
Mean	0	Sum Observations	0
Std Deviation	107235031	Variance	1.14994E16
Skewness	61.4395121	Kurtosis	4107.45143
Uncorrected SS	5.91297E19	Corrected SS	5.91297E19
Coeff Variation	.	Std Error Mean	1495300.3

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.433029	Pr> D	<0.0100
Cramer-von Mises	W-Sq	352.1443	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	1671.197	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** dour **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	14	9.686E15	6.919E14	10.25	<.0001
Residual	5129	1.629E17	3.176E13		
Uncorrected Total	5143	1.726E17			
Corrected Total	5142	1.671E17			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	865762	1060376	-1213065	2944589
sumcasht_1	0.000251	0.000378	-0.00049	0.000992
diffcasht_1	-0.00006	0.000057	-0.00018	0.000048
ddmpat_1	0.0146	0.0224	-0.0292	0.0584
ddmpat_12	3.91E-11	1.28E-10	-212E-12	2.9E-10
DTDEPMA	461104	318147	-162612	1084820
DZPF	63314.1	410126	-740722	867350
realr	3675150	21757054	-3.898E7	46329029
FAAB	-1186295	162022	-1503933	-868657
Public	4038992	576978	2907848	5170136
ruralare	-53932.7	241757	-527888	420023
largcity	194704	173949	-146317	535724
market	4.3597E8	2.0779E8	28598316	8.4333E8
marketw	26167997	15111923	-3458355	55794349

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(0.000251-0.00006) = ddmcasht_1*(1.91E-4)
dcasht_1*(0.000251+0.00006) = dcasht_1*(3.11E-4)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	208	4.04	208	4.04
1	4935	95.96	5143	100.00

Table 17c: Estimating the level of negative *dour*

Influence diagnostics (observations with negative *dour* ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	15253	99.14	15253	99.14
1	133	0.86	15386	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	15000	97.49	15000	97.49
1	386	2.51	15386	100.00

Normality tests (observations with negative *dour* ***)***

Moments

N	15386	Sum Weights	15386
Mean	0	Sum Observations	0
Std Deviation	20042026.3	Variance	4.01683E14
Skewness	-27.417342	Kurtosis	1604.97596
Uncorrected SS	6.17989E18	Corrected SS	6.17989E18
Coeff Variation	.	Std Error Mean	161576.711

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.426812	Pr> D	<0.0100
Cramer-von Mises	W-Sq	984.955	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	4666.55	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative** dour **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	13	5.618E16	4.321E15	2823.77	<.0001
Residual	15373	2.468E16	1.605E12		
Uncorrected Total	15386	8.086E16			
Corrected Total	15385	7.907E16			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-364671	70543.9	-502948	-226394
sumcasht_1	-0.00702	0.000262	-0.00753	-0.00650
diffcasht_1	-0.00179	0.000045	-0.00188	-0.00170
ddmpat_1	-0.00040	0.00101	-0.00237	0.00158
DTDEPMA	5239.5	35807.9	-64949.6	75428.6
DZPF	49005.8	30856.8	-11478.2	109490
realr	945978	1278703	-1560478	3452433
FAAB	128361	21677.5	85870.2	170853
Public	-315251	93143.8	-497827	-132675
ruralare	7924.2	32247.8	-55286.4	71134.8
largcity	-28112.8	23067.6	-73328.9	17103.2
market	1.689E8	36157322	98024329	2.3977E8
marketw	-4.151E8	7005189	-4.289E8	-4.014E8

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.00702-0.00179) = ddmcasht_1*(-8.81E-3)
dcasht_1*(-0.00702+0.00179) = dcasht_1*(-5.23E-3)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	1742	11.32	1742	11.32
1	13644	88.68	15386	100.00

Table 17d: **Association of Predicted and Observed** *dour*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
dourt	-56919.11	-4511238208	32177604.04	-4079179385	7282713000
pdourt	-46371.94	-3675300733	3862075.22	-820076286	329076094

Table 18: Estimation results for GC

Table 18a: Logistic model with complementary log-log function (Gompertz) (observations with positive GC)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	6948
2	0	72309

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	47096.512	45743.589
SC	47105.793	45919.917
-2LogL	47094.512	45705.589

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	1388.9237	18	<.0001
Score	2014.3317	18	<.0001
Wald	1406.4084	18	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-2.3275	0.0888	687.1010	<.0001
OIBDt	-6.56E-9	4.98E-10	172.9532	<.0001
OIBDt2	7.36E-19	2.75E-19	7.1332	0.0076
OIBDt3	7.65E-30	3.91E-29	0.0382	0.8450
FIt	3.08E-11	8.32E-11	0.1367	0.7116
FEt	7.75E-10	2.13E-10	13.2240	0.0003
TDEPMAt	9.147E-9	6.87E-10	177.2565	<.0001
TDEPMAt2	-122E-20	1.01E-19	146.0587	<.0001
EDEPBUt	3.208E-8	5.32E-9	36.3699	<.0001
EDEPBUt2	-945E-19	3.69E-17	6.5791	0.0103
ZPFt	-2.22E-8	5.658E-9	15.3672	<.0001
dourt	7.09E-11	2.76E-10	0.0662	0.7969
dgnp	6.5E-13	1.26E-12	0.2683	0.6045
FAAB	-0.4256	0.0258	271.9200	<.0001
Public	1.2128	0.0540	503.8085	<.0001
ruralare	-0.0887	0.0418	4.5119	0.0337
largcity	0.0532	0.0261	4.1335	0.0420

market	-10.2263	21.8315	0.2194	0.6395
marketw	3.6123	1.6850	4.5957	0.0321

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.4	Somers'D	0.276
Percent Discordant	32.8	Gamma	0.297
Percent Tied	6.8	Tau-a	0.044
Pairs	502402932	c	0.638

Table 18b: Logistic model with complementary log-log function (Gompertz) (observations with negative and zero GC)

Number of Response Levels:	2
Number of Observations:	72309
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	8356
2	0	63953

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	51772.852	47051.501
SC	51782.040	47216.897
-2LogL	51770.852	47015.501

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	4755.3509	17	<.0001
Score	6127.3778	17	<.0001
Wald	3615.0100	17	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-1.2440	0.0802	240.6491	<.0001
OIBDt	9.291E-9	4.27E-10	474.2146	<.0001
OIBDt2	-478E-20	3.23E-19	219.3506	<.0001
FIt	1.22E-10	1.1E-10	1.2268	0.2680
FEt	3.27E-10	2.22E-10	2.1604	0.1416
TDEPMAt	-2.27E-9	9.92E-10	5.2448	0.0220
TDEPMAt2	1.57E-18	1.11E-18	1.9809	0.1593
EDEPBU	-2.22E-8	4.215E-9	27.6656	<.0001

EDEPBUt2	5.72E-17	2.62E-17	4.7518	0.0293
ZPFt	6.57E-12	1.85E-10	0.0013	0.9717
dourt	-912E-13	4.12E-10	0.0491	0.8246
dgnp	-604E-14	1.14E-12	27.8405	<.0001
FAAB	-1.3687	0.0287	2271.2765	<.0001
Public	0.5467	0.0714	58.6406	<.0001
ruralare	-0.1423	0.0382	13.8936	0.0002
largcity	-0.1418	0.0237	35.7945	<.0001
market	-1.7044	1.9761	0.7439	0.3884
marketw	9.3471	1.9670	22.5810	<.0001

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.0	Somers'D	0.507
Percent Discordant	23.4	Gamma	0.520
Percent Tied	2.6	Tau-a	0.104
Pairs	534391268	c	0.753

GCt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	8356	10.54	8356	10.54
zero	63953	80.69	72309	91.23
positive	6948	8.77	79257	100.00

pGCt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
negative	8425	10.63	8425	10.63
zero	63897	80.62	72322	91.25
positive	6935	8.75	79257	100.00

Table 18c: Estimating the level of positive GC

Influence diagnostics (observations with positive GC)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	6844	98.50	6844	98.50
1	104	1.50	6948	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	6775	97.51	6775	97.51
1	173	2.49	6948	100.00

Normality tests (observations with positive *GC* ***)***

Moments

N	6948	Sum Weights	6948
Mean	0	Sum Observations	0
Std Deviation	119632404	Variance	1.43119E16
Skewness	15.1287295	Kurtosis	447.46677
Uncorrected SS	9.94249E19	Corrected SS	9.94249E19
Coeff Variation	.	Std Error Mean	1435221.48

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.363569	Pr> D	<0.0100
Cramer-von Mises	W-Sq	353.7129	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	1709.712	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *GC* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	19	3.137E19	1.651E18	6514.17	<.0001
Residual	6929	1.825E18	2.634E14		
Uncorrected Total	Total	6948	3.319E19		
Corrected Total	Total	6947	3.271E19		

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	2386081	1434699	-426419	5198581
OIBDt	-0.4343	0.00953	-0.4530	-0.4156
OIBDt2	2.34E-10	2.52E-12	2.29E-10	2.39E-10
OIBDt3	-292E-22	1.31E-21	-318E-22	-267E-22
FIt	0.0252	0.00242	0.0204	0.0299
FEt	0.2931	0.00741	0.2786	0.3076
TDEPMAt	0.5699	0.0140	0.5424	0.5974
TDEPMAt2	-113E-12	2.16E-12	-117E-12	-108E-12
EDEPBUt	1.4073	0.1020	1.2073	1.6072
EDEPBUt2	-1.39E-8	3.97E-10	-1.47E-8	-1.31E-8
ZPFt	-0.0715	0.1570	-0.3792	0.2363
dourt	0.00924	0.0149	-0.0199	0.0384
dgnp	0.000015	0.000021	-0.00003	0.000055
FAAB	-728042	432317	-1575532	119448
Public	5696818	955587	3823541	7570095
ruralare	-441121	691727	-1797142	914900
largcity	886165	439085	25408.4	1746923
market	1.4414E8	4.6143E8	-7.604E8	1.0487E9
marketw	-7.027E7	48127416	-1.646E8	24074573

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	976	14.05	976	14.05
1	5972	85.95	6948	100.00

Table 18d: Estimating the level of negative GC

Influence diagnostics (observations with negative GC ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	8192	98.04	8192	98.04
1	164	1.96	8356	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	7915	94.72	7915	94.72
1	441	5.28	8356	100.00

Normality tests (observations with negative GC ***)***

Moments

N	8356	Sum Weights	8356
Mean	0	Sum Observations	0
Std Deviation	44720799.4	Variance	1.99995E15
Skewness	-6.6058086	Kurtosis	440.033355
Uncorrected SS	1.67096E19	Corrected SS	1.67096E19
Coeff Variation	.	Std Error Mean	489226.916

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.325112	Pr> D	<0.0100
Cramer-von Mises	W-Sq	381.522	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	1845.003	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative GC ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	18	7.795E18	4.331E17	8131.80	<.0001
Residual	8338	4.408E17	5.286E13		
Uncorrected Total	8356	8.236E18			
Corrected Total	8355	7.748E18			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	931181	602967	-250805	2113167
OIBDt	-0.8272	0.00448	-0.8360	-0.8185
OIBDt2	-465E-15	1.23E-12	-288E-14	1.95E-12
FIt	-0.1481	0.00337	-0.1547	-0.1415
FEt	0.1743	0.00711	0.1604	0.1882
TDEPMAt	0.7714	0.00941	0.7529	0.7898
TDEPMAt2	4.03E-11	4.44E-12	3.16E-11	4.91E-11
EDEPBUt	2.0254	0.0557	1.9162	2.1346
EDEPBUt2	4.176E-9	3.58E-10	3.473E-9	4.878E-9
ZPFt	0.00308	0.00126	0.000612	0.00554
dourt	-0.00008	0.00377	-0.00747	0.00730
dgnp	-0.00001	8.745E-6	-0.00003	3.963E-6
FAAB	843477	223967	404439	1282515
Public	-2780722	640667	-4036611	-1524834
ruralare	9559.5	292804	-564419	583538
largcity	-553991	183095	-912910	-195073
market	-1.203E8	38649632	-1.96E8	-4.45E7
marketw	8873028	19936058	-3.021E7	47953368

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	2171	25.98	2171	25.98
1	6185	74.02	8356	100.00

Table 18e: **Association of Predicted and Observed** *GC*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
GCt	219327.03	17383202532	59654357.81	-3525685159	6504432000
pGCt	-123440.77	-9783545452	43737442.14	-1877356792	4572186841

Table 19: Estimation results for *OA*

Table 19a: Multinomial Logit

Response Profile

Ordered Value	Prob	Total Frequency
1	Positive	21431
2	Zero	49465
3	Negative	8361

Analysis Of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Wald 95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept1	-1.2269	0.02676	-1.2793	-1.1744	2101.78	0
Intercept2	0.80162	0.02625	0.75017	0.85307	932.602	8E-205
dourt	-54E-11	252E-12	-1E-9	-45E-12	4.57336	0.03247
GCt	1.64E-9	357E-13	1.57E-9	1.71E-9	2097.37	0
DTDEPMA	-0.1227	0.01066	-0.1436	-0.1018	132.461	119E-32
DZPF	0.08816	0.01512	0.05854	0.11779	34.0173	5.46E-9
realr	-1.6564	0.51864	-2.6729	-0.6399	10.2001	0.00140
FAAB	0.44790	0.00883	0.43059	0.46520	2574.19	0
Public	-0.0229	0.03218	-0.0860	0.04017	0.50646	0.47668
ruralare	0.03657	0.01374	0.00965	0.06349	7.08704	0.00776
largcity	0.01973	0.00918	0.00173	0.03772	4.61786	0.03164
market	0.29826	1.84284	-3.3136	3.91016	0.02620	0.87142
marketw	-3.6165	1.47518	-6.5078	-0.7252	6.01000	0.01423

NOTE: The scale parameter was held fixed.

Table 19b: Estimating the level of positive *OA*

Influence diagnostics (observations with positive OA)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	21311	99.44	21311	99.44
1	120	0.56	21431	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	20736	96.76	20736	96.76
1	695	3.24	21431	100.00

Normality tests (observations with positive *OA****)***

Moments

N	21431	Sum Weights	21431
Mean	0	Sum Observations	0
Std Deviation	56469117	Variance	3.18876E15
Skewness	1.61986239	Kurtosis	3820.96086
Uncorrected SS	6.83352E19	Corrected SS	6.83352E19
Coeff Variation	.	Std Error Mean	385735.65

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.411253	Pr> D	<0.0100
Cramer-von Mises	W-Sq	1370.461	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	6518.492	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *OA****)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	12	4.507E19	3.756E18	269134	<.0001
Residual	21419	3.224E17	1.505E13		
Uncorrected Total	Total	21431	4.539E19		
Corrected Total	Total	21430	4.488E19		

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	739683	176450	393821	1085546
dourt	0.00759	0.00165	0.00436	0.0108
GCt	0.9970	0.000591	0.9958	0.9981
DTDEPMA	41894.5	71497.9	-98249.3	182038
DZPF	-67079.7	87736.9	-239054	104894
realr	-2140373	3415546	-8835221	4554474
FAAB	-543333	55238.6	-651607	-435059
Public	380865	185482	17299.1	744431
ruralare	12606.5	84778.4	-153568	178781
largcity	153402	58745.9	38253.8	268551
market	64384715	71802596	-7.636E7	2.0513E8
marketw	38827730	7164785	24783958	52871502

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	1176	5.49	1176	5.49
1	20255	94.51	21431	100.00

Table 19c: Estimating the level of negative *OA*

Influence diagnostics (observations with negative OA)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	8288	99.13	8288	99.13
1	73	0.87	8361	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	7941	94.98	7941	94.98
1	420	5.02	8361	100.00

Normality tests (observations with negative OA)

Moments

N	8361	Sum Weights	8361
Mean	0	Sum Observations	0
Std Deviation	33213221	Variance	1.10312E15
Skewness	-25.576124	Kurtosis	1072.17196
Uncorrected SS	9.22207E18	Corrected SS	9.22207E18
Coeff Variation	.	Std Error Mean	363230.145

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.402025	Pr> D	<0.0100
Cramer-von Mises	W-Sq	494.0111	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	2366.3	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative OA)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	12	1.553E19	1.294E18	148243	<.0001
Residual	8349	7.449E16	8.922E12		
Uncorrected Total	8361	1.56E19			
Corrected Total	8360	1.462E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-241958	225182	-683379	199463
dourt	-0.00093	0.000397	-0.00171	-0.00015
GCt	0.9907	0.000799	0.9892	0.9923
DTDEPMA	248546	97542.6	57334.8	439757
DZPF	67000.4	128273	-184451	318452
realr	1052639	4304127	-7384672	9489950
FAAB	-73380.5	84118.8	-238277	91516.2
Public	-843524	217175	-1269249	-417800
ruralare	37518.7	114097	-186144	261182
largcity	-143196	71995.0	-284327	-2065.2

market	2190197	68766262	-1.326E8	1.3699E8
marketw	-5321761	7604676	-2.023E7	9585563

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	758	9.07	758	9.07
1	7603	90.93	8361	100.00

Table 19d: **Association of Predicted and Observed** *OA*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
OAt	512504.50	40619569009	54179761.73	-3525685000	4928903000
pOAt	499497.95	39588709026	56171053.50	-3494626993	6494558514

Table 20: Estimation results for *TL*

Table 20a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	50072
2	0	29185

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	104305.52	101567.16
SC	104314.80	101771.33
-2LogL	104303.52	101523.16

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	2780.3640	21	<.0001
Score	2356.2655	21	<.0001
Wald	2318.4964	21	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-0.2181	0.0349	39.0556	<.0001
OIBDt	2.906E-9	2.49E-10	136.2547	<.0001
OIBDt2	-857E-21	9.14E-20	87.8694	<.0001
FIt	3.55E-10	1.15E-10	9.5357	0.0020
FIt2	-323E-22	1.92E-20	2.8451	0.0917
FEt	-2.25E-9	2.96E-10	57.8206	<.0001
FEt2	4.97E-19	8.31E-20	35.8427	<.0001
TDEPMAt	-529E-12	3.99E-10	1.7634	0.1842
TDEPMAt2	4.83E-19	1.52E-19	10.1483	0.0014
EDEPBUt	-1.52E-8	3.191E-9	22.7443	<.0001
EDEPBUt2	9.04E-17	2.23E-17	16.3917	<.0001
dourt	-531E-12	3.14E-10	2.8518	0.0913
dourt2	5.44E-20	4.99E-20	1.1882	0.2757
ZPFt	4.24E-10	4.55E-10	0.8693	0.3511
PALLOt_1	1.918E-8	2.254E-9	72.3795	<.0001
dgnp	8.06E-13	4.92E-13	2.6854	0.1013
FAAB	0.4153	0.00962	1864.3708	<.0001
Public	-0.1320	0.0389	11.5290	0.0007

ruralare	-0.0277	0.0150	3.4301	0.0640
largcity	-0.0603	0.0102	35.0336	<.0001
market	1.2303	2.2983	0.2866	0.5924
marketw	-4.4881	1.9980	5.0456	0.0247

Association of Predicted Probabilities and Observed Responses

Percent Concordant	61.7	Somers'D	0.265
Percent Discordant	35.2	Gamma	0.273
Percent Tied	3.1	Tau-a	0.123
Pairs	1461351320	c	0.632

TLt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	29185	36.82	29185	36.82
positive	50072	63.18	79257	100.00

pTLt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	29212	36.86	29212	36.86
positive	50045	63.14	79257	100.00

Table 20b: Estimating the level of positive TL

Influence diagnostics (observations with positive TL ***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	49670	99.20	49670	99.20
1	402	0.80	50072	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	48782	97.42	48782	97.42
1	1290	2.58	50072	100.00

Normality tests (observations with positive TL ***)***

Moments

N	50072	Sum Weights	50072
Mean	0	Sum Observations	0
Std Deviation	13180242.3	Variance	1.73719E14
Skewness	25.3392452	Kurtosis	1649.96601
Uncorrected SS	8.69827E18	CorrectedSS	8.69827E18
Coeff Variation	.	StdErrorMean	58901.4418

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.409425	Pr> D	<0.0100
Cramer-von Mises	W-Sq	3285.739	Pr> W-Sq	<0.0050

Anderson-Darling	A-Sq	15514.31	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** TL **)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	22	8.983E17	4.083E16	61499.3	<.0001
Residual	50050	3.461E16	6.915E11		
Uncorrected Total	50072	9.329E17			
Corrected Total	50071	9.277E17			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-87622.7	29926.4	-146280	-28965.4
OIBDt	0.0909	0.000549	0.0899	0.0920
OIBDt2	-116E-13	1.62E-13	-119E-13	-113E-13
FIt	0.0244	0.000451	0.0235	0.0253
FIt2	4.71E-12	1.34E-13	4.45E-12	4.98E-12
FEt	-0.0239	0.000770	-0.0254	-0.0223
FEt2	2.64E-12	1.72E-13	2.31E-12	2.98E-12
TDEPMAt	-0.0625	0.000891	-0.0643	-0.0608
TDEPMAt2	1.25E-11	2.87E-13	1.19E-11	1.31E-11
EDEPBUt	-0.2130	0.00476	-0.2224	-0.2037
EDEPBUt2	9.57E-10	2.98E-11	8.98E-10	1.015E-9
dourt	0.0482	0.000696	0.0469	0.0496
dourt2	3.9E-11	6.52E-13	3.77E-11	4.03E-11
ZPFt	-0.3147	0.00254	-0.3197	-0.3097
PALLOt_1	0.3773	0.00302	0.3714	0.3833
dgnp	2.253E-6	4.18E-7	1.434E-6	3.073E-6
FAAB	-11317.6	8182.6	-27355.8	4720.6
Public	726252	46570.8	634971	817533
ruralare	-23180.7	12333.5	-47355.0	993.6
largcity	42745.8	8602.1	25885.3	59606.3
market	83539967	1707803	80192592	86887343
marketw	-1499138	2355516	-6116060	3117784

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	9747	19.47	9747	19.47
1	40325	80.53	50072	100.00

Table 20c: Association of Predicted and Observed *TL*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
TLt	858500.88	68042204487	13990142.60	-384369286	1639649673
pTLt	564695.84	44756098057	14201793.66	-1382131232	3003022239

Table 21: Estimation results for *OTA*

Table 21a: Multinomial Logit

Response Profile

Ordered Value	Prob	Total Frequency
1	Positive	61241
2	Zero	6763
3	Negative	11253

Analysis Of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Wald 95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept1	0.37943	0.03100	0.31867	0.44019	149.815	19E-35
Intercept2	0.66734	0.03101	0.60656	0.72812	463.126	1E-102
PALLOt_1	4.31E-9	55E-11	3.23E-9	5.38E-9	61.3445	479E-17
ZPFt	635E-13	435E-12	-79E-11	915E-12	0.02134	0.88385
TDEPMAt	-12E-10	495E-12	-22E-10	-28E-11	6.37391	0.01158
TDEPMAt2	21E-19	238E-21	164E-20	257E-20	78.0882	985E-21
OIBDt	-19E-10	239E-12	-24E-10	-14E-10	62.9300	214E-17
OIBDt2	-87E-20	183E-21	-12E-19	-51E-20	22.2813	2.35E-6
EDEPBUt	8.89E-9	2.64E-9	3.72E-9	1.41E-8	11.3572	0.00075
EDEPBUt2	842E-20	19E-18	-29E-18	457E-19	0.19583	0.65811
dourt	1.22E-9	442E-12	348E-12	2.08E-9	7.54482	0.00602
TLt	1.45E-8	1.06E-9	1.24E-8	1.66E-8	187.580	107E-44
FIt	-95E-10	402E-12	-1E-8	-87E-10	553.789	19E-123
FEt	3.1E-9	274E-12	2.56E-9	3.64E-9	127.955	115E-31
dgnp	-31E-14	438E-15	-12E-13	552E-15	0.48772	0.48495
FAAB	0.24556	0.00865	0.22861	0.26252	806.108	25E-178
Public	-0.2513	0.03584	-0.3215	-0.1810	49.1544	237E-14
ruralare	0.01285	0.01359	-0.0138	0.03949	0.89332	0.34458
largcity	-0.0928	0.00912	-0.1107	-0.0749	103.563	252E-26
market	-1.8691	1.64330	-5.0899	1.35171	1.29369	0.25537
marketw	-1.6171	1.61233	-4.7772	1.54299	1.00596	0.31587

NOTE: The scale parameter was held fixed.

Table 21b: Estimating the level of positive *OTA*

Influence diagnostics (observations with positive OTA)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	60831	99.33	60831	99.33
1	410	0.67	61241	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	59583	97.29	59583	97.29
1	1658	2.71	61241	100.00

Normality tests (observations with positive** OTA**)

Moments

N	61241	Sum Weights	61241
Mean	0	Sum Observations	0
Std Deviation	22184118	Variance	4.92135E14
Skewness	67.4179778	Kurtosis	9515.34523
Uncorrected SS	3.01384E19	Corrected SS	3.01384E19
Coeff Variation	.	Std Error Mean	89643.9594

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.421039	Pr> D	<0.0100
Cramer-von Mises	W-Sq	4058.861	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	19176.48	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive** OTA**)

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	20	7.337E17	3.668E16	20000.4	<.0001
Residual	61221	1.142E17	1.866E12		
Uncorrected Total	61241	8.479E17			
Corrected Total	61240	8.233E17			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-113233	42514.3	-196563	-29903.3
PALLOt_1	0.0103	0.00177	0.00682	0.0138
ZPFt	-0.00958	0.00152	-0.0126	-0.00661
TDEPMAt	0.0353	0.00131	0.0327	0.0378
TDEPMAt2	-548E-13	1.14E-12	-571E-13	-526E-13
OIBDt	-0.0114	0.000606	-0.0126	-0.0102
OIBDt2	2.48E-11	6.03E-13	2.36E-11	2.6E-11
EDEPBUt	0.2952	0.00701	0.2815	0.3090
EDEPBUt2	3.03E-10	6.5E-11	1.75E-10	4.3E-10
dourt	0.000416	0.000503	-0.00057	0.00140
TLt	0.9736	0.00244	0.9689	0.9784
FIt	-0.0530	0.00124	-0.0555	-0.0506
FEt	0.2117	0.00119	0.2093	0.2140
dgnp	2.384E-6	5.971E-7	1.213E-6	3.554E-6
FAAB	-29556.0	11700.7	-52489.9	-6622.2
Public	345844	57195.4	233739	457950
ruralare	-9898.4	17758.6	-44705.8	24909.1
largcity	54708.4	12352.7	30496.5	78920.2
market	1439326	13031448	-2.41E7	26981468
marketw	14462364	4056972	6510543	22414186

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	5273	8.61	5273	8.61
1	55968	91.39	61241	100.00

Table 21c: Estimating the level of negative OA

Influence diagnostics (observations with negative OTA***)***

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	11112	98.75	11112	98.75
1	141	1.25	11253	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	10986	97.63	10986	97.63
1	267	2.37	11253	100.00

Normality tests (observations with negative OTA***)***

Moments

N	11253	Sum Weights	11253
Mean	0	Sum Observations	0
Std Deviation	114968523	Variance	1.32178E16
Skewness	-24.332356	Kurtosis	1115.10237
Uncorrected SS	1.48726E20	Corrected SS	1.48726E20
Coeff Variation	.	Std Error Mean	1083789.14

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.37438	Pr> D	<0.0100
Cramer-von Mises	W-Sq	627.5316	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	3008.647	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with negative OTA***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	19	1.068E20	5.623E18	29899.7	<.0001
Residual	11234	2.207E18	1.964E14		
Uncorrected Total	11253	1.09E20			
Corrected Total	11252	1.079E20			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-1387497	1008947	-3365247	590253
PALLOt_1	0.5766	0.0432	0.4920	0.6612
ZPFt	-0.3916	0.1166	-0.6203	-0.1630

TDEPMAt	0.2403	0.00940	0.2219	0.2587
OIBDt	-0.2255	0.00531	-0.2359	-0.2151
OIBDt2	-189E-13	2.86E-12	-245E-13	-133E-13
EDEPBUt	-0.7811	0.0952	-0.9677	-0.5944
EDEPBUt2	9E-10	1.012E-9	-1.08E-9	2.885E-9
dourt	0.1481	0.0122	0.1242	0.1719
TLt	-0.00597	0.0317	-0.0682	0.0562
FIt	-0.9500	0.00142	-0.9528	-0.9472
FEt	0.7590	0.00508	0.7490	0.7689
dgnp	-5.63E-6	0.000014	-0.00003	0.000022
FAAB	1501982	289810	933895	2070070
Public	2318879	776770	796246	3841512
ruralare	-15825.2	471132	-939344	907694
largcity	-718388	299151	-1304788	-131988
market	-3.193E8	29756314	-3.776E8	-2.609E8
marketw	29604280	39678823	-4.817E7	1.0738E8

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	1424	12.65	1424	12.65
1	9829	87.35	11253	100.00

Table 21d: **Association of Predicted and Observed** *OTA*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
OTAt	-2853018.20	-2.261217E11	121126109	-18291166560	3350581783
pOTAt	-3173443.69	-2.515176E11	119290963	-17855387358	1596805744

Table 22: Estimation results for $TDEP^{BU}$

Table 22a: Tobit Regression

Analysis of Parameter Estimates

Parameter	*Estimate*	*Std. Error*	*Marginal effects*	*Chi- Square*	*Pr > ChiSq*
Intercept	-2867269.002	122295.23172		549.69	<.0001
sumcasht_1	-0.000088805	0.0001189346		0.56	0.4553
diffcasht_1	-0.000074819	0.0000194164		14.85	0.0001
EDEPMAt	0.0527071077	0.0023413409	0.0272304101	506.77	<.0001
EDEPMAt2	-7.25922E-11	2.50861E-12	-3.53558E-11	837.36	<.0001
SMAt	-0.000963799	0.0002832952	-0.000482122	11.57	0.0007
I_MAt	0.0024001815	0.0009957932	0.0012025995	5.81	0.0159
BUt_1	0.0267410219	0.0002057947	0.0147042398	16884.5	<.0001
BUt_12	-9.20641E-13	3.10627E-14	-4.55927E-13	878.42	<.0001
dcat	0.0001868159	0.0000694165	0.0000934492	7.24	0.0071
dcat2	-1.3814E-15	2.939616E-16	-6.90369E-16	22.08	<.0001
dclt	0.0000516105	0.0000649497	0.0000258083	0.63	0.4268
ddmpat_1	0.0071342988	0.0047361765	0.0035674599	2.27	0.1320
ddmpat_12	-1.25175E-11	4.315741E-12	-6.25127E-12	8.41	0.0037
ddmpat_13	-8.74809E-23	1.517684E-21	-4.37408E-23	0.00	0.9540
dgnp	4.2052645E-6	1.7025703E-6	2.2768591E-6	6.10	0.0135
FAAB	-225275.5688	33437.284115	-109287.755	45.39	<.0001
Public	11367.175096	129910.29097	5683.9243021	0.01	0.9303
ruralare	520913.48196	43443.086905	265436.46244	143.78	<.0001
largcity	-1303571.438	38815.082765	-536873.574	1127.89	<.0001
market	10895185.155	3719704.9111	5449463.9923	8.58	0.0034
marketw	-9820538.588	6787126.3039	-4906597.119	2.09	0.1479
Scale	1713791.0346	13883.484898			

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(-0.000088805-0.000074819) = ddmcasht_1*(-1.63624E-4)
dcasht_1*(-0.000088805+0.000074819) = dcasht_1*(-1.396E-5)

The marginal effects of ddmcasht_1 and dcasht_1 are -0.000081794 and -6.992912E-6 respectively.

Association of Predicted Probabilities and Observed Responses

$TDEP^{BU}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	65044	82.07	65044	82.07
positive	14213	17.93	79257	100.00

$pTDEP^{BU}$	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	66444	83.83	66444	83.83
positive	12813	16.17	79257	100.00

Table 22b: *Association of Predicted and Observed* $TDEP^{BU}$

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
$TDEP^{BU}$	398612.07	31592796445	8500272.16	0	1263065150
$pTDEP^{BU}$	269651.15	21371741299	3271974.57	0	194470087

Table 23: Estimation results for P^{allo}

Table 23a: Estimating the level of P^{allo}

Influence diagnostics (observations with positive P^{allo})

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	79020	99.70	79020	99.70
1	237	0.30	79257	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	77380	97.63	77380	97.63
1	1877	2.37	79257	100.00

Normality tests (observations with positive P^{allo})

Moments

N	79257	Sum Weights	79257
Mean	0	Sum Observations	0
Std Deviation	5222190.38	Variance	2.72713E13
Skewness	-89.761265	Kurtosis	13457.9345
Uncorrected SS	2.16141E18	Corrected SS	2.16141E18
Coeff Variation	.	Std Error Mean	18549.5717

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.437701	Pr> D	<0.0100
Cramer-von Mises	W-Sq	5861.272	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	27454.25	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive P^{allo})

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	13	2.09E19	1.608E18	1.368E8	<.0001
Residual	79244	1.009E15	1.273E10		
Uncorrected Total	79257	2.09E19			
Corrected Total	79256	2.089E19			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-13137.3	2392.8	-17827.4	-8447.2
sumcasht_1	3.701E-6	3.016E-6	-2.21E-6	9.612E-6
diffcasht_1	1.526E-6	4.678E-7	6.092E-7	2.443E-6
ZPFt	0.7981	0.000173	0.7977	0.7984
dmpat_1	-0.00263	0.000206	-0.00303	-0.00222
MPAt	0.9988	0.000121	0.9985	0.9990
realr	102203	50202.2	3804.7	200600

FAAB	1203.4	838.5	-440.1	2847.0
Public	-2897.8	3292.1	-9350.4	3554.9
ruralare	451.6	1312.2	-2120.3	3023.5
largcity	-700.3	889.1	-2442.9	1042.3
market	-835002	522174	-1858479	188475
marketw	-549567	146666	-837037	-262096

To be able to interpret the coefficient in front of sumcasht_1 and diffcasht_1, we transform these variables to

ddmcasht_1*(3.701E-6+1.526E-6) = ddmcasht_1*(5.227E-6)
dcasht_1*(3.701E-6-1.526E-6) = dcasht_1*(2.17E-6)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	5010	6.32	5010	6.32
1	74247	93.68	79257	100.00

Table 23c: Association of Predicted and Observed P^{allo}

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
PALLOt	745856.28	59114331155	30745988.93	0	5230000000
pPALLOt	878936.47	69661867807	31187554.54	-7048922.26	5227393549

Table 24: Estimation results for *ROT*

Table 24a: Logistic model with complementary log-log function (Gompertz)

Number of Response Levels:	2
Number of Observations:	79257
Model:	binary cloglog

Response Profile

Ordered Value	*Prob*	*Total Frequency*
1	1	1104
2	0	78153

Model Fit Statistics

Criterion	*Intercept Only*	*Intercept and covariates*
AIC	11631.003	11341.213
SC	11640.283	11563.944
-2LogL	11629.003	11293.213

Testing Global Null Hypothesis: BETA=0

Test	*Chi-Square*	*DF*	*Pr > ChiSq*
Likelihood Ratio	335.7896	23	<.0001
Score	1004.9118	23	<.0001
Wald	423.8009	23	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	*Estimate*	*Std. Error*	*Chi-Square*	*Pr > ChiSq*
Intercept	-4.3922	0.2238	385.1497	<.0001
sumallozpft	1.847E-8	2.678E-9	47.5716	<.0001
diffallozpft	-1.3E-8	2.364E-9	30.4768	<.0001
TDEPMAt	1.514E-9	9.14E-10	2.7440	0.0976
TDEPMAt2	-711E-21	2.05E-19	11.9879	0.0005
OIBDt	1.024E-9	2.54E-10	16.2437	<.0001
OIBDt2	-158E-21	1.01E-19	2.4317	0.1189
EDEPBUt	-813E-12	1.458E-8	0.0031	0.9555
EDEPBUt2	4.72E-17	9.86E-17	0.2293	0.6321
OTAt	-118E-12	2.45E-10	0.2317	0.6302
OTAt2	-38E-21	1.01E-20	14.0670	0.0002
TDEPBUt	-2.11E-9	1.69E-8	0.0157	0.9004
TDEPBUt2	-617E-19	1.5E-16	0.1682	0.6817
dourt	2.26E-10	4.32E-10	0.2733	0.6011
TLt	-618E-12	1.356E-9	0.2076	0.6486
FIt	6.53E-10	2.5E-10	6.8017	0.0091
FEt	-107E-12	3.3E-10	0.1042	0.7468
dgnp	5.7E-12	3.16E-12	3.2655	0.0708

FAAB	-0.6675	0.0678	97.0723	<.0001
Public	0.7457	0.1488	25.1265	<.0001
ruralare	-0.2892	0.1102	6.8908	0.0087
largcity	-0.0927	0.0647	2.0573	0.1515
market	-6.1357	32.2844	0.0361	0.8493
marketw	-9.2759	4.9170	3.5589	0.0592

To be able to interpret the coefficient in front of sumallozpft and diffallozpft, we transform these variables to

pallot*(1.847E-8-1.3E-8) = pallot*(5.47E-9)
zpft*(1.847E-8+1.3E-8) = zpft*(3.147E-8)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	54.7	Somers'D	0.277
Percent Discordant	27.0	Gamma	0.339
Percent Tied	18.3	Tau-a	0.008
Pairs	86280912	c	0.638

ROTt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	78153	98.61	78153	98.61
positive	1104	1.39	79257	100.00

pROTt	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
zero	78153	98.61	78153	98.61
positive	1104	1.39	79257	100.00

Table 24b: Estimating the level of positive *ROT*

Influence diagnostics (observations with positive** ROT **)

outl	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	1079	97.74	1079	97.74
1	25	2.26	1104	100.00

leverage	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
0	1055	95.56	1055	95.56
1	49	4.44	1104	100.00

Normality tests (observations with positive *ROT* ***)***

Moments

N	1104	Sum Weights	1104
Mean	0	Sum Observations	0
Std Deviation	14620471.8	Variance	2.13758E14
Skewness	5.14564082	Kurtosis	153.238749
Uncorrected SS	2.35775E17	Corrected SS	2.35775E17
Coeff Variation	.	Std Error Mean	440024.492

Tests for Normality

Test	*Statistic*		*P Value*	
Kolmogorov-Smirnov	D	0.357956	Pr> D	<0.0100
Cramer-von Mises	W-Sq	57.31808	Pr> W-Sq	<0.0050
Anderson-Darling	A-Sq	273.5632	Pr> A-Sq	<0.0050

Schweppe Bounded-Influence Regression using IRLS (observations with positive *ROT* ***)***

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Approx Pr > F*
Regression	17	8.362E15	4.919E14	167.38	<.0001
Residual	1087	3.347E15	3.079E12		
Uncorrected Total	1104	1.171E16			
Corrected Total	1103	1.159E16			

Parameter	*Estimate*	*Approx Std Error*	*Approx 95% Confidence Limits*	
Intercept	-143891	445264	-1017579	729798
sumallozpft	0.000576	0.000235	0.000115	0.00104
diffallozpft	-0.0152	0.0105	-0.0357	0.00539
OIBDt	-0.00475	0.000790	-0.00630	-0.00320
EDEPBUt	0.0684	0.0274	0.0147	0.1221
TDEPBUt	-0.1382	0.0165	-0.1706	-0.1058
dourt	0.0141	0.00384	0.00660	0.0217
TLt	0.0317	0.00839	0.0152	0.0481
FIt	-0.00049	0.000348	-0.00117	0.000193
FEt	0.0611	0.00236	0.0564	0.0657
dgnp	2.93E-6	6.255E-6	-9.34E-6	0.000015
FAAB	-212959	121827	-452007	26088.0
Public	-313273	303278	-908360	281814
ruralare	173824	199326	-217289	564938
largcity	117139	116430	-111319	345596
market	-4.98E7	94929699	-2.361E8	1.3647E8
marketw	-2.043E7	12778744	-4.551E7	4640634

To be able to interpret the coefficient in front of P^{allo} and zpf, we transform these variables to

pallot*(0.000576-0.0152) = pallot*(-0.014624)
zpft*(0.000576+0.0152) = zpft*(0.015776)

Influential observations according to Schweppe-Huber estimation

InflSchweppe	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Frequency*
-1	153	13.86	153	13.86
1	951	86.14	1104	100.00

Table 24c: Association of Predicted and Observed *ROT*

Variable	*Mean*	*Sum*	*Std Dev*	*Minimum*	*Maximum*
ROTt	37410.36	2965032745	3650762.25	-45608295.00	799343359
pROTt	16271.27	1289612195	1187711.57	-15974950.98	243641121

Table 25: Simulation results using current rules, MSEK

Time	*1999*	2000	2001	2002	2003	2004
Assets						
CA	*2032624*	2157244	2626464	2991256	3285191	3576919
MA	*590199*	655068	737831	812583	885209	956082
BU	*526432*	523931	1034464	1255277	1323282	1387825
OFA	*3087234*	3697248	3325529	3441718	3646573	3658999
Total	*6236489*	7033491	7724288	8500834	9140255	9579825
Liabilities						
CL	*1596495*	1481818	2698440	3449106	3842111	4231449
LL	*2543005*	2607697	2729837	2833233	2917740	3000933
ASD	*214129*	228662	235496	240215	244820	248510
OUR	*10862*	10524	10760	10939	10987	11138
SC	*317548*	338181	361574	384821	406226	429033
RR	*350500*	614792	869063	1124540	1377906	1631679
URE	*1052945*	1558227	592813	201132	50403	-293084
PFt	*31417*	62525	56066	55975	56578	57604
PFt-1t	*27497*	31417	62525	56066	55975	56578
PFt-2t	*23364*	27497	31417	62525	56066	55975
PFt-3t	*25432*	23364	27497	31417	62525	56066
PFt-4t	*23350*	25432	23364	27497	31417	62525
PFt-5t	*19941*	23350	25432	23364	27497	31417
Total	*6236485*	7033486	7724284	8500830	9140251	9579823
Income statment						
OIBD	*270101*	280257	405649	416275	409780	417883
EDEPma	*93798*	104559	128439	132789	131337	132571
EDEPbu	*15125*	9084	20342	19701	18767	18602
OIAD	*161177*	166613	256867	263784	259675	266709
FI	*259111*	393226	388071	401018	415111	422190
FE	*178090*	374932	394776	395233	400774	406966
EBA	*242199*	184907	250162	269569	274013	281933
TDEPma-EDEPma	*20271*	14533	6835	4719	4605	3690
OA	*11725*	16510	10358	10814	11308	15190
zPF	*11363*	19941	23350	25432	23364	27497
Pallo	*31418*	62525	56066	55975	56578	57604
EBT	*213598*	144302	220969	245122	247502	263326
TL	*36709*	53841	40623	44448	52085	51485
NI	*176888*	90460	180345	200673	195417	211841
OTA	*-79555*	-50437	-60229	-58915	-52809	-51285
TDEPbu	*14844*	18459	15550	17189	18026	18764
Olt-1t	*190963*	229466	410084	488375	546493	606410
TAX	*33877*	52525	47097	47020	47527	48389
ROT	*1507*	4789	4254	4330	4546	4622
FTAX	*32369*	47735	42842	42689	42980	43766
Olt	*229466*	410084	488375	546493	606410	652192
NBI	*181228*	96566	178126	202432	204522	219559

Flow variabels						
MTDM	*272391*	301038	362240	431438	486150	545979
MCASH	*1156242*	1196292	1947417	1365778	1479387	1734711
I_ma	*139449*	190417	182579	183866	184779	185714
I_bu	*9201*	6583	530875	240514	86772	83146
CMA	*376070*	442780	523219	595880	665887	734543
dCA	*103801*	124619	469219	364792	293934	291728
dOFA	*1444874*	610014	-371718	116188	204855	12425
dCL	*138027*	-114676	1216622	750665	393005	389337
dLL	*1237940*	64691	122139	103395	84506	83193
dOUR	*-9665*	-338	235	179	47	151
dSC	*21499*	20633	23392	23246	21405	22807
dRR	*73096*	264292	254270	255477	253365	253772
dURE	*123155*	505282	-965414	-391681	-150728	-343487
DIV	*0*	211241	896035	386669	466808	656558
CASHFL	*12504*	-725966	678028	-557398	-284784	-157533
SMA	*1199*	20990	-28623	-23675	-19183	-17730
MPA	*30453*	62528	56067	55975	56579	57605
Financial ratio analysis (mean)						
CR	*1,273*	1,456	0,973	0,867	0,855	0,845
DR	*0,681*	0,599	0,720	0,756	0,756	0,772
DER	*2,324*	1,492	2,570	3,094	3,103	3,384
ECR	*0,293*	0,401	0,280	0,244	0,244	0,228
FQ	*-0,355*	-0,292	-0,380	-0,404	-0,397	-0,399
ICR	*2,360*	1,493	1,634	1,682	1,684	1,693
DI	*0,042*	0,089	0,071	0,062	0,058	0,055
ROE	*0,113*	0,097	0,097	0,108	0,100	0,105
ROI	*0,067*	0,080	0,083	0,078	0,074	0,072
EFFTAX	*0,134*	0,258	0,171	0,158	0,157	0,155
RROI	*0,056*	0,072	0,062	0,059	0,061	0,062
ER	*0,012*	0,007	0,022	0,019	0,013	0,010

Table 26: The forecasting accuracy in year 2000

Variable name	(A) The mean of the variables in the sample for year 2000 (using population weights), MSEK*	(B) The StD of the variables in the sample for year 2000 (using population weights), MSEK**	(C) The predicted mean of the variables for year 2000 (using population weights), MSEK	(D) The predicted StD of the variables for year 2000 (using population weights), MSEK	(t) Matched pairs test for testing the hypothesis of equal predicted and sample means***
EDEPMA	0,4471392	15,1496	0,457904	10,43943	-0,279593141
SMA	2,206196	1097,177	0,0919232	50,23465	0,919865648
IMA	2,944805	1096,308	0,8339081	10,14956	0,920048307
EDEPBU	0,0686039	1,820977	0,0397836	1,251001	6,233624567
IBU	0,1610988	29,31948	0,0288315	7,808098	2,08311001
dofa	-2,600112	2373,374	2,67147	87,50439	-1,060656145
dca	0,4167558	578,7162	0,5457548	26,74837	-0,106402568
dll	-3,445979	2104,863	0,2833092	23,35712	-0,846584237
dcl	0,1137608	432,9304	-0,5022092	30,05674	0,678253657
dsc	-0,0347799	48,83318	0,090361	5,457231	-1,216981548
drr	0,4564197	113,3334	1,157433	12,72334	-2,937267227
OIBD	1,398407	81,45908	1,227348	17,73441	0,98049523
FI	1,879987	153,795	1,722081	103,2923	0,407291989
FE	1,150322	55,32959	1,641965	112,4354	-1,874788165
TDEPMA	0,5383478	18,42992	0,5215482	6,93586	0,407669012
ZPF	0,0950842	8,400709	-0,0014825	0,4440188	5,48530102
dour	-0,0150904	2,215839	0,0873331	8,180809	-5,774636057
GC	0,0333825	45,0755	0,0324488	35,62177	0,007766018
OA	0,0675791	42,84226	0,072304	35,46962	-0,040593699
TL	0,2325289	7,468342	0,2357932	37,69122	-0,040595905
OTA	-0,6775173	147,5901	-0,2208842	19,76173	-1,465366983
TDEPBU	0,0863143	3,688432	0,0808414	1,308076	0,668259899
PALLO	0,2242709	32,20007	0,2738195	2,579636	-0,732959816
ROT	0,0018437	4,354507	0,0209764	7,967906	-1,006879944
TAX	0,2121678	6,360432	0,230027	2,168281	-1,269978841
FTAX	0,210324	5,995196	0,2090506	8,168275	0,060055457

* The mean of the variables are calculated by dividing the weighted sum of the variables in the sample by the population size
* The standard deviation of the variables are calculated by using the population size instead of the sample size
*** The matched pair test is performed as follows: t=(A-B)/sqrt(C**2/N+D**2/N) where N=228344 is the population size

Table 27: Simulation results for a proposed tax reduction by 3 per cent, MSEK

Time	***1999***	**2000**	**2001**	**2002**	**2003**	**2004**
FTAX	*32369*	47735	42842	37651	37888	38581
Financial ratio analysis (mean)						
CR	*1,273*	1,456	0,973	0,867	0,855	0,845
DR	*0,681*	0,599	0,720	0,754	0,754	0,770
DER	*2,324*	1,492	2,570	3,064	3,073	3,350
ECR	*0,293*	0,401	0,280	0,246	0,246	0,230
FQ	*-0,355*	-0,292	-0,380	-0,402	-0,395	-0,397
ICR	*2,360*	1,493	1,634	1,682	1,684	1,693
DI	*0,042*	0,089	0,071	0,062	0,058	0,055
ROE	*0,113*	0,097	0,097	0,108	0,099	0,105
ROI	*0,067*	0,080	0,083	0,078	0,074	0,072
EFFTAX	*0,134*	0,258	0,171	0,140	0,138	0,137
RROI	*0,056*	0,072	0,062	0,058	0,059	0,060
ER	*0,012*	0,007	0,022	0,020	0,015	0,012
The cost of the proposed tax rule						
Periodic net cost (FTAXP-FTAXC)	*0*	0	0	-5038	-5092	-5185

Table 28: Simulation results for an alternative macro economic development, MSEK

Time	*1999*	2000	2001	2002	2003	2004
Assets						
CA	*2032624*	2221820	2598926	2902462	3191587	3503591
MA	*590199*	657081	740315	816902	889744	960249
BU	*526432*	536109	772696	866402	929845	1020582
OFA	*3087234*	3595313	3455647	3538734	3642691	3609632
Total	*6236489*	7010323	7567584	8124500	8653867	9094054
Liabilities						
CL	*1596495*	1659088	2441466	2908981	3283529	3755241
LL	*2543005*	2615689	2726189	2829716	2914126	3012416
ASD	*214129*	228078	235643	240910	245399	248805
OUR	*10862*	10523	10747	10944	10988	11139
SC	*317548*	338283	360439	382206	403752	426611
RR	*350500*	614574	867768	1123479	1377735	1631547
URE	*1052945*	1349657	700156	375908	136321	-300250
PFt	*31417*	63367	54096	52611	53025	54027
PFt-1t	*27497*	31417	63367	54096	52611	53025
PFt-2t	*23364*	27497	31417	63367	54096	52611
PFt-3t	*25432*	23364	27497	31417	63367	54096
PFt-4t	*23350*	25432	23364	27497	31417	63367
PFt-5t	*19941*	23350	25432	23364	27497	31417
Total	*6236485*	7010319	7567581	8124496	8653863	9094052
Income statment						
OIBD	*270101*	297243	374587	378472	383985	400172
EDEPma	*93798*	107759	125730	129035	129819	132767
EDEPbu	*15125*	11365	17618	16140	16528	17266
OIAD	*161177*	178119	231238	233296	237637	250137
FI	*259111*	390325	391171	401751	411202	416433
FE	*178090*	376230	391535	386767	392678	400156
EBA	*242199*	192214	230874	248280	256161	266414
TDEPma-EDEPma	*20271*	13949	7564	5268	4488	3406
OA	*11725*	15759	11037	11481	11488	15018
zPF	*11363*	19941	23350	25432	23364	27497
Pallo	*31418*	63367	54096	52611	53025	54027
EBT	*213598*	150599	203601	227314	233500	251496
TL	*36709*	49267	45569	47152	50534	48403
NI	*176888*	101332	158032	180162	182965	203093
OTA	*-79555*	-54804	-54734	-56078	-54105	-54611
TDEPbu	*14844*	17806	16358	17456	17497	17937
Olt-1t	*190963*	229466	405453	495458	561425	623902
TAX	*33877*	53232	45442	44194	44542	45384
ROT	*1507*	4590	4370	4433	4558	4576
FTAX	*32369*	48642	41072	39760	39984	40807
FLOSS	*229466*	405453	495458	561425	623902	670200
NBI	*181228*	101956	162529	187554	193515	210688

Flow variabels						
MTDM	*272391*	306034	363000	431349	488300	549941
MCASH	*1156242*	1200432	1742993	1499123	1574693	1736179
I_ma	*139449*	190147	183248	183804	184119	185218
I_bu	*9201*	21042	254205	109846	79971	108003
CMA	*376070*	444945	525505	599533	669863	738386
dCA	*103801*	189196	377105	303535	289125	312004
dOFA	*1444874*	508078	-139665	83087	103956	-33058
dCL	*138027*	62593	782378	467515	374547	471711
dLL	*1237940*	72683	110499	103526	84410	98289
dOUR	*-9665*	-338	223	197	43	151
dSC	*21499*	20734	22156	21766	21546	22858
dRR	*73096*	264073	253194	255711	254255	253812
dURE	*123155*	296712	-649501	-324247	-239587	-436572
DIV	*0*	242270	737071	496321	541359	659371
CASHFL	*12504*	-511786	316566	-480980	-317473	-147910
SMA	*1199*	15507	-25716	-21817	-18541	-18054
MPA	*30453*	63371	54097	52611	53026	54028
Financial ratio analysis (mean)						
CR	*1,273*	1,339	1,064	0,998	0,972	0,933
DR	*0,681*	0,627	0,700	0,724	0,734	0,762
DER	*2,324*	1,682	2,337	2,620	2,754	3,196
ECR	*0,293*	0,373	0,300	0,276	0,266	0,238
FQ	*-0,355*	-0,310	-0,357	-0,366	-0,365	-0,376
ICR	*2,360*	1,511	1,590	1,642	1,652	1,666
DI	*0,042*	0,086	0,074	0,066	0,062	0,058
ROE	*0,113*	0,082	0,082	0,090	0,089	0,101
ROI	*0,067*	0,081	0,082	0,078	0,075	0,073
EFFTAX	*0,134*	0,253	0,178	0,160	0,156	0,153
RROI	*0,056*	0,072	0,063	0,062	0,062	0,061
ER	*0,012*	0,009	0,019	0,016	0,013	0,012
The cost of the proposed tax rule						
Periodic net cost (FTAXP-FTAXC)	*0*	907	-1770	-2929	-2996	-2959

References

Abel A. (1982); ”Dynamic Effects of Permanent and Temporary Tax Policies in a q-Theory Model of Investment”; Journal of Monetary Economics, May; 353-374.

Abel B. A. and O. J. Blanchard (1986); ”The Present Value of Profits and Cyclical Movements in Investments”; Econometrica; Vol. 54; No. 2; March; 249-273.

Altman I. E. (1968); ” Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bancruptcy”; *The Journal of Finance*; Vol. XXIII, No. 4; 589-609.

Altman I. E. (1984); ” A Further Empirical Investigation of the Bankruptcy Cost Question”; *Journal of Finance*; 1067-1089.

Amemiya T. (1985); Advansed Econometrics; Basil Blackwell Ltd..

Andersson K. and E. Norrman (1987); ”Capital Taxation and Neutrality - A Study of Tax Wedges With Special References to Sweden”, *Lund Economic Studies number 41*.

Apel, M. and Södersten J., ”Personal Taxation and Investments Incentives in a Small Open Economy”, Working Paper 1996:5, Department of Economics, Uppsala University.

Atkinson A. and J. Stiglitz (1980); ”Lectures on Public Economics”; Lecture five.

Auerbach J. A (1983); ”Taxation, Corporate Financial Policy and the Cost of Capital”; Journal of Economic Literature; Vol XXI; Page 905-940.

Auerbach J. A. (1984); ”Taxes, Firms Financial Policy and the Cost of Capital: An Empirical Analysis”; *Journal of Public Economics*; 27-57.

Auerbach J. A. and K. Hasset (1992); ”Tax Policy and Business Fixed Investment in the United States”; Journal of Public Finance 47; North-Holland; 141-170.

Beaton, A. E. and Tukey, J. W. (1974) “The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data”, *Technometrics*, vol. 16, 146-185.

Benninga S. (1989); Numerical Techniques in Finance; MIT Press; Cambridge; MA

Bergström V. and J. Södersten (1981:1); ”Double Taxation and Corporate Capital Cost”; in G. Eliasson and J. Södersten; eds.; Business Taxation, Finance and Firm Behavior-Proceedings of a Symposium at IUI, Stockholm, August 28-29,1978; IUI Conference Reports 1981:1; Stockholm; 163-190.

Bergström V. and J. Södersten (1981:1); ”Inflation, Taxation and Capital Cost”; in G. Eliassson and J. Södersten; eds.; Business Taxation, Finance and Firm Behavior- Proceedings of a Symposium at IUI, Stockholm, August 28-29,1978; IUI Conference Reports 1981:1; Stockholm; 223-266.

Bhattacharya, S. (1979), ”Imperfect Information ,Dividend Policy, and the ‘Bird in the Hand’ Fallacy”, *Bell Journal of Economics*, Spring, 259-270.

Black, F. and Scholes, M. (1974), ”The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns”, *Journal of Financial Economics*, Vol. 1, 1-22.

Black F. (1990); ”The Dividend Puzzle”; Published in Clifford W. and JR. Smith (1990)”The Modern Theory of Corporate Finance”, Second Edition; 215-220.

Blundell R., S. Bond, and C. Meghir (1996); ”Econometric Models of Company Investment”; in Matyas L. and P. Sevestre; Eds.; The Econometrics of Panel Data. A Handbook of the Theory with Applications; Second Addition; Advanced Studies in Theoretical and Applied Economics; Vol. 33; Dordrecht; Boston and London: Kluwer Academic; 685-710.

Boadway R. (1979); ”Investment Incentives, Corporate Taxation, and Efficiency in the Allocation of Capital”; Economic Journal 99, 470-481.

Boadway R. (1980); ”Corporate Taxation and Investment: A Synthesis of the Neo-Classical Theory”; Canadian Journal of Economics 13, 250-267.

Boadway W. R. and N. Bruce (1979); ”Depreciation and Interest Deductions and Effect of the Corporation Income Tax on Investment”; Journal of Public Economics 11; 93-105.

Boadway, R. and Bruce, N. (1992), ”Problems with Integrating Corporate and Personal Income Taxes in an Open Economy”, *Journal of Public Economics*, Vol. 48(1), June, 39-66.

Bond S. and C. Meghir (1994); ”Dynamic Investment Models and the Firm´s Financial Policy”; Review of Economic Studies 61; 197-222.

Brickely A. J. (1982); ”Shareholders Wealth, Information Signaling and the Specially Designated Dividend: An Empirical Study”; Journal of Financial Economics 12; 187-209.

Caballero J. R., E. A. R. Eduardo, and J. C. Haltiwanger (1995); Plant-level Adjustment and Aggregate Investment Dynamics”; Brookings Papers on Economic Activity 2; 1-54.

Carroll, R. J. and Ruppert, D. (1987) “Diagnostics and Robust Estimation when Transforming the Dependent Variable and the Response”, *Technometrics*, vol. 29, 287-299

Chrinko S. R. (1993); ” Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications”; Journal of Economic Litterature; Vol. XXXI; December; 1875-1911.

Clarke S. and A. M. Tobias (1995); ”Complexity in Corporate Modelling: A Review”; Business History; Vol. 37; No. 2; 17-44.

Cummins G. J., H. S. Harris, and K. A. Hasset (1994); ”Accounting Standards, Information Flow, and Firm Investment Behavior”; National Bureau of Economic Research, Inc.; Working Paper No. 4685.

Cummins G. J., K. A. Hasset and R. G. Hubbard (1994); ”A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments”; Brookings Papers on Economic Activity 2; 1-74.

Dorfman R (1969); ”An Economic Interpretation of Optimal Control Theory”; *American Economic Review*.

Downes H. D. (1973); ”Financial Policy Models: Theory and Practice”; Journal of Financial and Quantitative Analysis; Vol. 8; 691-709.

Eades M. K. (1982); ”Empirical Evidence on Dividends as a Signal of Firm Value”; *Journal of Financial and Quantitative Analysis*; Vol. XVII, No. 4, November; 471-500.

Easterbrook F. H. (1984); ”Two Agency-Cost Explanation of Dividends”, *American Economic Review 74*; 650-659.

Fomby B. T., R. Carter, and S. R. Johnson (1984); Advansed Econometric Methods; Spriger-Verlag.

Forrester J. W. (1961); Industrial Dynamics; MIT Press; Cambridge; MA

Forsling G. (1998); Utilization of Tax Allowances and Corporate Borrowing; Economic Studies 37; Department of Economics; Uppsala University.

Francis C. J. (1983); ”Financial Planning and Forecasting Models: A Review”; Journal of Economic Business; Vol. 35; 285-300.

Francis C. J. and D. R. Rowell (1978); ”A Simultaneous Equation Model of the Firm for Financial Analysis and Planning”; Financial Management; Vol. 7; 29-44.

Gentry A. J. (1979); ”Simulating the Strategic Financial Planning Process”; Eurpoean Journal of Operational Research; Vol. 3; 441-449.

Gentry A. J. and P. A. Stephen (1973); ”Simulating an EPS Growth Model”; Financial Management; Vol. 2; 68-75.

Gilchrist S. and C. Himmelberg (1998); ”Investment, Fundamentals and Finance”; National Bureau of Economic Research; Cambridge; MA.

Gordon, M. (1959), ”Dividends, Earnings and Stock Prices”, *Review of Economics and Statistics*, May, 99-105.

Gordon, R. and J. MacKie-Mason (1992), ”Tax Distortions to the Choice Of Organizational Form”, *National Bureau of Economic Research Working Paper No. 4227.*

Gordon, R. and J. MacKie-Mason (1993); ”How Much Do Taxes Discourage Incorporation?” National Bureau of Economic Research Working Paper No. 4227.

Graham, B. and Dodd, D. L. (1951), ”Security Analysis”, McGraw-Hill, New-York.

Greene H. W. (1993); Econometric Analysis; Second Edition; Macmillan Publishing Company.

Griffiths E. W., R. C. Hill, and George G. Judge (1993); Learning and Practicing Econometrics; John Wiley & Sons, Inc.; Chapter 17.

Gustafsson L., H. Lanshammar and B. Sandbald (1982); System och Modell: En Introduktion till Systemanalys; Student Litteratur

Hall R. and D. Jorgenson (1962); ”Tax Policy and Investment Behavior”; American Economic Review.

Hamilton, L. C. (1991) *Regression with Graphics*. Pacific Grove, CA: Brooks/Cole Publishing Company.

Hampel, F. R. (1978) ”Optimally Bounding the Gross-Error Sensitivity and the Influence of Position in Factor Space”, *In Proceedings of the ASA Statistical Computing Section*, American Statistical Association, Washington DC

Hampel, F. R. (1985) “The Breakdown Points of the Mean Combined With Some Rejection Rules”, *Technometrics*, vol. 27, 95-107

Handschin, E., Kohlas, J., Fiechter, A. and Schweppe, F. (1975) ”Bad Data Analysis for Power System State Estimation”, *IEEE Transaction in Power Apparatus and Systems*, PAS-94 (**2**), 329-337.

Hansen S. (1999); Essays on Finance, Taxation and Corporate Investment; Economic Studies 47; Department of Economics; Uppsala University; Essays 1 and 2.

Harris M. and A. Raviv (1991); ”The Theory of Finance”; Vol. XLVI; No. 1; 297-355.

Hayashi F. (1982); ”Tobin’s Marginal q and Average q: A Neo-Classical Interpretation”; Econometrica; January.

Haugen A. R. and L. W. Senbet (1986); ”Corporate Finance and Taxes: A Review”; Fiancial Management; 5-21.

Huber, P. J. (1964) “Robust Estimation of a Location Parameter”, *Annals of Mathematical Statistics*, vol. 35, 73-101.

Huber, P. (1983) ”Minimax Aspects of Bounded-Influence Regression”, *Journal of the American Statistical Association*, vol. 78 (381), 66-80 (with discussion).

Jensen M. C. and W. H. Meckling (1976); ”Theory of the Firm: Management Behavior, Agency Costs and Ownership Structure”; *Journal of Financial Economics 3*; 305-360.

Jensen M. C. (1986); ”Agency Costs of Free Cash Flow, Corporate Finance, and Takeover”; *American Economic Review 76*, 323-329

Kamien I. M. and N. L. Schwartz (1991); Dynamic Optimization- The Calculus of Variations and Optimal Control in Economics and Management; Second Edition; Part II; Sections 1-11.

Kanniainen V. and J Södersten (1994); ”Costs of Monitoring and Corporate Taxation” Journal of Public Economics 55, 307-321.

Kanniainen V. and J Södersten (1995); ”The Importance of Reporting Conventions for the Theory of Corporate Taxation”; Journal of Public Economics; 57; 417-430.

Karsten G. S. (1990); Business Forecasting and Ecoinomic Cycles; Lanham; New York; Chapter 6.

King A. M. (1974); ”Taxation and the Cost of Capital”; Review of Economic Studies 41; 21-35.

King A. M. (1975); ”Taxation, Corporate Financial Policy, and the Cost of Capital- A Comment”; *Journal of Public Economics*; 4; 271-279.

Klevmarken, A. N. (2001) “Micro Simulation – A Tool for Economic Analysis”, Working Paper No. 13, Department of Economics, Uppsala University.

Kopcke W. R. (1985); ”The Determinants of Investment Spending”; New England Economic Review; 19-35.

Krasker, W. S. and Welsch, R. E. (1982) “Efficient Bounded-Influence Regression Estimation”, *Journal of the American Statistical Association*, vol. 77, 595-604.

Krasker, W. S., Kuh, E. and Welsch, R. E. (1983) ”Estimation for Dirty Data and Flawed Models”, In Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics Vol. 1*, Chap. 11, Amsterdam: North Holland

Krause A. and R. Litzenberger (1973); ”A State Preference model of Optimal Financial Leverage”; Journal of Finance; Vol. 28; 911-922.

Kumar R. and P. Vrat (1989); ”Using Computer Models in Corporate Planning”; Long Range Planning, Vol. 22; No. 2; 114-120.

Li, G. (1985) ‘Robust Regression’. In (D. C. Hoaglin, F. Mollester and J. W. Tukey, eds.) *Exploring Data Tables, Trends and Shapes*, 281-340, New York: John Wiley & Sons.

Lyneis M. J. (1975); ”Designing Financial Policies to Deal with Limited Financial Resources”; Financial Management; Vol. 4; 3-24.

Miller M. (1977); ”Debt and Taxes” *Journal of Finance*; 261-275.

Miller, M and Scholes, M, (1978), ”Dividends and Taxes”, *Journal of Financial Economics*, Vol. 6, 333-364.

Miller, M and Scholes, M, (1982), ”Dividends and Taxes, Some Empirical Evidence”, *Journal of Political Economy*, Vol. 90, 1118-1141.

Modigliani F. and M. Miller (1958); ”The Cost of Capital, Corporation Finance, and the Theory of Investment”; *American Economic Review 48*; 261-297.

Miller, M and F Modigliani (1961), ”Dividen Policy, Growth, and the Valuation of Shares”, *Journal of Business*, Vol. 34, 411-433.

Modigliani F. and M. Miller (1963); ”Corporate Taxes and the Cost of Capital: A Correction” *American Economic Review*; 433-443.”

Myers S. C.(1984); ” The capital structure puzzle” *Journal of Finance*; 575-592.

Osterberg P. W. (1989); ”Tobin’s q, Investment, and The Endogenous Adjustment of Financial Structure”; Journal of Public Economics; 40; 243-318.

Poterba, J. M. and Summers, L. H. (1983), ”Dividend Taxes, Corporate Investment, and ´Q´”, Journal of Public Economics 22; North-Holland; 135-167.

Poterba, J. M. and Summers, L. H. (1985), ”The Economic Effects of Dividend Taxation”, in E. Altman and M. Subrahmanyam (eds.), *Recent Advances in Corporate Finance*, R. D. Irwin, 229-283.

Ross, S. A. (1977), ”The Determination of Financial Structure: The Incentive Signalling Approach”, *Bell Journal of Economics*, Spring, 23-40.

Sandmo A. (1974); ”Investment Incentives and the Corporate Income Tax”; *Journal of Political Economy 82*; 287-302.

Shahnazarian H. (1996); Three Essays on Corporate Taxation; Economic Studies 24; Department of Economics; Uppsala University.

Shahnazarian H. and Bo Stoltz (1999); ”En översikt av olika synsätt på dubbelbeskattningens ekonomiska effekter på den reala ekonomin”; Mimeo; Ministry of Finance; Sweden.

Sinn H.-W. (1987); Capital Income Taxation and Resource Allocation; North-Holland.

Sinn H.-W. (1990); ”Taxation and The Cost of Capital: The ‘Old View’, The ‘New View’ and ‘Another View’; *National Bureau of Economic Research Paper 2501*.

SOU (1998:136); ”Redovisning och aktiekapital i euro- Slutbetänkande av Utredningen om aktiekapital och redovisning i utländsk valuta”.

Stiglitz J. E. (1973); ”Taxation, Corporate Financial Policy, and the Cost of Capital”; *Journal of Public Economics 2*; 1-34.

Stiglitz E. J. (1985); ” Credit Markets and the Control of Capital”; *Journal of Money, Credit, and Banking*; Vol. 17.

Stoltz, B. (1997), ”Dividends, Taxes and Share Prices: Evidence using Swedish Data 1983-1996” i *Essays on Portfolio Behavior and Asset Pricing*, Economic Studies 30, Department of Economics, Uppsala University, 93-122.

Street, J. O., Carroll, R. J. and Ruppert, D. (1988) “A Note on Computing Robust Regression Estimates via Iteratively Reweighted Least Squares”, *The American Statistician*, vol. 42, 152-154.

Summers L. (1981); ”Taxation and Corporate Investment: A q Theory Approach”; Brookings Papers on Economic Activity 1.

Swedish Standard Industrial Classification 1992; Reports on Statistical Co-ordination (1992:4)

Södersten J. (1982); ”Accelerated Depreciation and the Cost of Capital”; *Scand. Journal of Economics 84 (1);* 111-115.

Tongeren F. W. van (1995); Microsimulation Modeling of the Corporate Firm- Exploring Micro-Macro Economic Relations; Springer

Warren M. J. (1974); ”An Operational Model for Security Analysis and Valuation”; Journal of Financial and Quantitative Analysis; Vol. 9; 395-422.

Warren M. J. and J. P. Shelton (1973); ”A Simultaneous Equation Approach To Financial Planning”; Journal of Finance; Vol. 28; 1123-1142

Warner J.; ”Bankruptcy Costs: Some Evidence” *Journal of Finance*; May 1977; 337-347.

Weston F. J. (1974); ”New Themes in Finance”; Journal of Finance; Vol. 29; 237-243.

Zodrow, G. R. (1991), ”On the ‘Traditional’ and ‘New’ Views of Dividend Taxation”, *National Tax Journal*, Vol. XLIV, No. 4, 497-509.

Previous titles in this series:

1. Effective Exchange Rate Index by *Thomas Franzén, Aleksander Markowski* and *Irma Rosenberg*, Stockholm 1980
2. Sweden's Invisible Foreign Trade in the 1970s by *Kerstin Joelson* and *Nils Eric Persson*, Stockholm 1984
3. Valutahantering i ett antal svenska företag av *Malin Hedman*, Stockholm 1986
4. The Forward Exchange Market by *Thomas Franzén*, Stockholm 1987
5. Distributional Effects of High Interest Rates by *Jonas Agell* and *Mats Dillén*, Stockholm 1988
6. Minimac – an Econometric Model of the Swedish Economy from a Central Bank's Perspective by *Aleksander Markowski*, Stockholm 1988
7. Penningpolitikens mål och medel 1955-1967 av *Kurt Eklöf*, Stockholm 1990
8. Tre valutakriser 1967-1977 av *Kurt Eklöf*, Stockholm 1990
9. EEA and the Financial Services Sectors in Sweden by *Emil Ems, Erik Blomberg* and *Eva Blixt*, Stockholm 1993
10. Financial Integration in Western Europe – Structural and Regulatory Consequences by *Emil Ems* (Editor), Stockholm 1994

Sveriges Riksbank
Visiting address: Brunkebergs torg 11
Mail address: se-103 37 Stockholm

Website: www.riksbank.se
Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31
E-mail: registratorn@riksbank.se

Riksbank Printing Office Stockholm 2004

ISSN 0349-3253